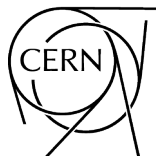


## **2016 European School of High-Energy Physics**

Skeikampen, Norway  
15 – 28 June 2016

Editors: M. Mulders  
G. Zanderighi



CERN Yellow Reports: School Proceedings  
Published by CERN, CH-1211 Geneva 23, Switzerland

ISBN 978-92-9083-476-2 (paperback)

ISBN 978-92-9083-477-9 (PDF)

ISSN 2519-8041 (Print)


ISSN 2519-805X (Online)

DOI <https://doi.org/10.23730/CYRSP-2017-005>

Accepted for publication by the CERN Report Editorial Board (CREB) on 14 December 2017

Available online at <http://publishing.cern.ch/> and <http://cds.cern.ch/>

Copyright © CERN, 2017

 Creative Commons Attribution 4.0

Knowledge transfer is an integral part of CERN's mission.

CERN publishes this volume Open Access under the Creative Commons Attribution 4.0 license (<http://creativecommons.org/licenses/by/4.0/>) in order to permit its wide dissemination and use.

The submission of a contribution to a CERN Yellow Report series shall be deemed to constitute the contributor's agreement to this copyright and license statement. Contributors are requested to obtain any clearances that may be necessary for this purpose.

This volume is indexed in: CERN Document Server (CDS), INSPIRE, Scopus.

This volume should be cited as:

Proceedings of the 2016 European School of High-Energy Physics, Skeikampen, Norway, 15 – 28 June 2016, edited by M. Mulders and G. Zanderighi, CERN Yellow Reports: School Proceedings, Vol. 5/2017, CERN-2017-009-SP (CERN, Geneva, 2017), <https://doi.org/10.23730/CYRSP-2017-005>

A contribution in this volume should be cited as:

[Author name(s)], in Proceedings of the 2016 European School of High-Energy Physics, Skeikampen, Norway, 15 – 28 June 2016, edited by M. Mulders and G. Zanderighi, CERN Yellow Reports: School Proceedings, Vol. 5/2017, CERN-2017-009-SP (CERN, Geneva, 2017), pp. [first page]–[last page], <https://doi.org/10.23730/CYRSP-2017-005>. [first page]



## **Abstract**

The European School of High-Energy Physics is intended to give young physicists an introduction to the theoretical aspects of recent advances in elementary particle physics. These proceedings contain lecture notes on the theory of the Weak interaction and Higgs physics, flavour physics and CP violation, neutrinos, theories beyond the Standard Model, physics at the LHC Run-2 and beyond, practical statistics for high-energy physicists, and cosmology and dark matter.



## Preface

The twenty-fourth event in the series of the European School of High-Energy Physics took place in Skeikampen, Norway, from 15 to 28 June 2016. It was organized jointly by CERN, Geneva, Switzerland, and JINR, Dubna, Russia, with support from the universities of Bergen and Oslo. The local organization team was chaired by Prof. Heidi Sandaker. The other members of the local committee were: Trygve Buanes, Gerald Eigen, Tomas Gonzalo, Eirik Gramstad, Børge Hovden, Abram Krislock, Anna Lipniacka, Bertrand Martin dit Latour, Farid Ould-Saada, Are Raklev and Alex Read.

A total of 98 students of 35 different nationalities attended the school, mainly from institutes in member states of CERN and/or JINR, but also some from other regions. The participants were generally students in experimental High-Energy Physics in the final years of work towards their PhDs.

The School was hosted at the Thon Hotel Skeikampen, about 40 km to the north of Lillehammer. According to the tradition of the school, the students shared twin rooms mixing participants of different nationalities.

A total of 30 lectures were complemented by daily discussion sessions led by six discussion leaders. The students displayed their own research work in the form of posters in an evening session in the first week, and the posters stayed on display until the end of the School. The full scientific programme was arranged in the on-site conference facilities.

The School also included an element of outreach training, complementing the main scientific programme. This consisted of a two-part course from the Inside Edge media training company. Additionally, students had the opportunity to act out radio interviews under realistic conditions based on a hypothetical scenario.

The students from each discussion group subsequently carried out a collaborative project, preparing a talk on a physics-related topic at a level appropriate for a general audience. The talks were given by student representatives of each group in an evening session in the second week of the School. A jury, chaired by Prof. Egil Lillestøl (a winner of the Norwegian Research Council award for communication of science), judged the presentations; other members of the jury were Gabriela Barenboim (lecturer at the School), Kate Ross (Schools Administrator), and Inga Hanne Dokka (a winner of the Norwegian Physical Society award excellence in teaching). We are very grateful to all of these people for their help.

Our thanks go to the local-organization team and, in particular, to Heidi Sandaker, for all of their work and assistance in preparing the School, on both scientific and practical matters, and for their presence throughout the event. Our thanks also go to the efficient and friendly hotel management and staff who assisted the School organizers and the participants in many ways.

Very great thanks are due to the lecturers and discussion leaders for their active participation in the School and for making the scientific programme so stimulating. The students, who in turn manifested their good spirits during two intense weeks, appreciated listening to and discussing with the teaching staff of world renown.

We would like to express our strong appreciation to Professor Fabiola Gianotti, Director General of CERN, and Professor Victor Matveev, Director of JINR, for their lectures on the scientific programmes of the two organizations and for discussing with the School participants.

In addition to the rich academic programme, the participants enjoyed numerous sports, leisure and cultural activities in and around Skeikampen. There was a half-day excursion to Lillehammer, with its Olympic museum and Maihaugen open-air museum on the first Saturday afternoon. Then, during the full-day excursion to the impressive Jotunheimen Mountains, participants had the option to hike over the Besseggen mountain, to follow an alternative route along the lakeshore, or to just relax and enjoy the spectacular scenery. Sports and leisure activities in and around the hotel, as well as the excursions, provided an excellent environment for informal interactions between staff and students.

We are very grateful to Kate Ross and Tatyana Donskova for their untiring efforts in the lengthy preparations for and the day-to-day operation of the School. Their continuous care of the participants and their needs during the School was highly appreciated.

The success of the School was to a large extent due to the students themselves. Their poster session was very well prepared and highly appreciated, their group projects were a huge success, and throughout the School they participated actively during the lectures, in the discussion sessions and in the different activities and excursions.

Nick Ellis  
(On behalf of the Organizing Committee)









## People in the photograph

|    |                             |    |                         |     |                      |
|----|-----------------------------|----|-------------------------|-----|----------------------|
| 1  | Artem Maevskiy              | 36 | Moritz Habermehl        | 71  | Nicoletta Belloli    |
| 2  | Alfredo Urbano              | 37 | Baptiste Abeloos        | 72  | Renat Sadykov        |
| 3  | Igor Tkachev                | 38 | Gionata Luisoni         | 73  | Viacheslav Kaminskiy |
| 4  | Denis Korablev              | 39 | Nedaa-Alexandra Asbah   | 74  | Suzanne Klaver       |
| 5  | Tatiana Antoshkina          | 40 | Angela Burger           | 75  | Callum Kilby         |
| 6  | Aliaksandr Antoshkin        | 41 | Carla Marin Benito      | 76  | Thor Taylor          |
| 7  | Tatiana Ovsiannikova        | 42 | Andrea Mogini           | 77  | Mai El Sawy          |
| 8  | Kate Ross                   | 43 | Arthur Lesage           | 78  | Margherita Spalla    |
| 9  | Konstantin Treskov          | 44 | Noemi Calace            | 79  | Roberto Franceschini |
| 10 | Daniel Gonzalez             | 45 | Katarina Gajdosova      | 80  | Radim Slovak         |
| 11 | Francesco Cirotto           | 46 | Luigi Longo             | 81  | Giulio Dujany        |
| 12 | Marek Walczak               | 47 | Giorgia Rauco           | 82  | Juraj Smiesko        |
| 13 | Myriam Schoenenberger       | 48 | Viktar Kireyeu          | 83  | Jaime Norman         |
| 14 | Xavier Coubez               | 49 | David Hohn              | 84  | Ismet Siral          |
| 15 | Steffen Maeland             | 50 | Marilea Reale           | 85  | Simon Corrodi        |
| 16 | Ruiqi Zhang                 | 51 | Giuseppe Lerner         | 86  | John Anders          |
| 17 | Alberto Escalante del Valle | 52 | Chiara Rizzi            | 87  | Michael Pitt         |
| 18 | Lucien Lo                   | 53 | Gabriela Barenboim      | 88  | Karlis Dreimanis     |
| 19 | Monika Blanke               | 54 | Hadi Hassan             | 89  | Luke Vinton          |
| 20 | Marton Bartok               | 55 | Emanuele Michielin      | 90  | Miroslva Simko       |
| 21 | Yi-Ting Duh                 | 56 | Ksenia Gasnikova        | 91  | Ayse Bat             |
| 22 | Dominik Krauss              | 57 | Mariana Shopova         | 92  | Alessandra Baas      |
| 23 | Radoslav Marchevski         | 58 | Alessio Boletti         | 93  | Rafal Sikora         |
| 24 | Vit Kucera                  | 59 | Julian Garcia Pardinias | 94  | Boris Teyssier       |
| 25 | Marco A. Harrendorf         | 60 | Tao Huang               | 95  | Clement Camincher    |
| 26 | Abram Krislock              | 61 | Lukasz Fulek            | 96  | Mareike Meyer        |
| 27 | Alina Vishneva              | 62 | Oleksiy Fedorchuk       | 97  | Eloi Le Quilleuc     |
| 28 | Pirmin Berger               | 63 | Maria Vieites Diaz      | 98  | Mariya Ilieva        |
| 29 | Denys Denysiuk              | 64 | Danijela Bogavac        | 99  | Borge Hovden         |
| 30 | Meera Vieira Machado        | 65 | Andrey Sapronov         | 100 | Rizki Syarif         |
| 31 | Nicolas Lurkin              | 66 | Igor Shandrov           | 101 | Svende Braun         |
| 32 | Dirk Sammel                 | 67 | Andy Wharton            | 102 | Nick Ellis           |
| 33 | Tatyana Donskova            | 68 | Cora Fischer            | 103 | Martijn Mulders      |
| 34 | Ryne Carbone                | 69 | Rhys Owen               | 104 | Trygve Buanes        |
| 35 | Aurelie Bonhomme            | 70 | Alexander Olshevskiy    |     |                      |



# Photographs (montage)



## The 2016 European School of High-Energy Physics

Skeikampen, Norway 15 – 28 June 2016









# Contents

|  |     |
|--|-----|
| Preface  |     |
| <i>N. Ellis</i> .....                            | v   |
| Photograph of participants .....                 | vii |
| Photographs (montage) .....                      | x   |
| Lectures on the Theory of the Weak Interaction   |     |
| <i>M.E. Peskin</i> .....                         | 1   |
| Introduction to Flavour Physics and CP Violation |     |
| <i>M. Blanke</i> .....                           | 71  |
| Neutrinos: Fast & Curious                        |     |
| <i>G. Barenboim</i> .....                        | 101 |
| Beyond the Standard Model                        |     |
| <i>B.C. Allanach</i> .....                       | 123 |
| Physics at the LHC Run-2 and Beyond              |     |
| <i>A. Hoecker</i> .....                          | 153 |
| Practical Statistics for High Energy Physicists  |     |
| <i>L. Lista</i> .....                            | 213 |
| Cosmology and Dark Matter                        |     |
| <i>I. Tkachev</i> .....                          | 259 |
| Organizing Committee .....                       | 295 |
| Local Organizing Committee .....                 | 295 |
| List of Lecturers .....                          | 295 |
| List of Discussion Leaders .....                 | 295 |
| List of Students .....                           | 296 |
| List of Posters .....                            | 297 |



# Lectures on the Theory of the Weak Interaction

*M. E. Peskin*

SLAC, Stanford University, Menlo Park, CA, USA

## Abstract

I review aspects of the theory of the weak interaction in a set of lectures originally presented at the 2016 CERN-JINR European School of Particle Physics. The topics discussed are: (1) the experimental basis of the  $V-A$  structure of the weak interaction; (2) precision electroweak measurements at the  $Z$  resonance; (3) the Goldstone Boson Equivalence Theorem; (4) the Standard Model theory of the Higgs boson; (5) the future program of precision study of the Higgs boson.

## Keywords

Lectures;  $V-A$ ;  $Z$  boson;  $W$  boson; Higgs boson.

## 1 Introduction

Today, all eyes in particle physics are on the Higgs boson. This particle has been central to the structure of our theory of weak interactions ever since Weinberg and Salam first wrote down what we now call the Standard Model of this interaction in 1967 [1, 2]. As our understanding of particle physics developed over the following decades, what lagged behind was our knowledge of this particle and its interactions. Increasingly, the remaining mysteries of particle physics became centered on this particle and the Higgs field of which it is a part.

In 2012, the Higgs boson was finally discovered by the ATLAS and CMS experiments at the LHC [3, 4]. Finally, we have the opportunity to study this particle in detail and to learn some of its secrets by direct observation. Many students at this summer school, and many others around the world, are involved in this endeavor. So it is worthwhile to review the theory of the Higgs boson and the broader theory of weak interactions in which it is embedded. That is the purpose of these lectures.

To learn where we are going, it is important to understand thoroughly where we have been. For this reason, the first half of this lecture series is devoted to historical topics. In Section 2, I review the basic formulae of the Standard Model and set up my notation. An important property of the Standard Model is that, unexpectedly at first sight, charge-changing weak interactions couple only to left-handed-polarized fermions. This structure, called the  $V-A$  interaction, is the reason that we need the Higgs field in the first place. In Section 3, I review the most convincing experimental tests of  $V-A$ . Section 4 reviews the precision measurements on the weak interaction made possible by the  $e^+e^-$  experiments of the 1990's at the  $Z$  resonance. These experiments confirmed the basic structure of the Standard Model and made the Higgs field a necessity.

One aspect of the Higgs field that is subtle and difficult to understand but very powerful its application is the influence of the Higgs field on the high-energy dynamics of vector bosons  $W$  and  $Z$ . Section 5 is devoted to this topic. The physics of  $W$  and  $Z$  bosons at high energy is full of seemingly mysterious enhancements and cancellations. The rule that explains these is the connection to the Higgs field through a result called the Goldstone Boson Equivalence Theorem, first enunciated by Cornwall, Levin, and Tiktopoulos and Vayonakis [5, 6]. In Section 5, I explain this theorem and illustrate the way it controls the energy-dependence of a number of interesting high-energy processes.

In Sections 6 and 7, I turn to the study of the Higgs boson itself. Section 6 is devoted to the Standard Model theory of the Higgs boson. I will review the general properties of the Higgs boson and explain in some details its expected pattern of decay models. Section 7 is devoted to the remaining

mysteries of the Higgs boson and the possibility of their elucidation through a future program of precision measurements.

## 2 Formalism of the Standard Model

To begin, I write the formalism of the Standard Model (SM) in a form convenient for the analysis given these lectures. The formalism of the SM is standard material for students of particle physics, so I assume that you have seen this before. It is explained more carefully in many textbooks (for example, [7, 8]).

### 2.1 Gauge boson interactions

The SM is a gauge theory based on the symmetry group  $SU(2) \times U(1)$ . A gauge theory includes interactions mediated by vector bosons, one boson for each generator of the gauge symmetry  $G$ . The coupling of spin 0 and spin  $\frac{1}{2}$  particles to these vector bosons is highly restricted by the requirements of gauge symmetry. The interactions of these fermions and scalars with one another is much less restricted, subject only to the constraints of the symmetry  $G$  as a global symmetry. Thus, the theory of fermions and vector bosons is extremely tight, while the introduction of a scalar field such as the Higgs field introduces a large number of new and somewhat uncontrolled interaction terms.

The SM contains 4 vector bosons corresponding to the 3 generators of  $SU(2)$  and 1 generator of  $U(1)$ . I will call these

$$A_\mu^a, \quad B_\mu, \quad (1)$$

with  $a = 1, 2, 3$ . These couple to fermion and scalar fields only through the replacement of the derivatives by covariant derivative

$$\partial_\mu \rightarrow D_\mu = (\partial_\mu - igA_\mu^a t^a), \quad (2)$$

where  $t^a$  is the generator of  $G$  in the representation to which the fermions or scalars are assigned. For the SM, fermion and scalar fields are assigned  $SU(2)$ , or weak isospin, quantum numbers 0 or  $\frac{1}{2}$  and a  $U(1)$ , or hypercharge, quantum number  $Y$ . The covariant derivative is then written more explicitly as

$$D_\mu = \partial_\mu - igA_\mu^a t^a - ig' B_\mu Y, \quad (3)$$

with

$$t^a = 0 \text{ for } I = 0, \quad t^a = \frac{\sigma^a}{2} \text{ for } I = \frac{1}{2}. \quad (4)$$

This formalism makes precise predictions for the coupling of the weak interaction vector bosons to quarks and leptons, and to the Higgs field. To obtain the masses of the vector bosons, we need to make one more postulate: The Higgs field obtains a nonzero value in the ground state of nature, the vacuum state, thus spontaneously breaking the  $SU(2) \times U(1)$  symmetry. This postulate is physically very nontrivial. I will discuss its foundation and implications in some detail in Section 7. However, for now, I will consider this a known aspect of the SM.

We assign the Higgs field  $\varphi$  the  $SU(2) \times U(1)$  quantum numbers  $I = \frac{1}{2}$ ,  $Y = \frac{1}{2}$ . The Higgs field is thus a spinor in isospin space, a 2-component complex-valued vector of fields

$$\varphi = \begin{pmatrix} \varphi^+ \\ \varphi^0 \end{pmatrix} \quad (5)$$

The action of an  $SU(2) \times U(1)$  transformation on this field is

$$\varphi \rightarrow \exp\left[i\alpha^a \frac{\sigma^a}{2} + i\beta \frac{1}{2}\right] \begin{pmatrix} \varphi^+ \\ \varphi^0 \end{pmatrix}. \quad (6)$$



If  $\varphi$  obtains a nonzero vacuum value, we can rotate this by an  $SU(2)$  symmetry transformation into the form

$$\langle\varphi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (7)$$

where  $v$  is a nonzero value with the dimensions of GeV. Once  $\langle\varphi\rangle$  is in this form, any  $SU(2) \times U(1)$  transformation will disturb it, except for the particular direction

$$\alpha^3 = \beta, \quad (8)$$

which corresponds to a  $U(1)$  symmetry generated by  $Q = (I^3 + Y)$ . We say that the  $SU(2) \times U(1)$  symmetry generated by  $(I^a, Y)$  is spontaneously broken, leaving unbroken only the  $U(1)$  subgroup generated by  $Q$ .

This already gives us enough information to work out the mass spectrum of the vector bosons. The kinetic energy term for  $\varphi$  in the SM Lagrangian is

$$\mathcal{L} = \left| D_\mu \varphi \right|^2 \quad (9)$$

Replacing  $\varphi$  by its vacuum value (7), this becomes

$$\mathcal{L} = \frac{1}{2} \begin{pmatrix} 0 & v \end{pmatrix} \left( g \frac{\sigma^a}{2} A_\mu^a + g' \frac{1}{2} B_\mu \right)^2 \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (10)$$

Multiplying this out and taking the matrix element, we find, from the  $\sigma^1$  and  $\sigma^2$  terms

$$\frac{g^2 v^2}{8} \left[ (A_\mu^1)^2 + (A_\mu^2)^2 \right], \quad (11)$$

and, from the remaining terms

$$\frac{v^2}{8} \left( -g A_\mu^3 + g' B_\mu \right)^2 \quad (12)$$

So, three linear combinations of the vector fields obtain mass by virtue of the spontaneous symmetry breaking. This is the mechanism of mass generation called the *Higgs mechanism* [9–11]. The mass eigenstates are

$$\begin{aligned} W^\pm &= (A^1 \mp i A^2) / \sqrt{2} & m_W^2 &= g^2 v^2 / 4 \\ Z &= (g A^3 - g' B) / \sqrt{g^2 + g'^2} & m_Z^2 &= (g^2 + g'^2) v^2 / 4 \\ A &= (g' A^3 + g B) / \sqrt{g^2 + g'^2} & m_A^2 &= 0 \end{aligned} \quad (13)$$

As we will see more clearly in a moment, the massless boson  $A$  is associated with the unbroken gauge symmetry  $Q$ . The combination of local gauge symmetry and the Higgs mechanism is the only known way to give mass to a vector boson that is consistent with Lorentz invariance and the positivity of the theory.

The linear combinations in (13) motivate the definition of the *weak mixing angle*  $\theta_w$ , defined by

$$\begin{aligned} \cos \theta_w &\equiv c_w = g / \sqrt{g^2 + g'^2} \\ \sin \theta_w &\equiv s_w = g' / \sqrt{g^2 + g'^2}. \end{aligned} \quad (14)$$

The factors  $c_w, s_w$  will appear throughout the formulae that appear in these lectures. For reference, the value of the weak mixing angle turns out to be such that

$$s_w^2 \approx 0.231 \quad (15)$$

I will describe the measurement of  $s_w$  in some detail in Section 3.

An important relation that follows from (13), (14) is

$$m_W = m_Z c_w . \quad (16)$$

This is a nontrivial consequence of the quantum number assignments for the Higgs field, and the statement that the masses of  $W$  and  $Z$  come only from the vacuum value of  $\varphi$ . Using the Particle Data Group values for the masses [12] and the value (15), we find

$$80.385 \text{ GeV} \approx 91.188 \text{ GeV} \cdot 0.877 = 79.965 \text{ GeV} . \quad (17)$$

so this prediction works well already at the leading order. We will see in Section 3 that, when radiative corrections are included, the relation (16) is satisfied to better than 1 part per mil.

Once we have the mass eigenstates of the vector bosons, the couplings of quarks and leptons to these bosons can be worked out from the expression (3) for the covariant derivative. The terms in (3) involving  $A_\mu^1$  and  $A_\mu^2$  appear only for  $I = \frac{1}{2}$  particles and can be recast as

$$-i \frac{g}{\sqrt{2}} (W_\mu^+ \sigma^+ + W_\mu^- \sigma^-) , \quad (18)$$

The  $W$  bosons couple only to  $SU(2)$  doublets, with universal strength  $g$ .

The terms with  $A_\mu^3$  and  $B_\mu$  can similarly be recast in terms of  $Z_\mu$  and  $A_\mu$ ,

$$\begin{aligned} -igA_\mu^3 - ig'B_\mu Y &= -i\sqrt{g^2 + g'^2} \left[ c_w(c_w Z_\mu + s_w A_\mu) I^3 + s_w(-s_w Z_\mu + c_w A_\mu) \right] \\ &= -i\sqrt{g^2 + g'^2} \left[ s_w c_w A_\mu (I^3 + Y) + Z_\mu (c_w^2 I^3 - s_w^2 Y) \right] \\ &= -i\sqrt{g^2 + g'^2} \left[ s_w c_w A_\mu (I^3 + Y) + Z_\mu (I^3 - s_w^2 (I^3 + Y)) \right] . \end{aligned} \quad (19)$$

We now see explicitly that the massless gauge boson  $A_\mu$  couples to  $Q = (I^3 + Y)$ , as we had anticipated. Its coupling constant is

$$e = s_w c_w \sqrt{g^2 + g'^2} = \frac{gg'}{\sqrt{g^2 + g'^2}} . \quad (20)$$

We can then identify this boson with the photon and the coupling constant  $e$  with the strength of electric charge. The quantity  $Q$  is the (numerical) electric charge of each given fermion or boson species. The expression (19) then simplifies to

$$-ieA_\mu Q - i \frac{e}{s_w c_s} Z_\mu Q_Z , \quad (21)$$

where the  $Z$  charge is

$$Q_Z = (I^3 - s_w^2 Q) . \quad (22)$$

To complete the specification of the SM, we assign the  $SU(2) \times U(1)$  quantum numbers to the quarks and leptons in each generation. As I will explain below, each quark or lepton is build up from fields of left- and right-handed chirality, associated with massless left- and right-handed particles and massless right- and left-handed antiparticles. For the applications developed in Sections 3–5, it will almost always be appropriate to ignore the masses of quarks and leptons, so these quantum number assignments will apply literally. The generation of masses for quarks and leptons is part of the physics of the Higgs field, which we will discuss beginning in Section 6.

In the SM, the left-handed fields are assigned  $I = \frac{1}{2}$ , and the right-handed fields are assigned  $I = 0$ . It is not so easy to understand how these assignments come down from fundamental theory. They are required by experiment, as I will explain in later in this section.



With this understanding, we can assign quantum numbers to the quarks and leptons as

$$\begin{aligned}
\nu_L : I^3 = +\frac{1}{2}, Y = -\frac{1}{2}, Q = 0 & & \nu_R : I^3 = 0, Y = 0, Q = 0 \\
e_L : I^3 = -\frac{1}{2}, Y = -\frac{1}{2}, Q = -1 & & e_R : I^3 = 0, Y = -1, Q = -1 \\
u_L : I^3 = +\frac{1}{2}, Y = \frac{1}{6}, Q = \frac{2}{3} & & u_R : I^3 = 0, Y = \frac{2}{3}, Q = \frac{2}{3} \\
d_L : I^3 = +\frac{1}{2}, Y = \frac{1}{6}, Q = -\frac{1}{3} & & d_R : I^3 = 0, Y = -\frac{1}{3}, Q = -\frac{1}{3}
\end{aligned} \tag{23}$$

The  $\nu_L$  and  $e_L$ , and the  $u_L$  and  $d_L$ , belong to the same  $SU(2)$  multiplet, so they must be assigned the same hypercharge  $Y$ . Note that (23) gives the correct electric charge assignments for all quarks and leptons. The  $\nu_R$  do not couple to the SM gauge fields and will play no role in the results reviewed in these lectures.

## 2.2 Massless fermions

The idea that massless fermions can be separated into left- and right-handed components will play a major role throughout these lectures. In this sentence, I introduce some notation that makes it especially easy to apply this idea.

To begin, write the the 4-component Dirac spinor and the Dirac matrices as

$$\Psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \quad \gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}, \tag{24}$$

with

$$\sigma^\mu = (1, \vec{\sigma})^\mu \quad \bar{\sigma}^\mu = (1, -\vec{\sigma})^\mu. \tag{25}$$

In this representation, the vector current takes the form

$$j^\mu = \bar{\Psi} \gamma^\mu \Psi = \psi_L^\dagger \bar{\sigma}^\mu \psi_L + \psi_R^\dagger \sigma^\mu \psi_R \tag{26}$$

and splits neatly into pieces that involve only the L or R fields. The L and R fields are mixed by the fermion mass term. In circumstances in which we can ignore the fermion masses, the L and R fermion numbers are separately conserved. We can treat  $\psi_L$  and  $\psi_R$  as completely independent species and assign them different quantum numbers, as we have already in (23). The label L, R is called *chirality*. For massless fermions, the chirality of the fields and the helicity of the particles are identical. For massive fermions, there is a change of basis from the chirality states to the helicity eigenstates.

The spinors for massless fermions are very simple. In the basis (24), we can write these spinors as

$$U(p) = \begin{pmatrix} u_L \\ u_R \end{pmatrix} \quad V(p) = \begin{pmatrix} v_R \\ v_L \end{pmatrix}. \tag{27}$$

For massless fermions, where the helicity and chirality states are identical, the spinors for a fermion with left-handed spin have  $u_R = 0$  and the spinors for an antifermion with right-handed spin have  $v_L = 0$ ; the opposite is true for a right-handed fermion and a left-handed antifermion. The nonzero spinor components for a massless fermion of energy  $E$  take the form

$$\begin{aligned}
u_L(p) &= \sqrt{2E} \xi_L & v_R(p) &= \sqrt{2E} \xi_L \\
u_R(p) &= \sqrt{2E} \xi_R & v_L(p) &= \sqrt{2E} \xi_R
\end{aligned} \tag{28}$$

where  $\xi_R$  is the spin-up and  $\xi_L$  is the spin-down 2-component spinor along the direction of motion. For example, for a fermion or antifermion moving in the  $\hat{3}$  direction,

$$\xi_L = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \xi_R = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \quad (29)$$

Spinors for other directions are obtained by rotating these according to the usual formulae for spin  $\frac{1}{2}$ . The reversal for antifermions can be thought of by viewing right-handed (for example) antifermions as holes in the Dirac sea of left-handed fermions. For a massive fermion moving in the  $\hat{3}$  direction, with

$$p^\mu = (E, 0, 0, p)^\mu, \quad (30)$$

the solutions to the Dirac equation are

$$\begin{aligned} U_L(p) &= \begin{pmatrix} \sqrt{E+p} \xi_L \\ \sqrt{E-p} \xi_L \end{pmatrix} & V_R(p) &= \begin{pmatrix} \sqrt{E+p} \xi_L \\ -\sqrt{E-p} \xi_L \end{pmatrix} \\ U_R(p) &= \begin{pmatrix} \sqrt{E-p} \xi_R \\ \sqrt{E+p} \xi_R \end{pmatrix} & V_L(p) &= \begin{pmatrix} \sqrt{E-p} \xi_R \\ -\sqrt{E+p} \xi_R \end{pmatrix}, \end{aligned} \quad (31)$$

with  $\xi_L, \xi_R$  given by (29). These formulae go over to (28) in the zero mass limit.

The matrix elements for creation or annihilation of a massless fermion pair will appear very often in these lectures. For annihilation of a fermion pair colliding along the  $\hat{3}$  axis,

$$\begin{aligned} \langle 0 | j^\mu | e_L^- e_R^+ \rangle &= v_R^\dagger \bar{\sigma}^\mu u_L \\ &= \sqrt{2E} \begin{pmatrix} -1 & 0 \end{pmatrix} (1, -\sigma^1, -\sigma^2, -\sigma^3) \sqrt{2E} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \end{aligned} \quad (32)$$

Note that I have rotated the  $e^+$  spinor appropriately by  $180^\circ$ . This gives

$$\langle 0 | j^\mu | e_L^- e_R^+ \rangle = 2E (0, 1, -i, 0)^\mu. \quad (33)$$

It is illuminating to write this as

$$\langle 0 | j^\mu | e_L^- e_R^+ \rangle = 2\sqrt{2E} \epsilon_-^\mu, \quad (34)$$

where

$$\epsilon_+^\mu = \frac{1}{\sqrt{2}} (0, 1, +i, 0)^\mu \quad \epsilon_-^\mu = \frac{1}{\sqrt{2}} (0, 1, -i, 0)^\mu \quad (35)$$

are the vectors of  $J^3 = \pm 1$  along the  $\hat{3}$  axis. The total spin angular momentum of the annihilating fermions ( $J = 1$ ) is transferred to the current and, eventually, to the final state.

More generally, we find

$$\begin{aligned} \langle 0 | j^\mu | e_R^- e_L^+ \rangle &= 2\sqrt{2E} \epsilon_+^\mu \\ \langle 0 | j^\mu | e_L^- e_R^+ \rangle &= 2\sqrt{2E} \epsilon_-^\mu \\ \langle e_R^- e_L^+ | j^\mu | 0 \rangle &= 2\sqrt{2E} \epsilon_+^{*\mu} \\ \langle e_L^- e_R^+ | j^\mu | 0 \rangle &= 2\sqrt{2E} \epsilon_-^{*\mu}. \end{aligned} \quad (36)$$

For an annihilation process such as  $e_L^- e_R^+ \rightarrow \mu_L^- \mu_R^+$  with annihilation by a current and creation by another current, the spinors appear as

$$(u_L^\dagger \bar{\sigma}^\mu v_R) (v_R^\dagger \bar{\sigma}_\mu u_L) = 2(2E)^2 \epsilon_-^{*\mu} \cdot \epsilon_- . \quad (37)$$

To evaluate this, rotate the  $\epsilon_-$  vector for the muons into the muon direction. If the muons come off at polar angle  $\theta$ , this gives

$$\epsilon_-^{I*} = \frac{1}{\sqrt{2}}(0, \cos \theta, -i, -\sin \theta) . \quad (38)$$

Then (37) becomes

$$2(2E)^2 \epsilon_-^{I*} \cdot \epsilon_- = s(1 + \cos \theta) = -2u , \quad (39)$$

in terms of the usual kinematic invariants  $s, t, u$ . Another way to write this is

$$|(u_L^\dagger \bar{\sigma}^\mu v_R)(v_R^\dagger \bar{\sigma}_\mu u_L)|^2 = 4(2p_{e^-} \cdot p_{\mu^+})(2p_{e^+} \cdot p_{\mu^-}) . \quad (40)$$

Similarly, for  $e_L^- e_R^+ \rightarrow \mu_R^- \mu_L^+$ ,

$$|(u_R^\dagger \bar{\sigma}^\mu v_L)(v_R^\dagger \bar{\sigma}_\mu u_L)|^2 = 4(2p_{e^+} \cdot p_{\mu^+})(2p_{e^-} \cdot p_{\mu^-}) . \quad (41)$$

It is a nice exercise to check these answers using the usual trace theorems. The trace theorems are more automatic, but the helicity formalism gives more physical insight.

### 3 Tests of the $V-A$ Interaction

The property that the  $W$  boson only couples to fermions of left-handed chirality is a crucial property of the SM. It is responsible for many of the surprising features of the weak interactions, both the most attractive and the most puzzling ones. It is therefore important to understand that this feature is extremely well supported experimentally. In this section, I review the most convincing experimental tests of this property.

#### 3.1 Polarization in $\beta$ decay

The first applications discussed in this section involve exchange of  $W$  bosons at low energy. In this limit, we can simplify the  $W$  propagator to a pointlike interaction

$$\frac{-i}{q^2 - m_W^2} \rightarrow \frac{i}{m_W^2} . \quad (42)$$

In this limit, the  $W$  exchange can be represented by the product of currents

$$\Delta \mathcal{L} = \frac{g^2}{2m_W^2} J_\mu^+ J^{-\mu} , \quad (43)$$

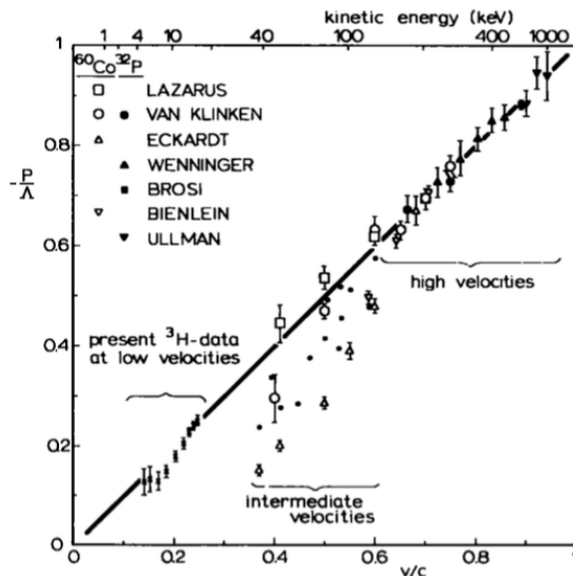
where

$$\begin{aligned} J_\mu^+ &= \nu_{eL}^\dagger \bar{\sigma}_\mu e_L + u_L^\dagger \bar{\sigma}_\mu d_L + \dots \\ J_\mu^- &= e_L^\dagger \bar{\sigma}_\mu \nu_{eL} + d_L^\dagger \bar{\sigma}_\mu u_L + \dots \end{aligned} \quad (44)$$

Here and henceforth in these lectures, I replace the label  $\psi$  with a label that gives the flavor quantum numbers of the field. In (44), I write explicitly the terms associated with the first generation quarks and leptons; the omitted terms are those for the higher generations. I ignore Cabibbo mixing, a reasonable approximation for the topics discussed in these lectures. I will also ignore the masses of the neutrinos.

The theory (43) is called the  $V-A$  interaction, since

$$u_L^\dagger \bar{\sigma}^\mu d_L = \bar{U} \gamma^\mu \frac{1 - \gamma^5}{2} D , \quad (45)$$



**Fig. 1:** Polarization of the electron emitted in  $\beta$  decay for a variety of  $\beta$  decay transitions in different nuclei, from [13].

the difference of a vector and an axial vector current. The coefficient in (43) is conventionally represented by the Fermi constant  $G_F$ ,

$$\frac{g^2}{2m_W^2} = \frac{4G_F}{\sqrt{2}}. \quad (46)$$

This interaction has *maximal parity violation* in charge-changing weak interactions.

The simplest consequence of  $V-A$  is that electrons emitted in  $\beta$  decay should be preferentially left-handed polarized. Since the energies of electrons in  $\beta$  decay are of order 1 MeV, it is typically not a good approximation to ignore the electron mass. However, since in  $V-A$  the electron is produced in the L chirality eigenstate, we can work out the polarization from the relative magnitude of the  $u_L$  terms in the left- and right-handed helicity massive spinors given in (31). The electron polarization, in the left-handed sense, is then given by

$$\text{Pol}(e^-) = \frac{(\sqrt{E+p})^2 - (\sqrt{E-p})^2}{(\sqrt{E+p})^2 + (\sqrt{E-p})^2} = \frac{p}{E} = \frac{v}{c}. \quad (47)$$

A data compilation is shown in Fig. 1 [13]. Careful experiments both at high and low electron energies verify the regularity (47).

### 3.2 Muon decay

The  $V-A$  interaction also has striking consequences for the electron energy and polarization in muon decay.

It is not difficult to work out the basic formulae for muon decay. In  $V-A$  theory, and ignoring the electron mass, muon decay has a massive muon at rest decaying to  $\nu_{\mu L} e_L^- \bar{\nu}_{e R}$ . For the muon at rest, averaged over polarizations, we find, instead of (40),

$$|(u_L^\dagger \bar{\sigma}^\mu v_R)(v_R^\dagger \bar{\sigma}_\mu u_L)|^2 = 2(2p_{e^-} \cdot p_\nu)(2p_{\bar{\nu}} \cdot p_{\mu^-}). \quad (48)$$

To integrate this over phase space, let

$$x_i = \frac{2p_i \cdot p_\mu}{p_\mu^2}, \quad (49)$$

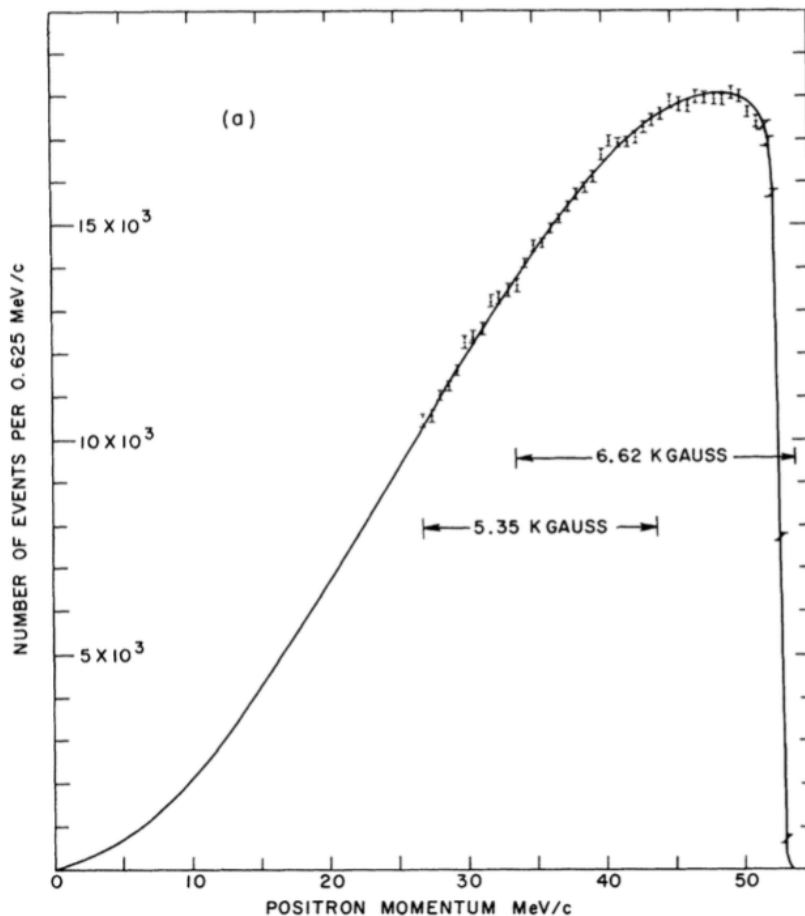


Fig. 2: Energy spectrum of  $e^+$  in  $\mu^+$  decay at rest, from [14].

where  $i = e, \nu, \bar{\nu}$ . Conservation of energy-momentum  $p_\mu = p_e + p_\nu + p_{\bar{\nu}}$  implies

$$x_e + x_\nu + x_{\bar{\nu}} = 2. \quad (50)$$

Each  $x_i$  takes the maximum value 1 when that massless particle recoils against the other two massless particles. Note also that

$$2p_e \cdot p_\nu = (p_e + p_\nu)^2 = (p_\mu - p_{\bar{\nu}})^2 = m_\mu^2(1 - x_{\bar{\nu}}). \quad (51)$$

Three-body phase space takes a simple form in the  $x_i$  variables,

$$\int d\Pi_3 = \frac{m_\mu^2}{128\pi^3} \int dx_e dx_{\bar{\nu}}. \quad (52)$$

Assembling the pieces, the muon decay rate is predicted to be

$$\Gamma = \frac{1}{2m_\mu} \left( \frac{4G_F}{\sqrt{2}} \right)^2 \frac{m_\mu^2}{128\pi^3} \int dx_e dx_{\bar{\nu}} 2m_\mu^4 x_{\bar{\nu}}(1 - x_{\bar{\nu}}). \quad (53)$$

The integral over  $x_{\bar{\nu}}$  is

$$\int_{1-x_e}^1 dx_{\bar{\nu}} x_{\bar{\nu}}(1 - x_{\bar{\nu}}) = \frac{1}{2}x_e^2 - \frac{1}{3}x_e^3. \quad (54)$$

Then finally we find for the electron energy distribution

$$\frac{d\Gamma}{dx_e} = \frac{G_F^2 m_\mu^5}{16\pi^3} \left( \frac{x_e^2}{2} - \frac{x_e^3}{3} \right). \quad (55)$$

This shape of this distribution is quite characteristic, with a double zero at  $E_e = 0$  and zero slope at the endpoint at  $E_e = m_\mu/2$ . Both effects are slightly rounded by radiative corrections, but, with these taken into account, the prediction agrees with the measured spectrum to high precision, as shown in Fig. 2 [14].

### 3.3 Pion decay

Charged pion decay is mediated by the  $V-A$  interaction

$$\frac{4G_F}{\sqrt{2}} (d_L^\dagger \bar{\sigma}^\mu u_L) \left( \nu_{eL}^\dagger \bar{\sigma}_\mu e_L + \nu_{\mu L}^\dagger \bar{\sigma}_\mu \mu_L \right) \quad (56)$$

At first sight, it might seem that the  $\pi^+$  must decay equally often to  $e^+$  and  $\mu^+$ . Experimentally, almost all pion decays are to  $\mu^+$ . Can this be reconciled with  $V-A$ ?

The pion matrix element is

$$\langle 0 | d_L^\dagger \bar{\sigma}^\mu u_L | \pi^+(p) \rangle = -i \frac{1}{2} F_\pi p^\mu, \quad (57)$$

where  $F_\pi$  is the pion decay constant, equal to 135 MeV. The matrix element of (56) then evaluates to

$$\frac{4G_F}{\sqrt{2}} \cdot \left( -\frac{i}{2} F_\pi \right) p^\mu U_{\nu L}^\dagger \bar{\sigma}_\mu V_{\ell^+}. \quad (58)$$

The pion is at rest, so

$$p^\mu \bar{\sigma}_\mu = m_\pi \cdot 1. \quad (59)$$

The neutrino is (essentially) massless and therefore must be left-handed. The pion has spin 0, so angular momentum requires that the  $\ell^+$  is also left-handed. But, from (31), the lepton spinor is then

$$V_L = \begin{pmatrix} \sqrt{E-p} \xi_R \\ \times \end{pmatrix} \quad (60)$$

The matrix element (58) reduces to

$$i \frac{4G_F}{\sqrt{2}} \cdot \left( \frac{1}{2} F_\pi \right) \sqrt{2E_\nu m_\pi} \sqrt{E_\ell - p_\ell}. \quad (61)$$

Two-body kinematics gives  $E_\nu = p_\nu = p_\ell = (m_\pi^2 - m_\ell^2)/2m_\pi$ . Then  $(E_\ell - p_\ell) = m_\ell^2/m_\pi^2$ . Phase space includes the factor  $2p_\ell/m_\pi$ , which brings another factor of  $(E_\ell - p_\ell)$ . Finally we find

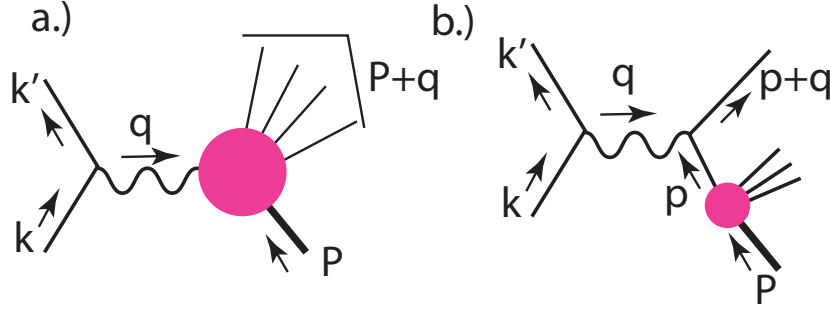
$$\Gamma(\pi^+ \rightarrow \ell^+ \nu) = \frac{G_F^2 m_\pi^3 F_\pi^2}{8\pi} \frac{m_\ell^2}{m_\pi^2} \left( 1 - \frac{m_\ell^2}{m_\pi^2} \right)^2. \quad (62)$$

The overall factor  $m_\ell^2/m_\pi^2$  comes from the matrix element (60). Angular momentum conservation requires the  $\ell^+$  to have the wrong helicity with respect to  $V-A$ , accounting for this suppression factor.

The result (62) leads to the ratio of branching fractions

$$\frac{BR(\pi^+ \rightarrow e^+ \nu_e)}{BR(\pi^+ \rightarrow \mu^+ \nu_\mu)} = \frac{m_e^2}{m_\mu^2} \left( \frac{m_\pi^2 - m_e^2}{m_\pi^2 - m_\mu^2} \right)^2 = 1.28 \times 10^{-4}, \quad (63)$$

in good agreement with the observed value  $1.23 \times 10^{-4}$ .



**Fig. 3:** Kinematics of neutrino deep inelastic scattering: (a) for neutrino scattering from a proton or heavy nucleus, (b) for neutrino scattering from a quark in the parton model description.

### 3.4 Neutrino deep inelastic scattering

The helicity structure of the  $V-A$  interaction is also seen in the energy distributions in deep inelastic neutrino scattering. For electrons, deep inelastic scattering is the scattering from a proton or nuclear target in which the momentum transfer is large and the target is disrupted to a high mass hadronic state. The kinematics is shown in Fig. 3(a). In the leading order of QCD, deep inelastic scattering is described by the scattering for the electron from a single quark in the parton distribution of the target. This kinematics is shown in Fig. 3(b).

Neutrino deep inelastic scattering experiments are done in the following way: One first creates a high-energy pion beam by scattering protons from a target. Then the pions are allowed to decay, producing a beam of neutrinos and muons. The beam is made to pass through a long path length of absorber to remove the muons and residual pions and other hadrons. Finally, the neutrinos are allowed to interact with a large-volume detector. A charged-current neutrino reaction then leads to a scattering event whose result is a  $\mu^\pm$ , depending on the charge of the decaying pion, and a high-multiplicity hadronic system.

If  $k$  is the initial momentum of the neutrino,  $k'$  is the final momentum of the muon, and  $P$  is the initial momentum of the target proton, we let  $q = (k - k')$  and define the Lorentz invariants

$$\begin{aligned} s &= (k + P)^2 & Q^2 &= -q^2 \\ x &= \frac{Q^2}{2P \cdot q} & y &= \frac{2P \cdot q}{2P \cdot k} \end{aligned} \quad (64)$$

We are interested in the deep inelastic limit  $Q^2 \gg P^2 = m_p^2$ . Then  $s \approx 2p \cdot k$  and  $Q^2 = xys$ . In the lab frame  $P = (m_p, \vec{0})$ , so  $y = q^0/k^0$ , the fraction of the initial neutrino energy transferred to the proton. To the extent that the initial neutrino energy  $k^0$  is known, all of the invariants  $x$ ,  $y$ , and  $Q^2$  can be determined by measurement of the final muon momentum.

At leading order in QCD, a deep inelastic reaction is an essentially elastic lepton-quark scattering, for example,  $\nu + d \rightarrow \mu^- + u$ . Using Feynman's parton model, which is also the basis for QCD predictions at hadron colliders, we model the proton or nuclear target as a collection of quarks and antiquarks that move collinearly and share the total momentum of the proton. Let  $p$  be the momentum of the initial quark, and approximate

$$p = \xi P, \quad (65)$$

where  $0 < \xi < 1$ . The quarks might also have transverse momentum relative to the proton, but this is ignorable if the momentum transfer  $Q^2$  from the neutrino scattering is large.

The final momentum of the quark is then  $p + q$ . The condition that this quark is on-shell is

$$0 = (p + q)^2 = 2p \cdot q + q^2 = 2\xi P \cdot q - Q^2. \quad (66)$$

Then

$$\xi = \frac{Q^2}{2P \cdot q} = x. \quad (67)$$

This is a remarkable result, also due to Feynman: To the leading order in QCD, deep inelastic scattering events at a given value of the invariant  $x$  arise from scattering from quarks or antiquarks in the proton with momentum fraction  $\xi = x$ .

We can now evaluate the kinematic invariants for a neutrino-quark scattering event. I call these  $\hat{s}$ ,  $\hat{t}$ ,  $\hat{u}$  to distinguish them from the invariants of neutrino-proton scattering. First,

$$\hat{s} = (p + k)^2 = 2p \cdot k = 2\xi P \cdot k = x s. \quad (68)$$

The momentum transfer can be evaluated from the lepton side, so

$$\hat{t} = q^2 = -Q^2. \quad (69)$$

Finally, for scattering of approximately massless particles,  $s + t + u = 0$ , so

$$\hat{u} = xs - Q^2 = xs(1 - y). \quad (70)$$

The aspect of the deep inelastic scattering cross section that is most important for the subject of this lecture is the distribution in  $y$ . To begin, consider the deep inelastic scattering of a  $\nu_\mu$ . The quark-level reaction is

$$\nu + d \rightarrow \mu^- + u \quad (71)$$

In the  $V-A$  theory, the  $\nu$  and the  $d$  must be left-handed. Similarly to (41),

$$|(u_L^\dagger(\mu^-)\bar{\sigma}^\mu u_L(\nu))(u_L^\dagger(u)\bar{\sigma}_\mu u_L(d))|^2 = 4(2p_{\mu^-} \cdot p_u)(2p_\nu \cdot p_d) = 4\hat{s}^2. \quad (72)$$

On the other hand, antineutrino scattering from a quark, which proceeds by the reaction

$$\bar{\nu} + u \rightarrow \mu^+ + d, \quad (73)$$

is, in  $V-A$  theory, the scattering of a right-handed  $\bar{\nu}$  and a left-handed  $u$ . Then

$$|(v_R^\dagger(\mu^+)\bar{\sigma}^\mu v_R(\bar{\nu}))(u_L^\dagger(u)\bar{\sigma}_\mu u_L(d))|^2 = 4(2p_{\mu^+} \cdot p_u)(2p_{\bar{\nu}} \cdot p_d) = 4\hat{u}^2. \quad (74)$$

Inserting (68), (70), we see that the dependence of the deep inelastic scattering cross section on  $y$  should be

$$\begin{aligned} \frac{d\sigma}{dy}(\nu p \rightarrow \mu^- X) &\sim \hat{s}^2 \sim 1 \\ \frac{d\sigma}{dy}(\bar{\nu} p \rightarrow \mu^+ X) &\sim \hat{u}^2 \sim (1 - y)^2. \end{aligned} \quad (75)$$

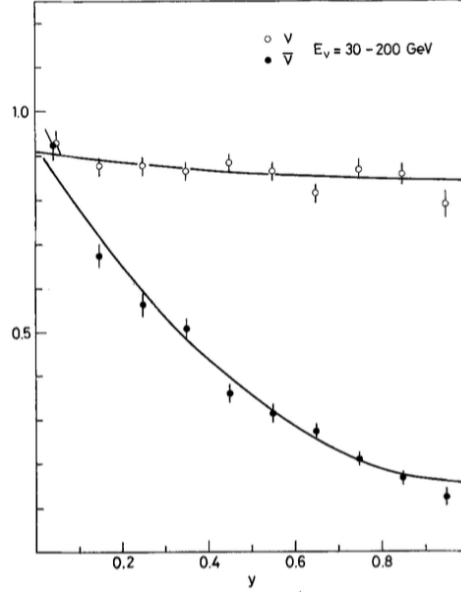
These results, which I have derived for a proton target, hold for any nuclear target under the assumption that we consider only scattering from quarks and not from antiquarks. For scattering from antiquarks, the dependence on  $y$  is reversed, with a  $(1 - y)^2$  dependence for neutrino scattering. The experimental result, from the CDHS experiment, a CERN neutrino experiment of the 1980's, is shown in Fig. 4 [15]. The  $y$  distribution for neutrino scattering is indeed almost flat, and that for antineutrino scattering is close to  $(1 - y)^2$ . The deviations from these ideal results are consistent with arising from the antiquark content of the proton and neutron.

The same regularity can be seen in collider physics. For example, the Standard Model predicts that, in quark-antiquark annihilation to a  $W$  boson,

$$\begin{aligned} \frac{d\sigma}{d\cos\theta}(d\bar{u} \rightarrow W^- \rightarrow \mu^- \bar{\nu}) &\sim u^2 \sim (1 + \cos\theta)^2 \\ \frac{d\sigma}{d\cos\theta}(u\bar{d} \rightarrow W^+ \rightarrow \mu^+ \nu) &\sim t^2 \sim (1 - \cos\theta)^2, \end{aligned} \quad (76)$$

and these distributions are well verified at the LHC [16, 17].





**Fig. 4:** Dependence on the variable  $y$  of the cross sections for neutrino and antineutrino scattering on an iron target, from [15].

### 3.5 $e^+e^-$ annihilation at high energy

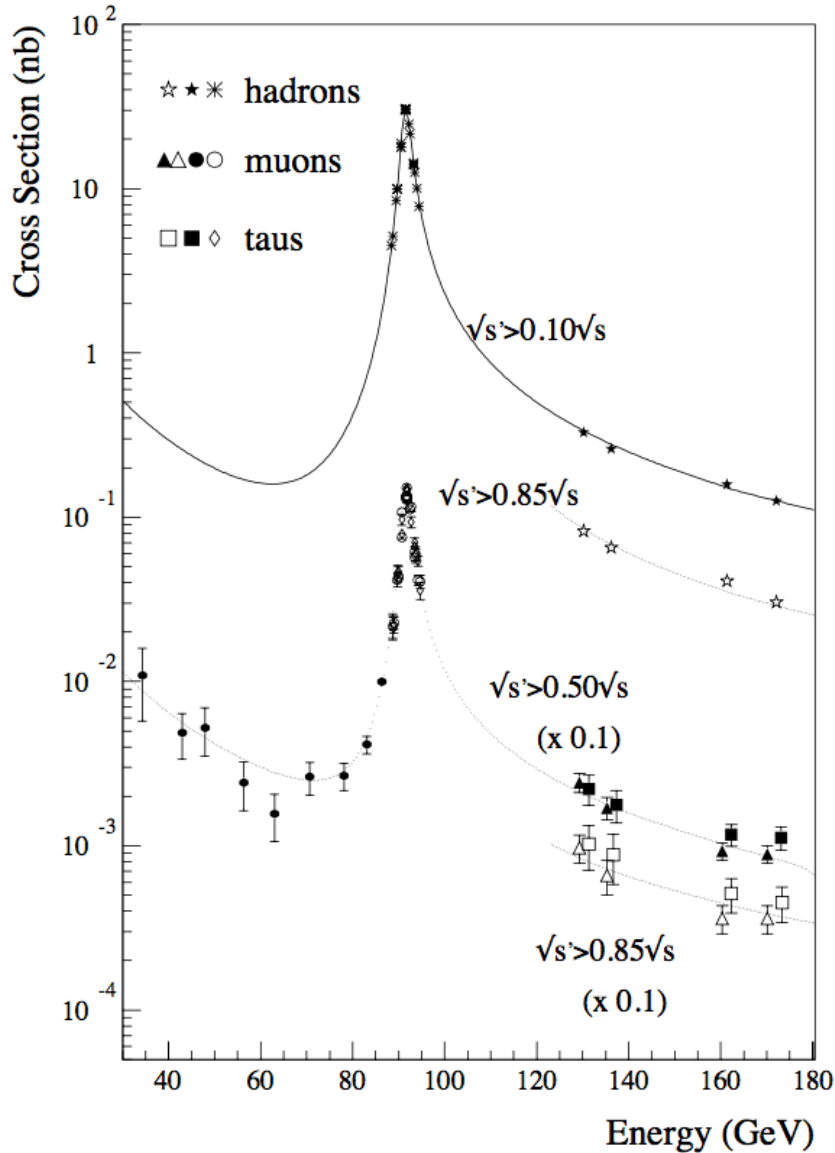
The angular distributions in annihilation through the neutral current are more complex, first, because of photon- $Z$  interference, and, second, because the weak neutral current couples to both left- and right-handed quarks and leptons.

To write formulae for the cross sections in  $e^+e^-$  annihilation to a fermion pair, it is simplest to begin with the cross sections for polarized initial and final states. Using the same principles for evaluating spinor products as before, it is not difficult to work these out. The general form of the differential cross sections is

$$\begin{aligned}
 \frac{d\sigma}{d\cos\theta}(e_L^-e_R^+ \rightarrow f_L\bar{f}_R) &= \frac{\pi\alpha^2}{2s} |sF_{LL}(s)|^2 (1 + \cos\theta)^2 \\
 \frac{d\sigma}{d\cos\theta}(e_R^-e_L^+ \rightarrow f_L\bar{f}_R) &= \frac{\pi\alpha^2}{2s} |sF_{RL}(s)|^2 (1 - \cos\theta)^2 \\
 \frac{d\sigma}{d\cos\theta}(e_L^-e_R^+ \rightarrow f_R\bar{f}_L) &= \frac{\pi\alpha^2}{2s} |sF_{LR}(s)|^2 (1 - \cos\theta)^2 \\
 \frac{d\sigma}{d\cos\theta}(e_R^-e_L^+ \rightarrow f_R\bar{f}_L) &= \frac{\pi\alpha^2}{2s} |sF_{RR}(s)|^2 (1 + \cos\theta)^2.
 \end{aligned} \tag{77}$$

The form factors  $F_{IJ}(s)$  reflect photon $\gamma$ - $Z$  interference, with the  $p\gamma$  charges  $Q$  and the  $Z$  charges  $Q_Z$  in (22). Using the subscript  $f$  to denote the flavor and chirality of the fermion,

$$\begin{aligned}
 F_{LL}(s) &= \frac{Q_f}{s} + \frac{(1/2 - s_w^2)(I_f^3 - s_w^2 Q_f)}{s_w c_w} \frac{1}{s - m_Z^2} \\
 F_{RL}(s) &= \frac{Q_f}{s} + \frac{(-s_w^2)(-s_w^2 Q_f)}{s_w c_w} \frac{1}{s - m_Z^2} \\
 F_{LR}(s) &= \frac{Q_f}{s} + \frac{1/2 - s_w^2}{s_w c_w} (I_f^3 - s_w^2 Q_f) \frac{1}{s - m_Z^2} \\
 F_{RR}(s) &= \frac{Q_f}{s} + \frac{(-s_w^2)(-s_w^2 Q_f)}{s_w c_w} \frac{1}{s - m_Z^2}.
 \end{aligned} \tag{78}$$

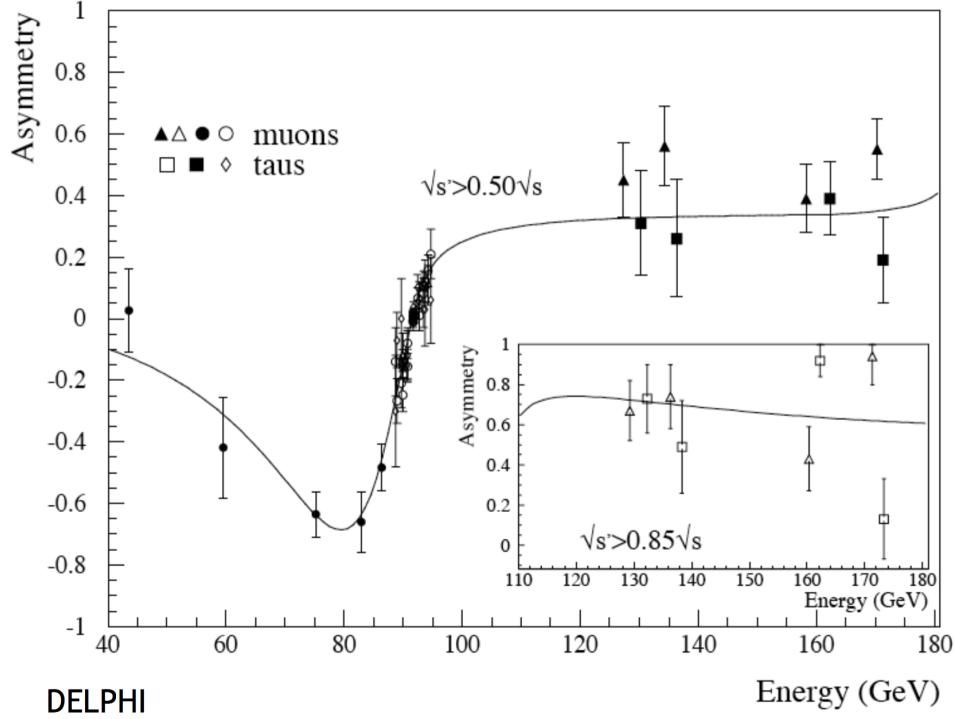


**Fig. 5:** Total cross section for  $e^+e^- \rightarrow \text{hadrons}$ ,  $e^+e^- \rightarrow \mu^+\mu^-$ , and  $e^+e^- \rightarrow \tau^+\tau^-$ , as a function of center of mass energy, as measured by the DELPHI experiment at the collider LEP [18]. The continuous lines are the predictions of the SM.

The total cross sections predicted from these formulae for  $e^+e^- \rightarrow \text{hadrons}$ ,  $e^+e^- \rightarrow \mu^+\mu^-$ , and  $e^+e^- \rightarrow \tau^+\tau^-$  are shown in Fig. 5 and compared to data from the DELPHI experiment at the CERN  $e^+e^-$  collider LEP. The resonance at the center of mass energy of 91 GeV is of course the  $Z$  boson.

Notice that, for  $s > m_Z^2$ , we have constructive interference in the LL and RR polarization states and destructive interference for RL and LR. Then in an experiment with unpolarized beams (as in the program of  $e^+e^-$  experiments at LEP), the LL and RR modes should dominate and produce a positive forward-backward asymmetry in the angular distribution. This behavior is actually seen in the data. Figure 6 shows the forward-backward asymmetry in  $e^+e^- \rightarrow \mu^+\mu^-$  and  $e^+e^- \rightarrow \tau^+\tau^-$  measured by the DELPHI experiment at LEP [18]. The solid line is the prediction of the SM.

It is interesting to explore the high-energy limits of the expressions (78). Begin with  $F_{RL}(s)$ ,



**Fig. 6:** Forward-backward asymmetry in the reactions  $e^+e^- \rightarrow \mu^+\mu^-$  and  $e^+e^- \rightarrow \tau^+\tau^-$ , as a function of center of mass energy, as measured by the DELPHI experiment at the collider LEP [18]. .

corresponding to  $e_R^-e_L^+ \rightarrow f_L\bar{f}_R$ . In the limit  $s \gg m_Z^2$  and inserting  $Q = I_f^3 + Y$ , this becomes

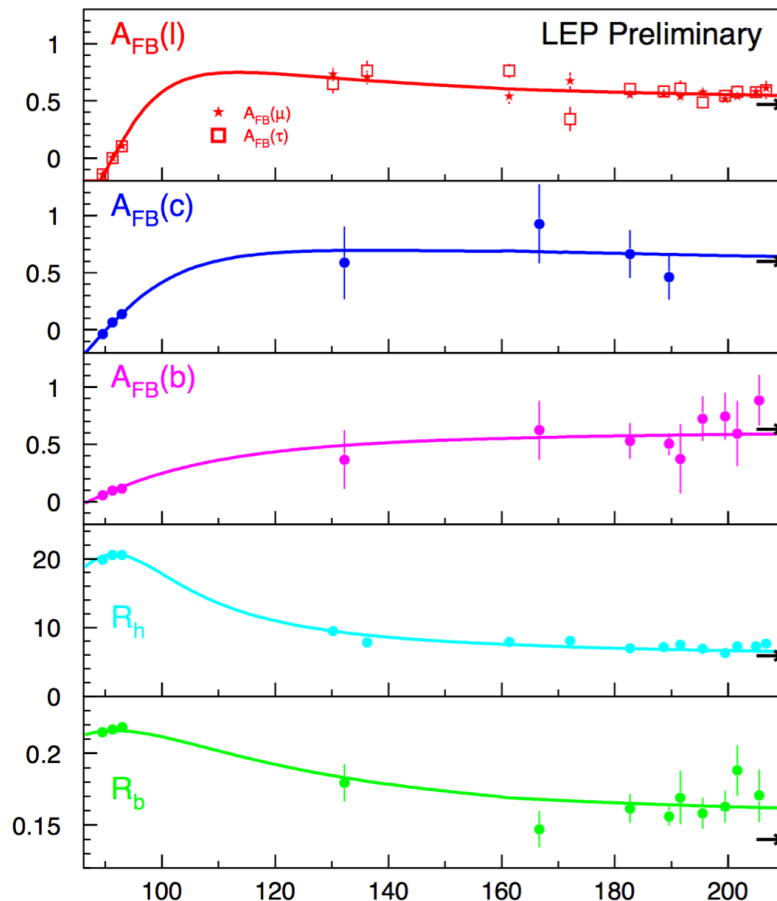
$$\begin{aligned}
 F_{RL} &\rightarrow \frac{s_w^2 c_w^2 (I_f^3 + Y_f) - s_w^2 I_f^3 + s_w^4 (I_f^3 + Y_f)}{s_w^2 c_w^2 s} \\
 &= \frac{s_w^2 Y_f}{s_w^2 c_w^2 s} \\
 &= \frac{1}{e^2} \left( \frac{g'^2 Y_{eR} Y_f}{s} \right). \tag{79}
 \end{aligned}$$

The expression in parentheses is exactly the amplitude for  $s$ -channel exchange of the  $U(1)$  boson  $B$  in the situation in which the original  $SU(2) \times U(1)$  symmetry was not spontaneously broken. So we see that the full gauge symmetry is restored at high energies.

Here is the same analysis for  $F_{LL}(s)$ :

$$\begin{aligned}
 F_{RL} &\rightarrow \frac{s_w^2 c_w^2 (I_f^3 + Y_f) + (1/2 - s_w^2)(I_f^3 - s_w^2(I_f^3 + Y_f))}{s_w^2 c_w^2 s} \\
 &= \frac{(1/2)c_w^2 I_f^3 + (1/2)s_w^2 Y_f}{s_w^2 c_w^2 s} \\
 &= \frac{1}{e^2} \left( \frac{g^2 I_{eL}^3 I_f^3}{s} + \frac{g'^2 Y_{eR} Y_f}{s} \right). \tag{80}
 \end{aligned}$$

Now the result is a coherent sum of  $A^3$  and  $B$  exchanges in the  $s$ -channel. Again, this is the result expected in a theory of unbroken  $SU(2) \times U(1)$ .



**Fig. 7:** Compilation of preliminary LEP measurements of the forward-backward asymmetry in lepton,  $c$ , and  $b$  pair production, the hadron to lepton ratio  $R_h$  and the  $b$  to all hadron ratio  $R_b$  [19]. The solid curves show the SM prediction. The arrows at the right are the predictions of unbroken  $SU(2) \times U(1)$ .

It is interesting to compare the values of ratios and asymmetries measured at LEP to the asymptotic values predicted by unbroken  $SU(2) \times U(1)$ . This comparison is shown in Fig. 7 from a compilation of preliminary LEP results [19]; final LEP results on 2-fermion processes are collected in [20]. The arrows at the extreme right show the values for restored  $SU(2) \times U(1)$ . The calculation of  $R_b$  involves a top quark box diagram that does not yet reach its asymptotic limit at 200 GeV. It is remarkable that, for all other observables, the LEP measurements at center of mass energies of 200 GeV are already close to the asymptotic values predicted at high energy.

#### 4 Precision electroweak measurements at the $Z$ resonance

It is possible to test the SM theory of the weak interactions more incisively by focusing more tightly on the properties of the  $Z$  boson. The  $Z$  boson appears as a resonance in  $e^+e^-$  annihilation. In the 1990's, the accelerators LEP at CERN and SLC at SLAC tuned their energies to the  $Z$  boson resonance to produce large numbers of  $Z$  bosons at rest in the lab, in an appropriate setting for precision measurements. In this section, I review the results of these precision measurements, which continue to provide important constraints on the SM and its generalizations.

#### 4.1 Properties of the $Z$ boson in the Standard Model

My discussion will be based on the leading order matrix elements for  $Z$  decay to  $f_L \bar{f}_R$  and  $f_R \bar{f}_L$ . It is straightforward to work these out based on the spinor matrix elements computed in Section 2.2. The leading order matrix element for  $Z$  decay to  $f_L \bar{f}_R$  is

$$\mathcal{M}(Z \rightarrow f_L \bar{f}_R) = i \frac{g}{c_w} Q_{Zf} u_L^\dagger \bar{\sigma}^\mu v_R \epsilon_{Z\mu}, \quad (81)$$

with

$$Q_Z = I^3 - s_w^2 Q, \quad (82)$$

as in (22). Using (36) for the spinor matrix element, this becomes

$$\mathcal{M} = i \frac{g}{c_w} \sqrt{2} m_Z \epsilon_-^* \cdot \epsilon_Z. \quad (83)$$

Square this and average over the direction of the fermion, or, equivalently, average over three orthogonal directions for the  $Z$  polarization vector. The result is

$$\langle |\mathcal{M}|^2 \rangle = \frac{2}{3} \frac{g^2}{c_w^2} Q_{Zf}^2 m_Z^2. \quad (84)$$

Then, since

$$\Gamma(Z \rightarrow f_L \bar{f}_R) = \frac{1}{2m_Z} \frac{1}{8\pi} \langle |\mathcal{M}|^2 \rangle, \quad (85)$$

we find

$$\Gamma(Z \rightarrow f_L \bar{f}_R) = \frac{\alpha_w m_Z}{6c_w^2} Q_{Zf}^2 N_f, \quad (86)$$

where

$$\alpha_w = \frac{g^2}{4\pi} \quad (87)$$

and

$$N_f = \begin{cases} 1 & \text{lepton} \\ 3(1 + \alpha_s/\pi + \dots) & \text{quark} \end{cases} \quad (88)$$

accounts the number of color states and the QCD correction. The same formula holds for the  $Z$  width to  $f_R \bar{f}_L$ .

To evaluate this formula, we need values of the weak interaction coupling constants. The electromagnetic coupling  $\alpha$  is famously close to  $1/137$ . However, in quantum field theory,  $\alpha$  is a running coupling constant that becomes larger at small distances. At a scale of  $Q = m_Z$ ,  $\alpha(Q) = 1/129$ . Later in the lecture, I will defend a value of the weak mixing angle

$$s_w^2 = 0.231. \quad (89)$$

Then the  $SU(2)$  and  $U(1)$  couplings take the values

$$\alpha_w = \frac{g^2}{4\pi} = \frac{1}{29.8} \quad \alpha' = \frac{g'^2}{4\pi} = \frac{1}{99.1} \quad (90)$$

It is interesting to compare these values to other fundamental SM couplings taken at the same scale  $Q = m_Z$ ,

$$\alpha_s = \frac{1}{8.5} \quad \alpha_t = \frac{y_t^2}{4\pi} = \frac{1}{12.7}. \quad (91)$$

All of these SM couplings are roughly of the same order of magnitude.

Using (89) or (90), we can tabulate the values of the  $Z$  couplings to left- and right-handed fermions,

| species | $Q_{ZL}$                          | $Q_{ZR}$            | $S_f$ | $A_f$ |
|---------|-----------------------------------|---------------------|-------|-------|
| $\nu$   | $+\frac{1}{2}$                    | -                   | 0.250 | 1.00  |
| $e$     | $-\frac{1}{2} + s_w^2$            | $+s_W^2$            | 0.126 | 0.15  |
| $u$     | $+\frac{1}{2} - \frac{2}{3}s_w^2$ | $-\frac{2}{3}s_W^2$ | 0.143 | 0.67  |
| $d$     | $-\frac{1}{2} + \frac{1}{3}s_w^2$ | $+\frac{1}{3}s_W^2$ | 0.185 | 0.94  |

In this table, the quantities evaluated numerically are

$$S_f = Q_{ZL}^2 + Q_{ZR}^2 \quad A_f = \frac{Q_{ZL}^2 - Q_{ZR}^2}{Q_{ZL}^2 + Q_{ZR}^2}. \quad (92)$$

The quantity  $S_f$  gives the contribution of the species  $f$  to the total decay rate of the  $Z$  boson. The quantity  $A_f$  gives the polarization asymmetry for  $f$ , that is, the preponderance of  $f_L$  over  $f_R$ , in  $Z$  decays,

## 4.2 Measurements of the $Z$ properties

It is possible to measure many of the total rates and polarization asymmetries for individual species in a very direct way through experiments on the  $Z$  resonance. This subject is reviewed in great detail in the report [22]. Values of the  $Z$  observables given below are taken from this reference unless it is stated otherwise.

The  $S_f$  are tested by the measurement of the  $Z$  resonance width and its branching ratios. Using (86), we find for the total width of the  $Z$

$$\Gamma_Z = \frac{\alpha_w m_Z}{6c_w^2} \left[ 3 \cdot 0.25 + 3 \cdot 0.126 + 2 \cdot (3.1) \cdot 0.144 + 3 \cdot (3.1) \cdot 0.185 \right]. \quad (94)$$

The four terms denote the contributions from 3 generations of  $\nu$ ,  $e$ ,  $u$ , and  $d$ , minus the top quark, which is too heavy to appear in  $Z$  decays. The numerical prediction is

$$\Gamma_Z = 2.49 \text{ GeV} \quad (95)$$

The separate terms in (94) give the branching ratios

$$\begin{aligned} BR(\nu_e \bar{\nu}_e) &= 6.7\% & BR(e^+ e^-) &= 3.3\% \\ BR(u \bar{u}) &= 11.9\% & BR(d \bar{d}) &= 15.3\% \end{aligned} \quad (96)$$

The measured value of the total width, whose extraction I will discuss in a moment, is

$$\Gamma_Z = 2.4952 \pm 0.0023 \text{ GeV} . \quad (97)$$

This is in very good agreement with (95), with accuracy such that a valid comparison with theory requires the inclusion of electroweak radiative corrections, with typically are of order 1%. The measurements of branching ratios and polarization asymmetry that I review later in this section are also of sub-% accuracy. At the end of this section, I will present a more complete comparison of theory and experiment, including radiative corrections to the theoretical predictions.

To begin our review of the experimental measurements, we should discuss the measurement of the  $Z$  resonance mass and width in more detail. Ideally, the  $Z$  is a Breit-Wigner resonance, with cross section shape

$$\sigma \sim \left| \frac{1}{s - m_Z^2 + im_Z \Gamma_Z} \right|^2. \quad (98)$$

At first sight, it seems that we can simply read off the  $Z$  mass as the maximum of the resonance and the width as the observed width at half maximum. However, we must take into account that the resonance is distorted by initial-state radiation. As the electron and positron collide and annihilate into a  $Z$ , they can radiate hard collinear photons. Because of this, the resonance is pushed over to higher energies, an effect that shifts the peak and creates a long tail on above the resonance. The magnitude of the photon radiation is given by the parameter

$$\beta = \frac{2\alpha}{\pi} \left( \log \frac{s}{m_e^2} - 1 \right) = 0.108 \quad \text{at } s = m_Z^2 \quad (99)$$

In addition, since the  $Z$  is narrow, the effect of this radiation is magnified, since even a relatively soft photon can push the center of mass energy off of the resonance. The size of the correction can be roughly estimated as

$$-\beta \cdot \log \frac{m_Z}{\Gamma_Z} = 40\% . \quad (100)$$

To make a proper accounting of this effect, we need to include arbitrary numbers of radiated collinear photons. Fadin and Kuraev introduced the idea of viewing the radiated photons and the final annihilating electron as partons in the electron in the same way that quarks and gluons are treated as partons in the proton [21]. For the proton, the parton distribution is generated by non-perturbative effects, but for the electron the parton distributions are generated only by QED, so that they can be calculated as a function of  $\alpha$ . The result for the parton distribution of the electron in the electron, to order  $\alpha$ , is

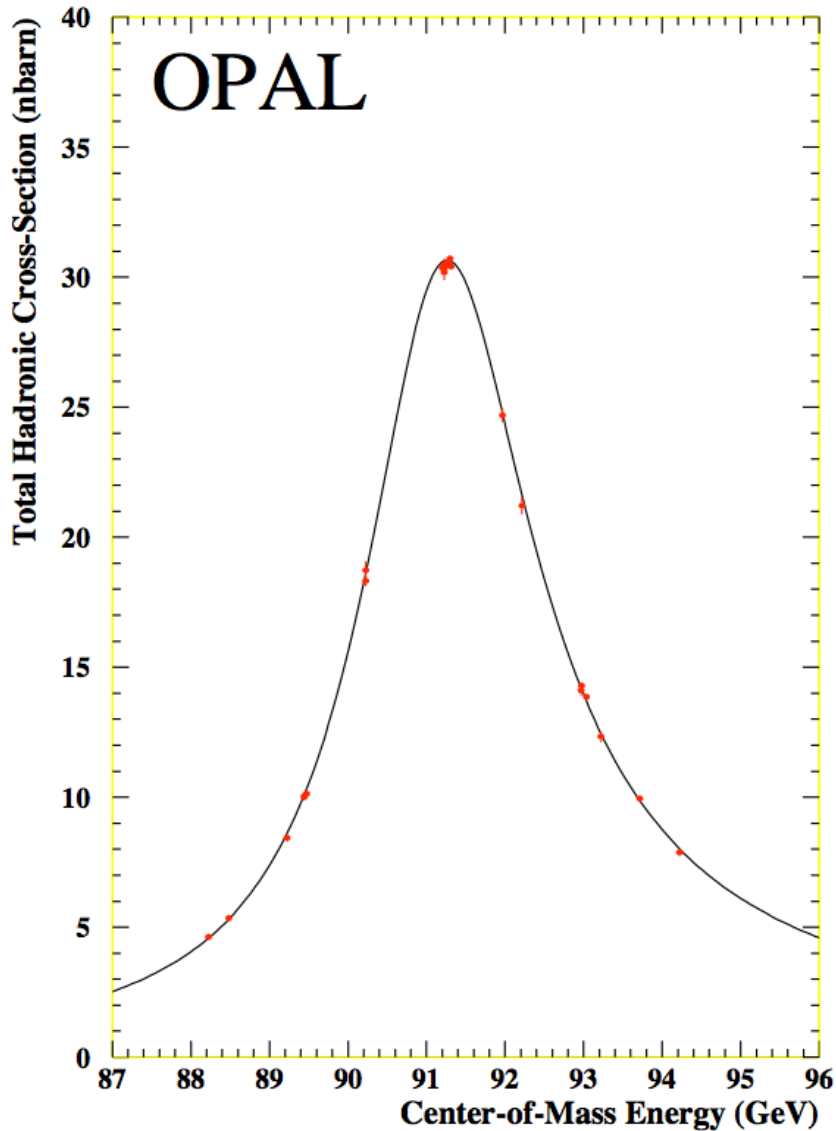
$$f_e(z, s) = \frac{\beta}{2} (1-z)^{\beta/2-1} \left( 1 + \frac{3}{8}\beta \right) - \frac{1}{4}\beta(1+z) + \dots , \quad (101)$$

where  $z$  is the momentum fraction of the original electron carried into the  $e^+e^-$  annihilation to a  $Z$  boson. The cross section for producing a  $Z$  boson would then be a convolution of the Breit-Wigner cross section (98) with the parton distribution (101) and the corresponding distribution for the positron. For the LEP experiments, this theory was extended to include two orders of subleading logarithms and finite corrections of order  $\alpha^2$  [23].

The experimental aspects of the measurement of the  $Z$  resonance lineshape were also very challenging; see Section 2.2 of [22]. Careful control was needed for point-to-point normalization errors across the  $Z$  resonance. The absolute energy of the LEP ring was calibrated using resonant depolarization of a single electron beam and then corrected for two-beam effects. This calibration was found to depend on the season and the time of day. Some contributing effects were the changes in the size of the LEP tunnel due to the annual change in the water level in Lake Geneva and current surges in the LEP magnets due to the passage to the TGV leaving Geneva for Paris.

Some final results for the resonance line shape measurement are shown in Figs. 8, 9. The first of these figures shows the measurements by the OPAL experiment over the resonance and the detailed agreement of the shape between theory and experiment [24]. The second shows the combination of the resonance height and width measurements from the four LEP experiments ALEPH, DELPHI, L3, and OPAL [22]. In this figure, the lower curve is the radiatively corrected result; the higher curve is the inferred Breit-Wigner distribution excluding the effects of radiative corrections.

The measurement of branching ratios is more straightforward. It is necessary only to collect  $Z$  decay events and sort them into categories. The various types of leptonic and hadronic decay modes have very different, characteristic forms. Typical events are shown in Fig. 10 for hadronic,  $e^+e^-$ ,  $\mu^+\mu^-$ , and  $\tau^+\tau^-$  decays [25]. The major backgrounds are from Bhabha scattering and 2-photon events. These do not resemble  $Z$  decay events and are rather straightforwardly separated. Nonresonant  $e^+e^-$  annihilations are also a small effect, generally providing backgrounds at only the level of parts per mil. An exception is the  $Z$  decay to  $\tau^+\tau^-$ , which can be faked by hadronic  $e^+e^-$  annihilations with radiation to provide a background level of a few percent. Still, these high signal to background ratios are completely different from the situation at the LHC and enable measurements of very high precision.



**Fig. 8:** Resonance line shape of the  $Z$  in  $e^+e^-$  annihilation, as measured by the OPAL experiment [24].

Two particular branching ratios merit special attention. First, consider  $Z$  decays to invisible final states. The SM includes  $Z$  decays to 3 species of neutrino, with a total branching ratio of 20%. Even though these decays are not seen in the detector, the presence of invisible final states affects the resonance lineshape by increasing the  $Z$  width and decreasing the  $Z$  peak height to visible modes such as hadrons. Measurement of the resonance parameters then effectively gives the number of light neutrinos into which the  $Z$  can decay. The result is

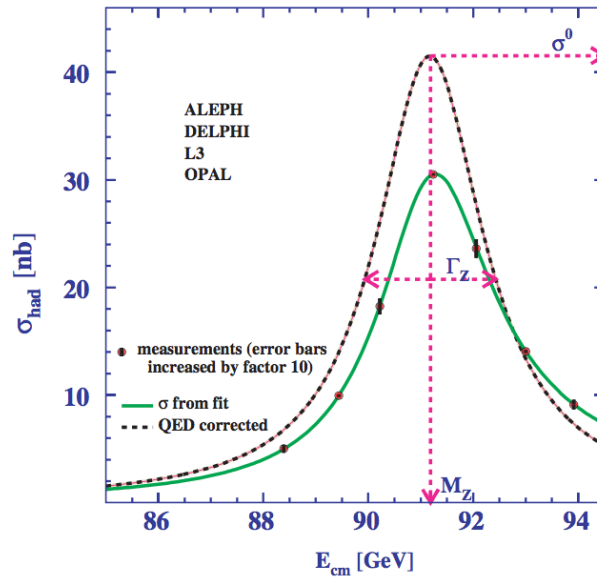
$$n_\nu = 2.9840 \pm 0.0082, \quad (102)$$

strongly constraining extra neutrinos or more exotic neutral particles.

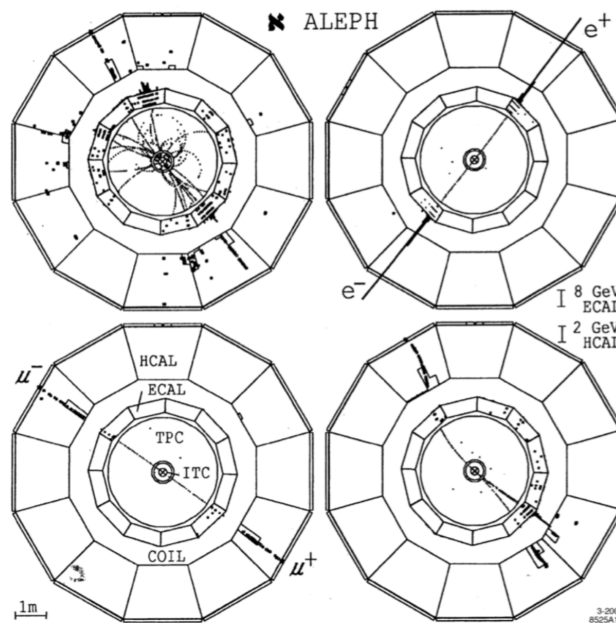
Second, the  $Z$  branching ratio to  $b$  quarks is of special interest, for two reasons. First, the  $b$  belongs to the same  $SU(2) \times U(1)$  multiplet as the top quark, and, even in the SM, there is a relatively large radiative correction due to top quark loops, from the diagrams shown in Fig. 11. These produce

$$Q_{ZbL} = -\left(\frac{1}{2} - \frac{1}{3}s_w^2 - \frac{\alpha}{16\pi s_w^2} \frac{m_t^2}{m_W^2}\right), \quad (103)$$

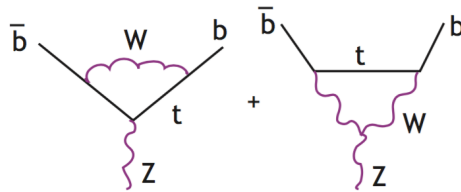




**Fig. 9:** Resonance line shape of the  $Z$  in  $e^+e^-$  annihilation, as measured by the four LEP experiments, from [22]. The dotted curve shows the zeroth-order resonance line shape of the  $Z$  resonance. The solid line shows the Standard Model prediction including initial-state radiative corrections.



**Fig. 10:** Typical  $e^+e^- \rightarrow Z$  events corresponding to the  $Z$  decays to hadrons, to  $e^+e^-$ , to  $\mu^+\mu^-$ , and to  $\tau^+\tau^-$ , from [25].



**Fig. 11:** Diagrams containing the top quark which give a relatively large correction to the partial width for  $Z \rightarrow b\bar{b}$ .

a shift of about  $-2\%$ . More generally, the  $b$  is a third-generation particle that might have a nontrivial coupling to new, heavier, particles.

An observable that specifically tracks this effect is

$$R_b = \frac{\Gamma(Z \rightarrow b\bar{b})}{\Gamma(Z \rightarrow \text{hadrons})}. \quad (104)$$

At leading order, we predict  $R_b = 0.22$ , but in the full SM this value should be reduced according to (103).  $Z$  decays to  $b\bar{b}$  could be identified by vertex tags. The SLD detector at SLAC included a pixel vertex detector capable of separating decays to  $b$  and  $c$  by vertex mass and by the presence of tertiary charm decay vertices in  $b$  jets. Fig. 12(a) shows the signal and background separation in the OPAL experiment [26]. Fig. 12(b) shows a corresponding result from SLD, in which the observed vertex mass was used to discriminate between the  $c$  and  $b$  contributions [27]. The final LEP and SLC results gave

$$\begin{aligned} R_b &= 0.21629 \pm 0.00066 \\ R_c &= 0.1721 \pm 0.0030, \end{aligned} \quad (105)$$

confirming the shift predicted by (103) and demonstrating consistency with the SM also for  $Z \rightarrow c\bar{c}$ .

While the total rates for the  $Z$  decay to the various species have similar values, the asymmetries listed in (92) vary over a wide range, from 15% for the charged leptons to almost maximal for the  $d$ -type quarks. The SM predicts these disparate values from a common value of  $s_w^2$ .

There are three very different methods to measure the lepton asymmetries  $A_e$ . First, the  $A_e$  can be found from the forward-backward asymmetry for  $e^+e^- \rightarrow f\bar{f}$  at the  $Z$ . Second,  $A_e$  can be determined from the final-state polarization effects in the decays of  $\tau^+\tau^-$  produced at the  $Z$ . Finally,  $A_e$  can be measured directly from the rate for  $Z$  production from polarized electron beams.

For unpolarized beams, the angular distribution for  $e^+e^- \rightarrow f\bar{f}$  can be found from (77). On the  $Z$  resonance, the distribution takes the form

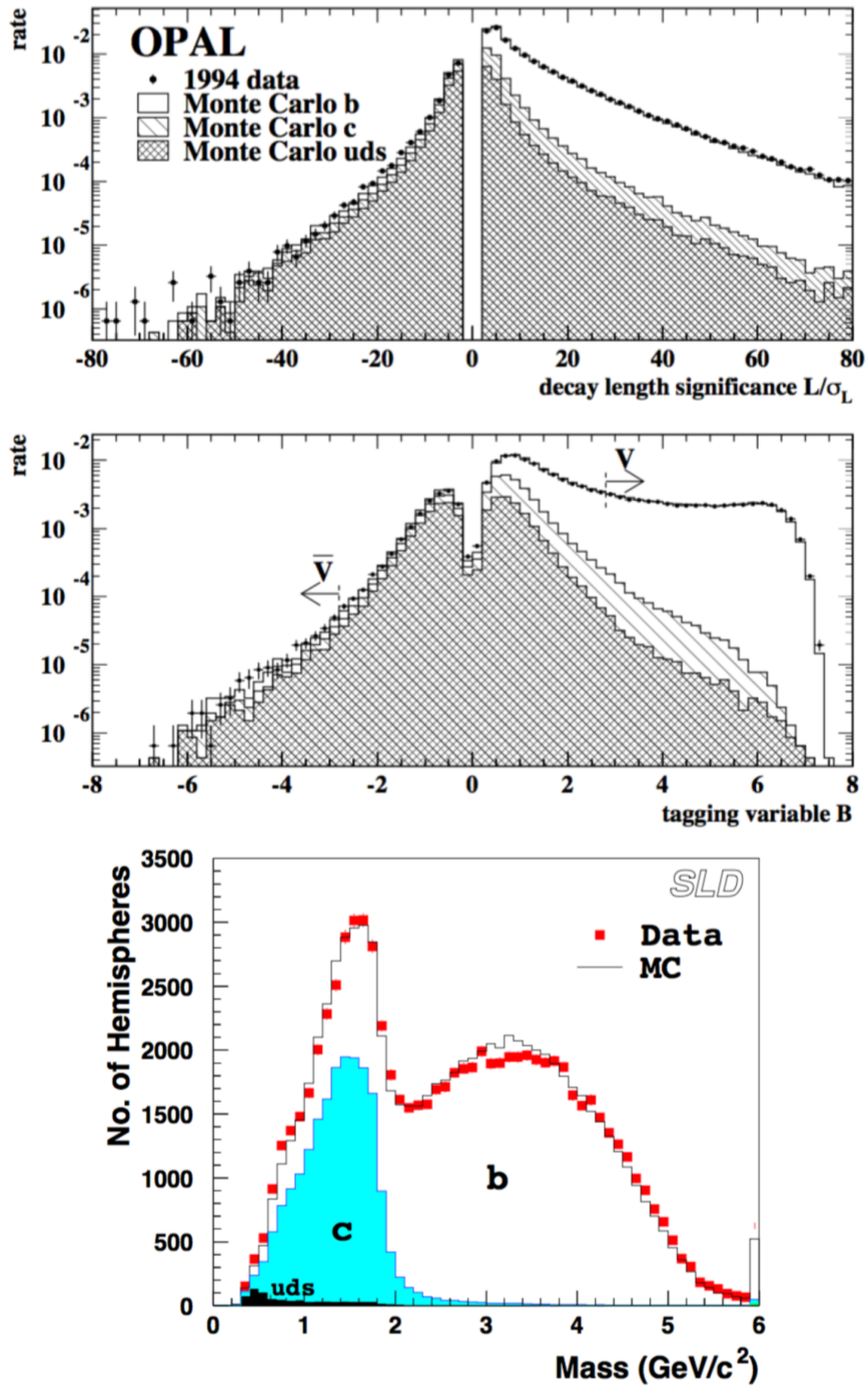
$$\begin{aligned} \frac{d\sigma}{d\cos\theta} &= \left(\frac{1+A_e}{2}\right)\left(\frac{1+A_f}{2}\right)(1+\cos\theta)^2 + \left(\frac{1-A_e}{2}\right)\left(\frac{1+A_f}{2}\right)(1-\cos\theta)^2 \\ &\quad + \left(\frac{1+A_e}{2}\right)\left(\frac{1-A_f}{2}\right)(1-\cos\theta)^2 + \left(\frac{1-A_e}{2}\right)\left(\frac{1-A_f}{2}\right)(1+\cos\theta)^2. \end{aligned} \quad (106)$$

The forward-backward asymmetry predicted by this expression is

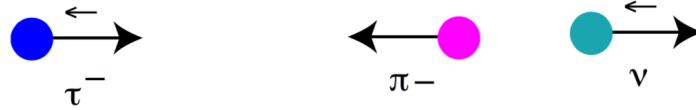
$$A_{FB} = \frac{3}{4}A_eA_f \quad (107)$$

Especially for  $b$  quarks, which have an almost maximal asymmetry, the dependence of this quantity on  $s_w^2$  is mainly through  $A_e$ .

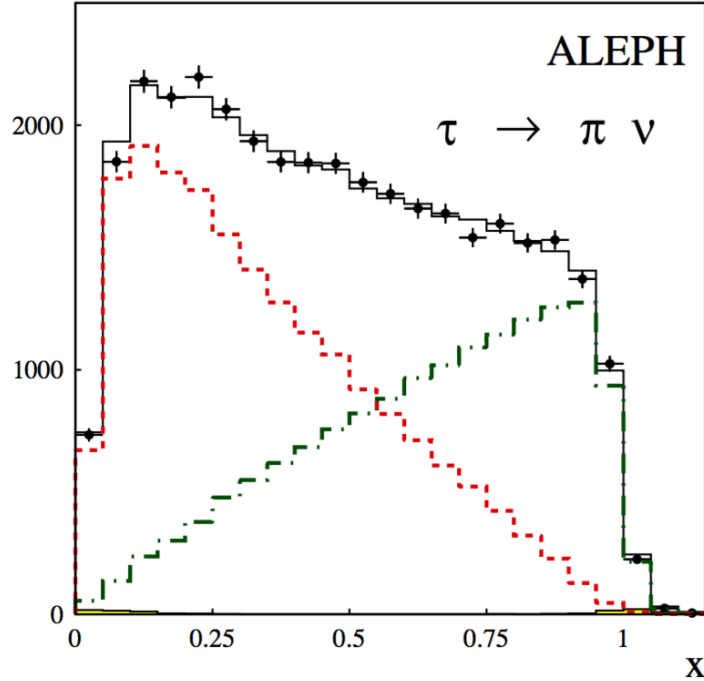
The value of  $A_e$  determines the polarization of  $\tau$  leptons produced in  $Z$  decays, and this polarization becomes visible through the  $V-A$  structure of the  $\tau$  decays. The easiest case to understand is the



**Fig. 12:** Measurements of the  $b$  and  $c$  branching fractions of the  $Z$ . Top: Distributions in decay length significance and the  $b$  quark tagging variable, from the OPAL experiment, showing the relative contributions of light quarks,  $c$ , and  $b$ , from [26]. Right: Vertex mass distribution from the SLD experiment, showing the contributions from  $c$  and  $b$  meson decays, from [27].



**Fig. 13:** Kinematics of  $\tau \rightarrow \nu\pi$  decay.



**Fig. 14:** Pion energy spectrum in  $\tau \rightarrow \nu\pi$  decays at the  $Z$  resonance, from [28]. The ordinate  $x = 2E_\pi/m_Z$ . The separate contributions from  $\tau_L$  and  $\tau_R$  decays are indicated.

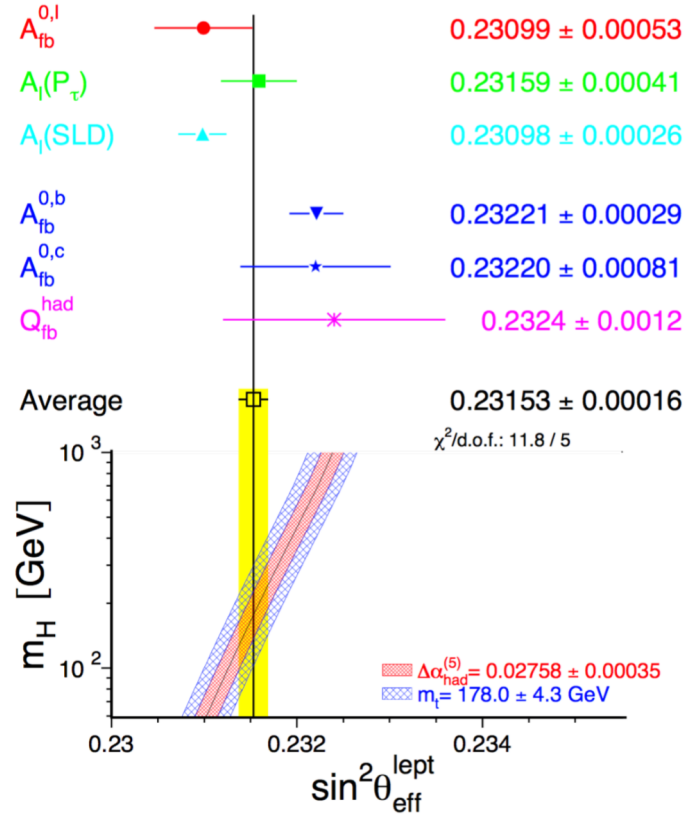
decay  $\tau^- \rightarrow \nu_\tau\pi^-$ . Since the neutrino is always left-handed and the pion has zero spin, a  $\tau^-$  at rest with  $S^3 = -\frac{1}{2}$  will decay to a forward neutrino and a backward  $\pi^-$ , as shown in Fig. 13. When the  $\tau^-$  is boosted, a left-handed  $\tau$  will decay to a high-energy neutrino and a slow pion. A right-handed  $\tau$  will decay to a low-energy neutrino and a fast pion. More generally, if  $x$  is the fraction of the  $\tau$  momentum carried by the  $\pi^-$ ,

$$\tau_L : \frac{d\Gamma}{dx} \sim (1-x) \quad \tau_R : \frac{d\Gamma}{dx} \sim x. \quad (108)$$

Similar asymmetries appear in the other  $\tau$  decay modes. Fig. 14 shows the distributions measured by the ALEPH experiment for  $\tau \rightarrow \pi\nu$ , compared to the expected distributions from  $\tau_L$  and  $\tau_R$ . The 15% asymmetry is apparent. The SM also predicts a correlation between polarization and  $\cos\theta$  that can be used to improve the  $s_w^2$  measurement.

The SLC produced  $e^+e^- \rightarrow Z$  events using linear acceleration of the electrons. This technique allowed the preservation of electron polarization from the source to the collisions. The experiment was conducted by flipping the the electron polarization in each bunch randomly, and measuring the correlation between the polarization orientation and the total  $Z$  production rate—measured 4 km downstream of the source. This gave a direct measurement [29]

$$A_e = 0.1516 \pm 0.0021 \quad (109)$$



**Fig. 15:** Summary of  $A_\ell$  measurements at the  $Z$  resonance from different observables, from [22].

Figure 15 shows the summary of the various determinations of  $s_w^2$  from the leptonic asymmetries [22]. The measurements are statistically consistent and lead to a very precise value.

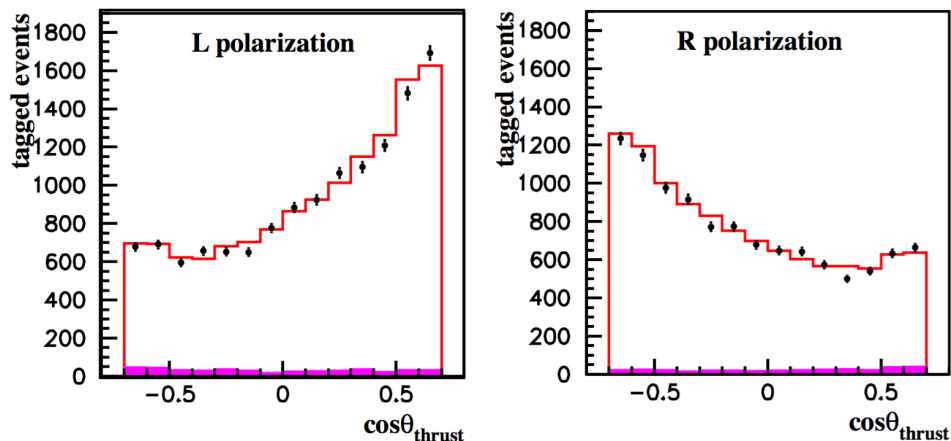
The prediction that the  $b$  asymmetry is close to maximal implies that the angular distribution of  $e^+e^- \rightarrow b\bar{b}$  at the  $Z$  should show a large dependence on beam polarization. The distribution should be close to  $(1 + \cos\theta)^2$  for a left-handed polarized beam and close to  $(1 - \cos\theta)^2$  for a right-handed polarized beam. The distributions measured by the SLD experiment at the SLC for left- and right-handed beams are shown in Fig. 16. Allowing for the expected confusion in separating  $b$  and  $\bar{b}$  jets, the results are consistent with a high  $b$  polarization in  $Z$  decays. The difference in normalization of the two distributions reflects the 15% asymmetry in the production cross section.

Figure 17 shows a summary of the precision measurements of the properties of the  $Z$  boson [22]. The measured values listed in the first column are compared to the values from the best fit to the SM, including one-loop radiative corrections. The bars show the deviations from the SM prediction, in units of the  $\sigma$  of the measurement. This is an impressive confirmation of the  $SU(2) \times U(1)$  weak interaction model.

### 4.3 Constraints on oblique radiative corrections

From the excellent agreement of the  $Z$  measurements with the SM, it is possible to put general constraints on possible new particles coupling to the weak interactions.

To explain this, we should first discuss the properties of one-loop corrections to the  $SU(2) \times U(1)$  predictions in more detail. The SM contains a large number of parameters. However, the predictions



**Fig. 16:** Angular distribution of  $e^+e^- \rightarrow Z \rightarrow b\bar{b}$  events measured by the SLD experiment for left- and right-handed polarized beams, from [30].

discussed in this Section depend, at the three level, only on the three parameters

$$g, g', v. \quad (110)$$

The loop corrections will include divergences, including quadratically divergent corrections to  $v^2$ . However, because the  $SU(2) \times U(1)$  theory is renormalizable, once these three parameters are fixed, all of the 1-loop corrections must be finite. Then each specific reaction acquires a finite prediction, which is a testable consequence of the SM.

Different schemes are used to fix the three underlying divergent amplitudes. Each gives different expressions for the measurable cross sections. Three common schemes are

- applying  $\overline{MS}$  subtraction, as in QCD
- fixing  $\alpha(m_Z), m_Z, m_W$  to their measured values (Marciano-Sirlin scheme) [32]
- fixing  $\alpha(m_Z), m_Z, G_F$  to their measured values (on shell  $Z$  scheme)

In the  $\overline{MS}$  scheme, used by the Particle Data Group, the  $\overline{MS}$  parameters  $g, g'$ , and  $v$  are unphysical but can be defined as the values that give the best fit to the corpus of SM measurements [31].

The various schemes for renormalizing the  $SU(2) \times U(1)$  model lead to different definitions of  $s_w^2$  that are found in the literature. In the Marciano-Sirlin scheme, we define  $\theta_w$  by

$$c_w \equiv m_W/m_Z. \quad (111)$$

This leads to

$$s_w^2 = 0.22290 \pm 0.00008. \quad (112)$$

We will see in Section 4 that the relation (111) is often needed to insure the correct behavior in high-energy reactions of  $W$  and  $Z$ , so it is useful that this relation is insured at the tree level. Thus, the Marciano-Sirlin definition of  $\theta_w$  is the most common one used in event generators for LHC. However, one should note that the value (112) is significantly different from the value (89) that best represents the sizes of the  $Z$  cross sections and asymmetries.

In the on-shell  $Z$  scheme,  $\theta_w$  is defined by

$$\sin^2 2\theta_w = (2c_w s_w)^2 \equiv \frac{4\pi\alpha(m_Z)}{\sqrt{2}G_F m_Z^2}, \quad (113)$$

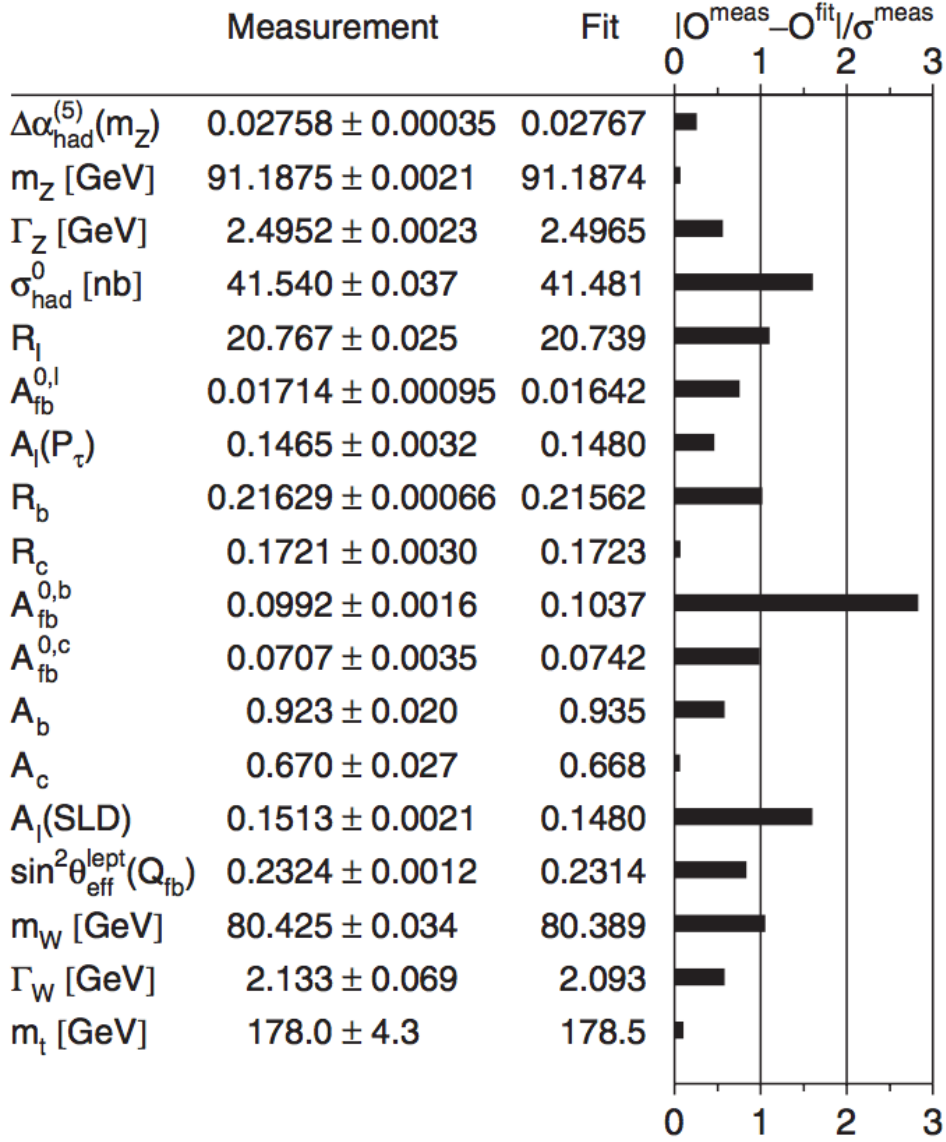


Fig. 17: Summary of precision electroweak measurements at the  $Z$  resonance, from [22].

leading to

$$s_w^2 = 0.231079 \pm 0.000036. \quad (114)$$

This definition gives at tree level a value that is much closer to (89). All three values of  $\sin^2\theta_w$  lead to the same predictions for the relation of observables to observables after the (scheme-dependent) finite 1-loop corrections are included.

One particular class of radiative corrections is especially simple to analyze. If new particles have no direct coupling to light fermions, they can appear in radiative corrections to the  $Z$  observables only through vector boson vacuum polarization amplitudes. Effects of this type are called *oblique* radiative corrections. These effects can be analyzed in a quite general way.

There are four electroweak vacuum polarization amplitudes  $\Pi_{AB}(q^2)$ . I will notate them as shown in Fig. 18. The subscripts 1, 3 refer to the weak isospin currents  $j^{\mu a}$ ,  $a = 1, 3$ ; the subscript  $Q$  refers to the electromagnetic current. The  $Z$  vacuum polarizations are found from these elements using (82). If the particles in the loop have large masses  $M$ , we can Taylor expand the vacuum polarization amplitudes



$$\begin{aligned}
A \text{ wavy line} \text{---} \text{blue circle} \text{---} A &= ie^2 \Pi_{QQ} g^{\mu\nu} \\
Z \text{ wavy line} \text{---} \text{blue circle} \text{---} A &= i \frac{e^2}{s_w c_w} (\Pi_{3Q} - s_w^2 \Pi_{QQ}) g^{\mu\nu} \\
Z \text{ wavy line} \text{---} \text{blue circle} \text{---} Z &= i \frac{e^2}{s_w^2 c_w^2} (\Pi_{33} - 2s_w^2 \Pi_{3Q} + s_w^2 \Pi_{QQ}) g^{\mu\nu} \\
W \text{ wavy line} \text{---} \text{blue circle} \text{---} W &= i \frac{e^2}{s_w^2} \Pi_{11} g^{\mu\nu}
\end{aligned}$$

**Fig. 18:** Vector boson vacuum polarization diagrams..

in powers of  $q^2/M^2$ . Up to order  $q^2/M^2$ , we find

$$\begin{aligned}
\Pi_{QQ}(q^2) &= Aq^2 + \dots \\
\Pi_{3Q}(q^2) &= Bq^2 + \dots \\
\Pi_{33}(q^2) &= C + Dq^2 + \dots \\
\Pi_{11}(q^2) &= E + Fq^2 + \dots
\end{aligned} \tag{115}$$

There are six constants in this set of formulae. Three of them are fixed by the renormalizations of  $g$ ,  $g'$ ,  $v$ . This leaves 3 finite combinations of vacuum polarization amplitudes will be predicted in any new physics model. These combinations are canonically defined as [33]

$$\begin{aligned}
S &= \frac{16\pi}{m_Z^2} \left[ \Pi_{33}(m_Z^2) - \Pi_{33}(0) - \Pi_{3Q}(m_Z^2) \right] \\
T &= \frac{4\pi}{s_w^2 m_W^2} \left[ \Pi_{11}(0) - \Pi_{33}(0) \right] \\
U &= \frac{16\pi}{m_Z^2} \left[ \Pi_{11}(m_Z^2) - \Pi_{11}(0) - \Pi_{33}(m_Z^2) + \Pi_{33}(0) \right]
\end{aligned} \tag{116}$$

In [33], the amplitudes appearing in (116) are the new physics contributions only, but other analyses, for example, [31], use different conventions. The three parameters in (116) have clear physical interpretations.  $T$  parametrizes the size of weak isospin violating corrections to the relation  $m_W = m_Z c_w$ .  $S$  parametrizes the  $q^2/M^2$  corrections.  $U$  requires both effects and is predicted to be very small in most new physics models.

The leading oblique corrections to electroweak observables can then be expressed as linear shifts proportional to  $S$  and  $T$ . For example,

$$\begin{aligned}
\frac{m_W^2}{m_Z^2} - c_0^2 &= \frac{\alpha c_w^2}{c_w^2 - s_w^2} \left( -\frac{1}{2} S + c_w^2 T \right) \\
s_*^2 - s_0^2 &= \frac{\alpha}{c_w^2 - s_w^2} \left( -\frac{1}{2} \frac{1}{4} S - s_w^2 c_w^2 T \right),
\end{aligned} \tag{117}$$

where  $s_0$ ,  $c_0$  are the values of  $s_w$  and  $c_w$  in the on-shell  $Z$  scheme and  $s_*$  is the value of  $s_w$  used to evaluate the  $Z$  asymmetries  $A_f$ . By fitting to the formulae such as (117), we can obtain general constraints that can be applied to a large class of new physics models.

Some guidance about the expected sizes of  $S$  and  $T$  is given by the result for one new heavy electroweak doublet,

$$S = \frac{1}{6\pi} \quad T = \frac{|m_U^2 - m_D^2|}{m_Z^2}. \tag{118}$$



A complete heavy fourth generation gives  $S = 0.2$ . The effects of the SM top quark and Higgs boson can also be expressed approximately in the  $S, T$  framework,

$$\begin{aligned} \text{top :} \quad S &= \frac{1}{6\pi} \log \frac{m_t^2}{m_Z^2} & T &= \frac{3}{16\pi s_w^2 c_w^2} \frac{m_t^2}{m_Z^2} \\ \text{Higgs :} \quad S &= \frac{1}{12\pi} \log \frac{m_h^2}{m_Z^2} & T &= -\frac{3}{16\pi c_w^2} \log \frac{m_h^2}{m_Z^2} \end{aligned} \quad (119)$$

The appearance of corrections proportional to  $m_t^2/m_Z^2$ , which we have already seen in (103), will be explained in Section 5.

Figure 19 shows the progress of the  $S, T$  fit with our improved understanding of the SM. Figure 19(a) reflects the situation in 1991, before the discovery of the top quark [33]. The two vertical lines to the left are predictions of the SM with a varying top quark mass. Values of  $m_t$  in the range of 170–180 GeV are highly favored by the precision electroweak data. The measurement of  $S$ , even without the value of  $m_t$ , strongly constrained the “technicolor” models of electroweak symmetry breaking. (I will describe these models at the end of Section 7.2.) Figure 19(b) shows the  $S, T$  fit in 2008. The solid curve shows the predictions of the SM with a variable Higgs boson mass. Values of the Higgs mass close to 100 GeV are strongly favored. Figure 19(c) shows the current  $S, T$  fit [34]. The fit is in good agreement with the SM with the now-measured values of  $m_t$  and  $m_h$ . It also is in substantial tension with the presence of a fourth generation of quarks and leptons.

## 5 The Goldstone Boson Equivalence Theorem

In this section, I will describe the properties of the weak interactions at energies much greater than  $m_W$  and  $m_Z$ . Some new conceptual issues appear here. These affect the energy-dependence of  $W$  and  $Z$  boson reactions at high energy and the parametrization of possible effects of new physics. I will introduce a way of thinking that can be used as a skeleton key for understanding these issues, called the Goldstone Boson Equivalence Theorem.

### 5.1 Questions about $W$ and $Z$ bosons at high energy

To begin this discussion, I will raise a question, one that turns out to be one of the more difficult questions to answer about spontaneously broken gauge theories.

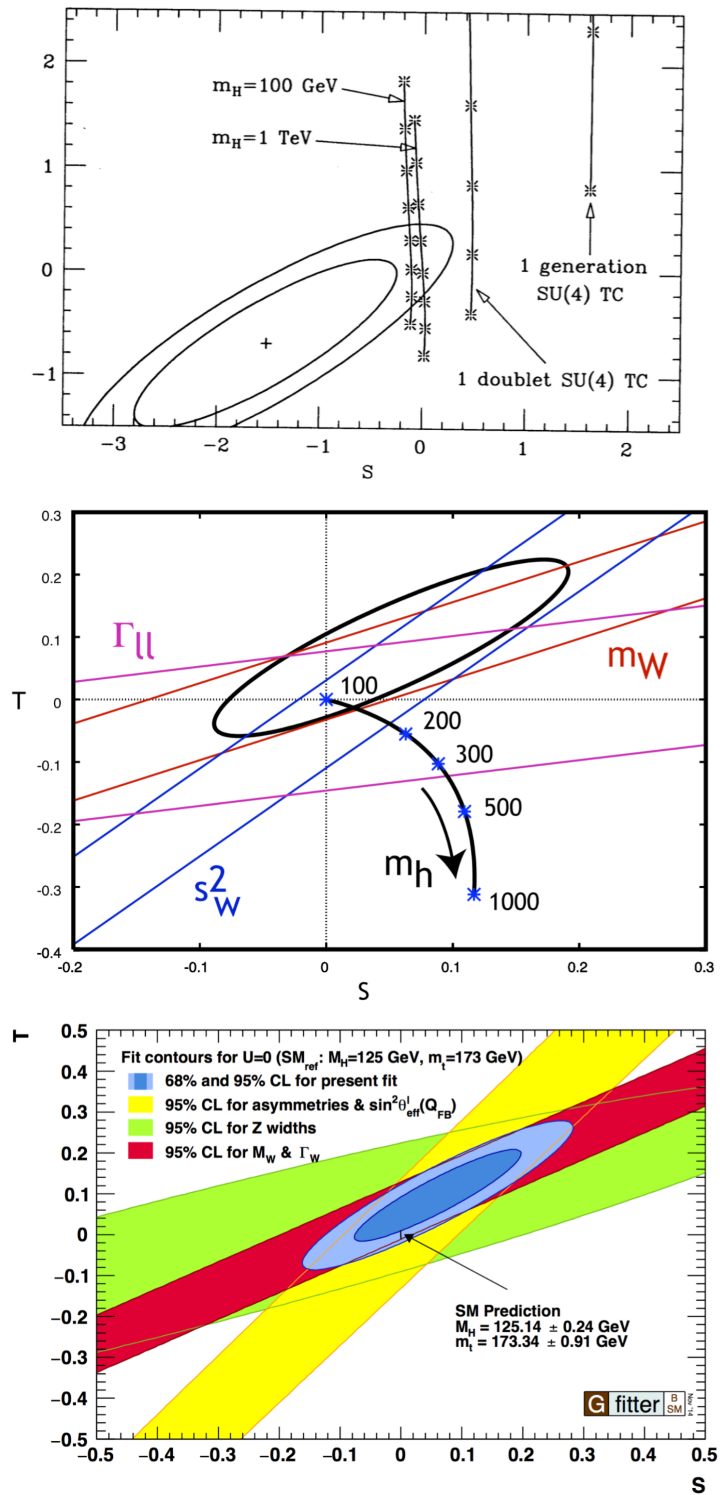
In its rest frame, with  $p^\mu = (m, 0, 0, 0)^\mu$ , a massive vector boson has 3 polarization states, corresponding to the 3 orthogonal spacelike vectors

$$\begin{aligned} \epsilon_+^\mu &= \frac{1}{\sqrt{2}}(0, 1, +i, 0)^\mu \\ \epsilon_0^\mu &= (0, 0, 0, 1)^\mu \\ \epsilon_-^\mu &= \frac{1}{\sqrt{2}}(0, 1, -i, 0)^\mu . \end{aligned} \quad (120)$$

These vectors represent the states of the vector boson with definite angular momentum  $J^3 = +1, 0, -1$ .

Now boost along the  $\hat{3}$  axis to high energy,  $p^\mu = (E, 0, 0, p)^\mu$ . The boosts of the polarization vectors in (120) are

$$\begin{aligned} \epsilon_+^\mu &= \frac{1}{\sqrt{2}}(0, 1, +i, 0)^\mu \\ \epsilon_0^\mu &= \left(\frac{p}{m}, 0, 0, \frac{E}{m}\right)^\mu \\ \epsilon_-^\mu &= \frac{1}{\sqrt{2}}(0, 1, -i, 0)^\mu . \end{aligned} \quad (121)$$



**Fig. 19:** Allowed domain for the  $S, T$  parameters in three different eras: in 1991, before the discovery of the top quark [33]; in 2008, before the discovery of the Higgs boson; today [34].

The transverse polarization vectors  $\epsilon_+$ ,  $\epsilon_-$  are left unchanged by the boost. However, for the longitudinal polarization vector  $\epsilon_0$ , the components grow without bound. At very high energy

$$\epsilon_0^\mu \rightarrow \frac{p^\mu}{m}. \quad (122)$$

Another way to understand this is to recall that the polarization sum for a massive vector boson is written covariantly as

$$\sum_i \epsilon_i^\mu \epsilon_j^\nu = - \left( g^{\mu\nu} - \frac{p^\mu p^\nu}{m^2} \right). \quad (123)$$

In the rest frame of the vector boson, this is the projection onto the 3 spacelike polarization vectors. For a highly boosted vector boson, however, the second term in parentheses in this expression has matrix elements that grow large in the same way as (122).

This potentially leads to very large contributions to amplitudes for high-energy vector bosons, even threatening violation of unitarity. An example of this problem is found in the production of a pair of massive vector bosons in  $e^+e^-$  annihilation. The amplitude for production of a pair of scalar bosons in QED is

$$i\mathcal{M}(e^+e^- \rightarrow \phi^+\phi^-) = -i \frac{e^2}{s} (2E) \sqrt{2} \epsilon_- \cdot (k_- - k_+), \quad (124)$$

where  $k_+$ ,  $k_-$  are the scalar particle momenta. In  $e^+e^- \rightarrow W^+W^-$ , we might expect that this formula generalizes to

$$i\mathcal{M}(e^+e^- \rightarrow \phi^+\phi^-) = i \frac{e^2}{s} (2E) \sqrt{2} \epsilon_- \cdot (k_+ - k_-) \epsilon^*(k_+) \cdot \epsilon^*(k_-). \quad (125)$$

where  $\epsilon(k_+)$ ,  $\epsilon(k_-)$  are the  $W^+$  and  $W^-$  polarization vectors. For longitudinally polarized  $W$  bosons, this extra factor becomes

$$\frac{k_+ \cdot k_-}{m_W^2} = \frac{s - 2m_W^2}{2m_W^2} \quad (126)$$

at high energy. This growth of the production amplitude really would violate unitarity.

This raises the question: Are the enhancements due to  $\epsilon_0 \sim p/m$  at high energy actually present? Do these enhancements appear always, sometimes, or never?

The answer to this question is given by the Goldstone Boson Equivalence Theorem (GBET) of Cornwall, Levin, and Tiktopoulos and Vayonakis [5, 6].

When a  $W$  boson or other gauge boson acquires mass through the Higgs mechanism, this boson must also acquire a longitudinal polarization state that does not exist for a massless gauge boson. The extra degree of freedom is obtained from the symmetry-breaking Higgs field, for which a Goldstone boson is gauged away. When the  $W$  is at rest, it is not so clear which polarization state came from the Higgs field. However, for a highly boosted  $W$  boson, there is a clear distinction between the transverse and longitudinal polarization states. The GBET states, in the limit of high energy, the couplings of the longitudinal polarization state are precisely those of the original Goldstone boson,

$$\mathcal{M}(X \rightarrow Y + W_0^+(p)) = \mathcal{M}(X \rightarrow Y + \pi^+(p)) \left( 1 + \mathcal{O}\left(\frac{m_W}{E_W}\right) \right) \quad (127)$$

The proof is too technical to give here. Some special cases are analyzed in Chapter 21 of [7]. A very elegant and complete proof, which accounts for radiative corrections and includes the possibility of multiple boosted vector bosons, has been given by Chanowitz and Gaillard in [35]. Both arguments rely in an essential way on the underlying gauge invariance of the theory.

In the rest of this section, I will present three examples that illustrate the various aspects of this theorem.

## 5.2 $W$ polarization in top quark decay

The first application is the theory of the polarization of the  $W$  boson emitted in top quark decay,  $t \rightarrow bW^+$ .

It is straightforward to compute the rates for top quark decay to polarized  $W$  bosons. These rates follow directly from the form of the  $V-A$  coupling. The matrix element is

$$i\mathcal{M} = i\frac{g}{\sqrt{2}}u_L^\dagger(b)\bar{\sigma}^\mu u_L(t)\epsilon_\mu^*. \quad (128)$$

In evaluating this matrix element, I will ignore the  $b$  quark mass, a very good approximation. I will use coordinates in which the  $t$  quark is at rest, with spin orientation given by a 2-component spinor  $\xi$ , and the  $W^+$  is emitted in the  $\hat{3}$  direction. The  $b$  quark is left-handed and moves in the  $-\hat{3}$  direction. Then the spinors are

$$u_L(b) = \sqrt{2E_b} \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad u_L(t) = \sqrt{m_t}\xi. \quad (129)$$

For a  $W_-^+$ ,

$$\bar{\sigma} \cdot \epsilon_-^* = \frac{1}{\sqrt{2}}(\sigma^1 + i\sigma^2) = \sqrt{2}\sigma^+ \quad (130)$$

and so the amplitude is

$$i\mathcal{M} = ig\sqrt{2m_tE_b}\xi_2. \quad (131)$$

with, from 2-body kinematics,  $E_b = (m_t^2 - m_w^2)/2m_t$ . For a  $W_+^+$ , the sigma matrix structure is proportional to  $\sigma^-$  and the amplitude vanishes. For a  $W_0^+$ ,

$$\bar{\sigma} \cdot \epsilon_0^* = -\frac{p + E\sigma^3}{m_W} \quad (132)$$

and the amplitude is

$$i\mathcal{M} = ig\sqrt{2m_tE_b} \frac{m_t}{m_W} \xi_1. \quad (133)$$

Squaring these matrix elements, averaging over the  $t$  spin direction, and integrating over phase space, we find

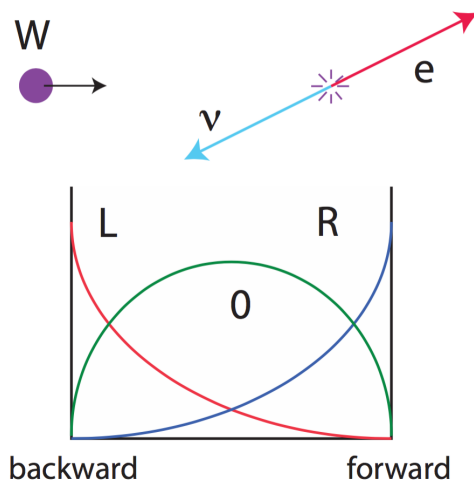
$$\begin{aligned} \Gamma(t \rightarrow bW_-^+) &= \frac{\alpha_w}{8} m_t \left(1 - \frac{m_W^2}{m_t^2}\right)^2 \\ \Gamma(t \rightarrow bW_+^+) &= 0 \\ \Gamma(t \rightarrow bW_0^+) &= \frac{\alpha_w}{8} m_t \left(1 - \frac{m_W^2}{m_t^2}\right)^2 \cdot \frac{m_t^2}{2m_W^2}. \end{aligned} \quad (134)$$

From these formulae, we see that the fraction of longitudinally polarized  $W$  bosons is

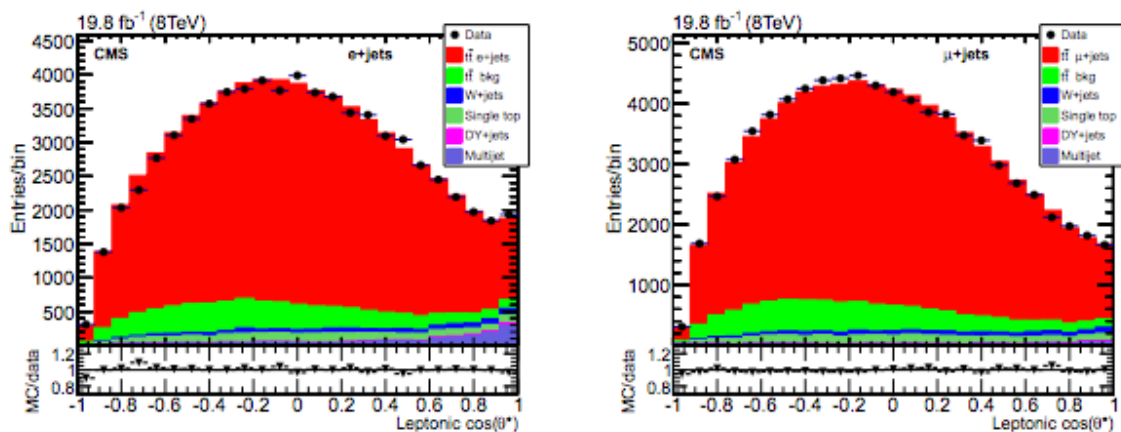
$$\frac{\Gamma(t \rightarrow bW_0^+)}{\Gamma(t \rightarrow bW^+)} = \frac{m_t^2/2m_W^2}{1 + m_t^2/2m_W^2} \approx 70\%. \quad (135)$$

The polarization of  $W$  bosons in  $t$  decay can be measured by reconstructing full  $pp \rightarrow t\bar{t} \rightarrow \ell\nu + 4$  jet events. Beginning in the  $t$  rest frame, we boost the leptonically decaying  $W$  to rest. The angular distribution of the decay lepton in the  $W$  frame is then given by for the three polarization states by

$$\frac{d\Gamma}{d\cos\theta_*} \sim \begin{cases} (1 + \cos\theta_*)^2 & + \\ \sin^2\theta_*/2 & 0 \\ (1 - \cos\theta_*)^2 & - \end{cases}, \quad (136)$$



**Fig. 20:** Angular distributions of  $\cos \theta_*$  in  $W$  boson decay for each of the three possible polarization states.



**Fig. 21:** CMS measurement of the  $\cos \theta_*$  distribution in top decay, compared to a simulation that represents the SM expectation [36].

where  $\theta_*$  is the angle between the boost direction and the lepton direction. These angular distributions, which are also a consequence of  $V-A$ , are illustrated in Fig. 20. The actual distributions measured in hadron collisions are distorted from the idealized ones, since leptons with  $\cos \theta_*$  near  $-1$ , which implies low lab-frame energy, have low acceptance. Figure 21 shows the  $\cos \theta_*$  distribution measured by the CMS experiment at the LHC and indicates an excellent agreement with the SM prediction [36].

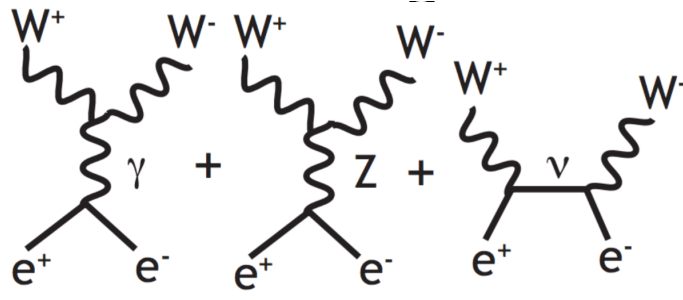
An interesting feature of this prediction is the form of the amplitude (133). This amplitude is enhanced by a factor  $m_t/m_W$ , just as we might have expected from (122). This behavior can be understood using the GBET. According to the GBET, we should find

$$i\mathcal{M}(t \rightarrow bW_0^+) \rightarrow i\mathcal{M}(t \rightarrow b\pi^+). \quad (137)$$

The amplitude for emission of a Higgs boson should be proportional to the top quark Yukawa coupling  $y_t$ , given by

$$m_t = \frac{y_t v}{\sqrt{2}}. \quad (138)$$

So the GBET predicts that the rate for  $t$  decay to a longitudinal  $W$  should be larger than the rate to a



**Fig. 22:** Feynman diagrams for the process  $e^+e^- \rightarrow W^+W^-$ .

transverse  $W$  by the factor

$$\frac{y_t^2}{g^2} = \frac{2m_t^2/v^2}{4m_W^2/v^2} = \frac{m_t^2}{2m_W^2}, \quad (139)$$

and this is exactly what we found in the explicit calculation.

### 5.3 High energy behavior in $e^+e^- \rightarrow W^+W^-$

The next example to study is the high energy behavior of the reaction

$$e^+e^- \rightarrow W^+W^-. \quad (140)$$

I argued earlier that the amplitude for this process cannot show the enhancement (122), at least in the most straightforward way, since this would lead to an amplitude that violates unitarity. Indeed, the prediction of the GBET is that

$$\mathcal{M}(e^+e^- \rightarrow W_0^+W_0^-) \rightarrow \mathcal{M}(e^+e^- \rightarrow \pi^+\pi^-). \quad (141)$$

Using (36), the high-energy limit of  $SU(2) \times U(1)$ , and the quantum numbers of the Higgs field  $(I, Y) = (\frac{1}{2}, \frac{1}{2})$ , we can readily work out that the right-hand side of (141) is, for an  $e_R^-e_L^+$  initial state,

$$i\mathcal{M} = -i(2E)\sqrt{2} \epsilon_+ \cdot (k_- - k_+) \cdot \frac{e^2}{2c_w^2} \frac{1}{s}, \quad (142)$$

and for an  $e_L^-e_R^+$  initial state,

$$i\mathcal{M} = -i(2E)\sqrt{2} \epsilon_- \cdot (k_- - k_+) \cdot \left( \frac{e^2}{4c_w^2} \frac{1}{s} + \frac{e^2}{4c_w^2} \frac{1}{s} \right), \quad (143)$$

where  $k_-$  and  $k_+$  are the final-state momenta. So it must be that the expression we guessed in (125) is either incorrect or is cancelled by other factors.

In the SM, the complete tree level amplitude for  $e^+e^- \rightarrow W^+W^-$  is given by a sum of three diagrams, shown in Fig. 22. It will be instructive to work out the sum of diagrams in a careful way. I will do this first for the initial state  $e_R^-e_L^+$ , for which the neutrino diagram does not appear.

The full matrix element involves the Yang-Mills vertex for the  $WW\gamma$  and  $WWZ$  interactions. It is

$$i\mathcal{M} = (-ie)(ie)2E\sqrt{2} \epsilon_{+\mu} \left[ \frac{-i}{s} + \frac{-s_w^2 c_w}{s_w c_w s_w} \frac{-i}{s - m_Z^2} \right] \cdot \left[ \epsilon^*(-) \epsilon^*(+) (k_- - k_+)^{\mu} + (-q - k_-) \epsilon^*(+) \epsilon^{*\mu}(-) + (k_+ + q) \epsilon^*(-) \epsilon^{*\mu}(+) \right],$$

(144)

where  $q = k_- + k_+$  and, in the second line,  $\epsilon^*(-)$  and  $\epsilon^*(+)$  are the  $W$  polarizations. To evaluate the high-energy limit for longitudinally polarized  $W$  bosons, send

$$\epsilon^*(-) \rightarrow \frac{k_-}{m_W} \quad \epsilon^*(+) \rightarrow \frac{k_+}{m_W} . \quad (145)$$

Then the second term in brackets becomes

$$\begin{aligned} & \frac{1}{m_W^2} \left[ k_- k_+ (k_- - k_+)^{\mu} - 2k_- k_+ k_-^{\mu} + 2k_+ k_- k_+^{\mu} \right] \\ &= -\frac{k_- k_+}{m_W^2} (k_- - k_+)^{\mu} = -\frac{s - 2m_W^2}{2m_W^2} (k_- - k_+)^{\mu} . \end{aligned} \quad (146)$$

This expression has the enhancement (126). However, there is a nice cancellation in the first term in brackets,

$$\left[ \frac{-i}{s} - \frac{-i}{s - m_Z^2} \right] = \frac{i m_Z^2}{s(s - m_Z^2)} . \quad (147)$$

Assembling the pieces and using  $m_W^2 = m_Z^2 c_w^2$ , we find

$$i\mathcal{M} = ie^2 2E\sqrt{2} \epsilon_{+\mu} (k_- - k_+)^{\mu} \left( -\frac{s - 2m_W^2}{2c_w^2 s(s - m_Z^2)} \right) , \quad (148)$$

which indeed agrees with (142) in the high energy limit.

For the  $e_L^- e_R^+$  case, the  $\gamma$  and  $Z$  diagrams do not cancel, and so the neutrino diagram is needed. The first two diagrams contribute

$$\begin{aligned} i\mathcal{M} &= (-ie)(ie)2E\sqrt{2} \epsilon_{+\mu} \left[ \frac{-i}{s} + \frac{(1/2 - s_w^2) c_w}{s_w c_w} \frac{-i}{s_w s - m_Z^2} \right] \\ &\cdot \left[ \epsilon^*(-)\epsilon^*(+)(k_- - k_+)^{\mu} + (-q - k_-)\epsilon^*(+)\epsilon^{*\mu}(-) + (k_+ + q)\epsilon^*(-)\epsilon^{*\mu}(+) \right] , \end{aligned} \quad (149)$$

After the reductions just described, there is a term in the high-energy behavior that does not cancel,

$$\begin{aligned} i\mathcal{M} &= ie^2 2E\sqrt{2} \epsilon_{-\mu} (k_- - k_+)^{\mu} \left[ \frac{1}{2s_W^2 s} \right] \left( -\frac{s}{2m_W^2} \right) \\ &= \frac{ie^2}{4s_w^2} 2E\sqrt{2} \epsilon_{-\mu} (k_- - k_+)^{\mu} \frac{1}{m_W^2} . \end{aligned} \quad (150)$$

We must add to this the neutrino diagram, which contributes

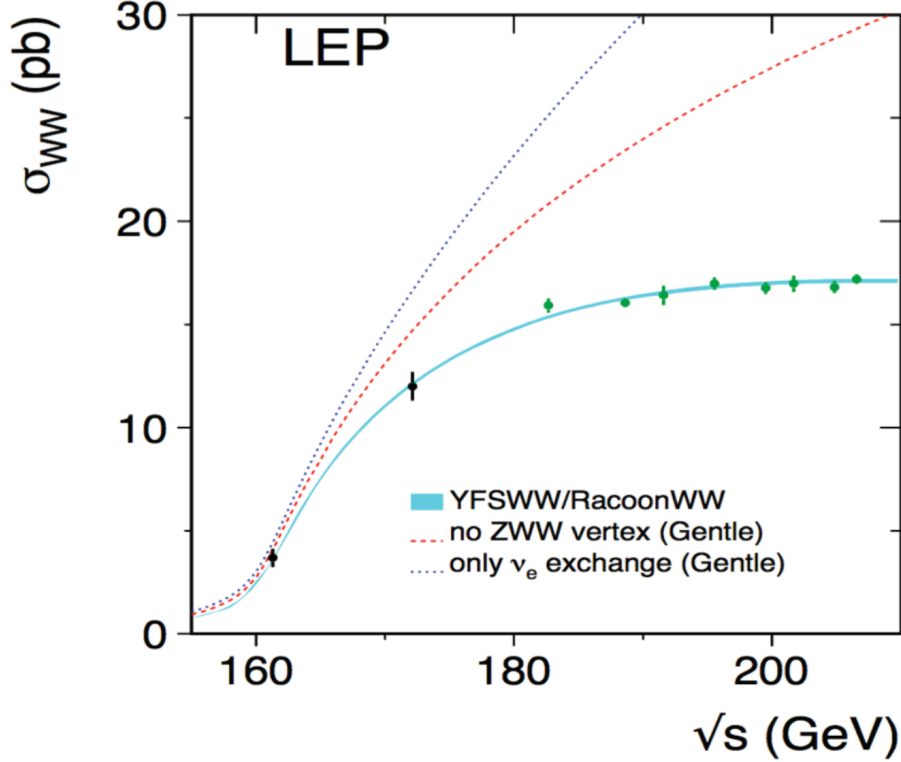
$$i\mathcal{M} = \left( i\frac{g}{\sqrt{2}} \right)^2 v_R(\bar{p})^{\dagger} \bar{\sigma} \cdot \epsilon^*(+) \frac{i\sigma \cdot (p - k_-)}{(p - k_-)^2} \bar{\sigma} \cdot \epsilon^*(-) u_L(p) . \quad (151)$$

Substituting  $\epsilon^*(-) \rightarrow k_-/m_W$ , the second half of this formula becomes

$$\frac{i\sigma \cdot (p - k_-)}{(p - k_-)^2} \bar{\sigma} \cdot \frac{k_-}{m_W} u(p) . \quad (152)$$

Since  $\bar{\sigma} \cdot p u_L(p) = 0$ , this can be written

$$\frac{i\sigma \cdot (p - k_-)}{(p - k_-)^2} \bar{\sigma} \cdot \frac{(k_- - p)}{m_W} u(p) = -\frac{i}{m_W^2} u(p) . \quad (153)$$



**Fig. 23:** Measurement of  $\sigma(e^+e^- \rightarrow W^+W^-)$  from the four LEP experiments, from [20].

Sending  $\epsilon^*(+) \rightarrow k_+/m_W = ((p + \bar{p})/2 + (k_+ - k_-)/2)/m_W$  and using  $(\bar{\sigma} \cdot p)u_L = v_R^\dagger(\bar{\sigma} \cdot \bar{p}) = 0$ , we finally find

$$i\mathcal{M} = -\frac{ie^2}{2s_w^2} 2E\sqrt{2} \epsilon_{-\mu} \frac{1}{2}(k_- - k_+)^\mu \frac{1}{m_W^2}, \quad (154)$$

and this indeed cancels the high-energy behavior (150) from the  $\gamma$  and  $Z$  diagrams. To fully verify (143), we would need to carry out this calculation more exactly to pick up all subleading terms at high energy. It does work out correctly, as was first shown by Alles, Boyer, and Buras [37].

The cross section for  $e^+e^- \rightarrow W^+W^-$  was measured by the LEP experiments. The result is shown in Fig. 23 [20]. The lowest, solid line is the prediction of the SM, including one-loop radiative corrections. It is in excellent agreement with the measurements. The upper curves show the effect of omitting, first, the  $Z$  diagram and, second, both the  $\gamma$  and  $Z$  diagrams. Apparently, the cancellation I have demonstrated here is important not only at very high energy but even in the qualitative behavior of the cross section quite close to threshold.

#### 5.4 Parametrizing corrections to the Yang-Mills vertex

The cancellation described in the previous section clearly requires the precise structure of the Yang-Mills vertex that couples three vector bosons. Before the LEP measurements, when the gauge boson nature of the  $W$  and  $Z$  was less clear, theorists suggested that the  $WW\gamma$  and  $WWZ$  vertices might be modified from the Yang-Mills form, and that such modifications could be tested by measurements of  $W$  reactions at high energy.

The most general Lorentz-invariant,  $CP$  conserving  $WW\gamma$  vertex in which the photon couples to



a conserved current has the form [38]

$$\Delta\mathcal{L} = e \left[ ig_{1A} A_\mu (W^-_\nu W^{+\mu\nu} - W^+_\nu W^{-\mu\nu}) + i\kappa_A A_{\mu\nu} W^-_\mu W^+_\nu + i\lambda_A \frac{1}{m_W^2} W^-_{\lambda\mu} W^{+\mu\nu} A_\nu^\lambda \right]. \quad (155)$$

In this formula, for each vector field,  $V_{\mu\nu} = (\partial_\mu V_\nu - \partial_\nu V_\mu)$ . We can write a similar generalization of the SM  $WWZ$  vertex, with parameters  $g_{1Z}$ ,  $\kappa_Z$ ,  $\lambda_Z$  and overall coupling  $ec_w/s_w$ . The choice

$$g_{1\gamma} = g_{1Z} = \kappa_A = \kappa_Z = 1 \quad \lambda_A = \lambda_Z = 0 \quad (156)$$

gives the SM coupling. If we relax the assumption of CP conservation, several more terms can be added.

It was quickly realized that any changes to the SM vertex produce extra contributions to the  $W$  production amplitudes that are enhanced by the factor  $s/m_W^2$ . In view of the discussion earlier in this section, this is no surprise. If the additional terms violate the gauge invariance of the theory, the GBET will not be valid, and the cancellations it requires will not need to occur. However, this idea would seem to be already excluded by the strong evidence from the precision electroweak measurements that the  $W$  and  $Z$  are the vector bosons of a gauge theory.

Still, there is a way to modify the  $WW\gamma$  and  $WWZ$  vertices in a way that is consistent with gauge invariance. It is certainly possible that there exist new heavy particles that couple to the gauge bosons of the SM. The quantum effects of these particles can be described as a modification of the SM Lagrangian by the addition of new gauge-invariant operators. This approach to the parametrization of new physics effects has become known as Effective Field Theory (EFT). The SM already contains the most general  $SU(2) \times U(1)$ -invariant operators up to dimension 4, but new physics at high energy can add higher-dimension operators, beginning with dimension 6.

There are many dimension 6 operators that can be added to the SM. Even for 1 generation of fermions, there are 84 independent dimension 6 operators, of which 59 are baryon-number and CP-conserving [39]. The theory of these operators has a complexity that I do not have room to explain here. It is possible to make many different choices for the basis of these operators, using the fact that combinations of these operators are set to zero by the SM equations of motion. The theory of EFT modifications of the SM is reviewed in [40] and, in rather more detail, in [41]. I will give only a simple example here.

Consider, then, adding to the SM the dimension-6 operators

$$\Delta\mathcal{L} = \frac{c_T}{2v^2} \Phi^\mu \Phi_\mu + \frac{4gg'}{m_W^2} \Phi^a W_{\mu\nu}^a B^{\mu\nu} + \frac{g^3 c_{3W}}{m_W^2} \epsilon^{abc} W_{\mu\nu}^a W^{b\nu\rho} W^{c\rho\mu}, \quad (157)$$

where, in this formula,  $W_{\mu\nu}^a$  and  $B_{\mu\nu}$  are the  $SU(2)$  and  $U(1)$  field strengths and  $\Phi_\mu$ ,  $\Phi^a$  are bilinears in the Higgs field,

$$\Phi_\mu = \varphi^\dagger D_\mu \varphi - (D_\mu \varphi)^\dagger \varphi \quad \Phi^a = \varphi^\dagger \frac{\sigma^a}{2} \varphi. \quad (158)$$

It can be shown that these shift the parameters of the  $WW\gamma$  and  $WWZ$  couplings to

$$\begin{aligned} g_{1Z} &= 1 + \left[ \frac{c_T}{2(c_w^2 - s_w^2)} - \frac{8s_w^2 c_{WB}}{c_w^2 (c_w^2 - s_w^2)} \right] \\ \kappa_A &= 1 - 4c_{WB} \\ \lambda_A &= -6g^2 c_{3W} \end{aligned} \quad (159)$$

The parameter  $g_{1A} = 1$  is not shifted; this is the electric charge of the  $W$  boson. The remaining two parameters obey

$$\kappa_Z = g_{1Z} - \frac{s_w^2}{c_w^2} (\kappa_A - 1) \quad \lambda_Z = \lambda_A. \quad (160)$$

It can be shown that the relations (160) are maintained for any set of dimension-6 perturbations of the SM. They may be modified by dimension-8 operators.

Dimension-6 operators also contribute to the  $S$  and  $T$  parameters discussed at the end of the previous section. From the perturbation (157),

$$\begin{aligned}\alpha S &= 32s_w^2 c_{WB} \\ \alpha T &= c_T\end{aligned}\tag{161}$$

Given that EFT is based on gauge-invariant Lagrangian, this formalism for parametrizing new physics can be worked out explicitly in great detail. QCD and electroweak radiative corrections can be included. The higher-dimension operators in the EFT must of course be renormalized according to some scheme, and the detailed formulae will depend on the scheme.

A dimension-6 operator has a coefficient with the units of  $(\text{GeV})^{-2}$ . Thus, the effects of such operators are suppressed by one factor of  $s/M^2$ , where  $M$  is then mass scale of new particles. Contributions from dimension-8 operators suppressed by  $(s/M^2)^2$ , and similarly for operators of still higher dimension. So, an analysis that puts constraints on dimension-6 operators, ignoring the effects of dimension-8 operators is properly valid only when  $s/M^2 \ll 1$ .

As a corollary to this point, I call your attention to a Devil's bargain that arises frequently in tests of the structure of  $W$  and  $Z$  vertices at hadron colliders. In  $pp$  collisions, the parton center of mass energy  $\hat{s}$  varies over a wide range. There is always a region of phase space where  $\hat{s}$  becomes extremely large. This is the region that has the greatest sensitivity to higher-dimension operators. It is tempting to apply event selections that emphasize this region to obtain the strongest possible limits.

However, this is exactly the region where operators of dimension 8 and higher might also be important. In many models, these give negative contribution. Then a parametrization that uses only dimension-6 operators leads to limits on their coefficients that are stronger than the limits that would be obtained in a more complete theory.

The question of how to interpret limits on dimension-6 EFT coefficients is now hotly debated in the literature. My personal position is on one extreme, that only analyses in which  $\hat{s}/M^2 \ll 1$  for all events included in the analysis should be trusted. The authors of [42] advocate for a much more aggressive approach. Experimenters who quote such limits should study this issue carefully.

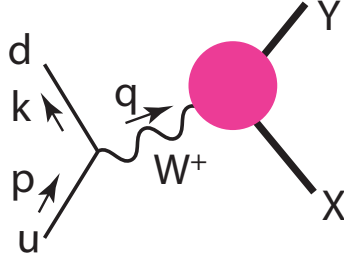
On the other hand, the SM itself makes precise predictions in all regions of  $\hat{s}$ . Your first priority should be to discover a deviation from these predictions. If you are able to demonstrate a substantial deviation from the SM predictions in any region of phase space, we can all have fun quarreling about the interpretation of this result.

## 5.5 $W$ parton distributions

As a final topic in this section, I will discuss a situation in which the GBET might be expected to apply, but it does not. This involves processes in which a  $W$  boson is radiated from a quark or lepton with small transverse momentum relative to the fermion direction. In QCD, the collinear radiation of gluons from initial quarks is essential in creating the observed quark and gluon parton distributions. In Section 4.2, we saw that collinear radiation of photons from initial electrons and positrons is also an important effect that makes qualitative changes in the  $Z$  resonance line shape. In this section, I will present the analogous theory for collinear  $W$  boson emission [43]. I will carry out the analysis for quark initial states, but the same theory applies to electron and positron initial states.

For definiteness, consider the following setup: An initial  $u$  quark, with momentum  $p$ , emits an almost collinear  $W^+$  boson, with momentum  $q$ ,

$$u(p) \rightarrow d(k) + W^+(q) .\tag{162}$$



**Fig. 24:** Kinematics of a process in which a  $W$  is emitted collinearly from a quark and then initiates a large-momentum-transfer reaction.

The  $W$  boson must be off-shell. This emission will be part of a process shown in Fig. 24, in which the virtual  $W$  collides with a parton from the other proton to initiate a hard-scattering reaction. An important class of processes of this type is  $WW$  scattering, including the reaction  $W^+W^- \rightarrow h$  that we will discuss in Section 6.2.

For  $W$  reactions that involve the Higgs boson, it will be important to have  $W$  bosons with longitudinal polarization. According to the GBET, a longitudinally polarized  $W$  boson should have a coupling equal to that of the corresponding Goldstone boson  $\pi^+$  from the Higgs sector. Then the study of high energy  $W$  boson reactions allows us to directly measure the strength of Higgs boson interactions. However, it is not clear that it is possible to radiate longitudinally polarized  $W$  bosons from initial quarks. A  $\pi^+$  couples to a light fermion with its Higgs Yukawa coupling, that is, negligibly, the radiation of longitudinally polarized  $W$  bosons would seem to be forbidden by the GBET.

To understand the correct story, we must compute the  $u \rightarrow Wd$  emission amplitude explicitly. In this calculation, I will take the  $W$  boson to be emitted approximately collinearly with the  $u$  quark. The analysis is very similar to calculation of the Altarelli-Parisi splitting functions that you will find, for example, in Chapter 17 of [7]. I will assume that the  $W$  has  $p_T \sim m_W \ll p_{\parallel}$ .

First, I write the momentum vectors for the quarks, taking the  $u$  quark to move in the  $\hat{3}$  direction and the  $d$  quark to carry away an energy fraction  $(1-x)$  and to have a small transverse momentum,

$$\begin{aligned} p &= (E, 0, 0, E) \\ k &= ((1-z)E, -p_T, 0, (1-z)E - \frac{p_T^2}{2(1-z)E}). \end{aligned} \quad (163)$$

The momentum  $k$  is on-shell to order  $p_T^2$ . The  $W$  momentum vector is then determined by momentum conservation

$$q = (zE, p_T, 0, zE + \frac{p_T^2}{2(1-z)E}). \quad (164)$$

The denominator of the  $W$  propagator is then

$$q^2 - m_W^2 = -p_T^2 - \frac{z}{(1-z)}p_T^2 - m_W^2 = -\left(\frac{p_T^2}{1-z} + m_W^2\right). \quad (165)$$

Next, we compute the matrix elements for  $W$  emission

$$i\mathcal{M} = ig u_L^\dagger(k) \bar{\sigma} \cdot \epsilon_W^* u_L(p) \quad (166)$$

to first order in  $(p_T, m_W)$ . The explicit form of the spinors is

$$u_L(k) = \sqrt{2(1-z)E} \begin{pmatrix} p_T/2(1-z) \\ 1 \end{pmatrix} \quad u_L(p) = \sqrt{2E} \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (167)$$

The  $W$  polarization vectors are

$$\epsilon_{\pm}^{*\mu} = (0, 1, \mp i, -p_T/zE)^\mu / \sqrt{2} \quad (168)$$

for the transverse polarizations, and

$$\epsilon_0^{*\mu} = (q, p_T, 0, zE)^\mu / m_W \quad (169)$$

for the longitudinal polarization state. In this formula

$$q = [(zE)^2 - m_W^2]^{1/2} = zE - \frac{m_W^2}{2zE} \quad (170)$$

Then

$$\begin{aligned} \bar{\sigma} \cdot \epsilon_+^* &= \frac{1}{\sqrt{2}} \begin{pmatrix} -p_T/zE & 0 \\ 2 & p_T/zE \end{pmatrix} \\ \bar{\sigma} \cdot \epsilon_-^* &= \frac{1}{\sqrt{2}} \begin{pmatrix} -p_T/zE & 2 \\ 0 & p_T/zE \end{pmatrix} \\ \bar{\sigma} \cdot \epsilon_0^* &= \frac{1}{m_W} \begin{pmatrix} q + zE & p_T \\ p_T & q - zE \end{pmatrix} \end{aligned} \quad (171)$$

With these ingredients, it is straightforward to work out the matrix elements for the three  $W$  polarization states,

$$i\mathcal{M}(u \rightarrow dW^+) = ig \cdot \begin{cases} \sqrt{1-z} p_T/z & + \\ \sqrt{1-z} p_T/z(1-z) & - \\ -\sqrt{1-z} m_W/\sqrt{2}z & 0 \end{cases} \quad (172)$$

We can convert these expressions to cross sections for complete  $W$ -induced processes. The cross section for a process  $uX \rightarrow dY$ , in the approximation in which the  $W$  is almost on shell, is given by

$$\begin{aligned} \sigma &= \frac{1}{2s} \int \frac{d^3k}{(2\pi)^3 2k} \int d\Pi_Y (2\pi)^4 \delta^{(4)}(p + p_X - k - p_Y) \\ &\quad \left| \mathcal{M}(u \rightarrow dW^+) \frac{1}{q^2 - m_W^2} \mathcal{M}(W^+ X \rightarrow Y) \right|^2 \end{aligned} \quad (173)$$

In the collinear kinematics, with  $\hat{s} = zs$

$$\frac{1}{2s} \int \frac{d^3k}{(2\pi)^3 2k} = \frac{1}{2\hat{s}/z} \int \frac{dz E d^2 p_T}{16\pi^3 E(1-z)} = \frac{1}{2\hat{s}} \int \frac{dz dp_T^2 \pi}{16\pi^3} \frac{z}{(1-z)} \quad (174)$$

Then, also using (165), (173) simplifies to

$$\begin{aligned} \sigma &= \int dz \int \frac{dp_T^2}{(4\pi)^2} \frac{z}{(1-z)} |\mathcal{M}(u \rightarrow dW^+)|^2 \frac{1}{p_T^2/(1-z) + m_W^2} \\ &\quad \cdot \frac{1}{2\hat{s}} \int d\Pi_Y (2\pi)^4 \delta^{(4)}(q + p_X - p_Y) |\mathcal{M}(W^+ X \rightarrow Y)|^2 \end{aligned} \quad (175)$$

The last line of (175) is  $\sigma(W^+(q) + X \rightarrow Y)$ . Then (175) has the form of a parton model cross section

$$\sigma(uX \rightarrow dY) = \int dz f_{W \leftarrow u}(z) \sigma(W^+ X \rightarrow Y) \quad (176)$$

where  $f_{W \leftarrow u}(z)$  is the parton distribution for a  $W$  boson in the  $u$  quark,

$$f_{W \leftarrow u}(z) = \int \frac{dp_T^2}{(4\pi)^2} \frac{z}{(1-z)} \frac{(1-z)^2}{(p_T^2 + (1-z)m_W^2)^2} |\mathcal{M}(u \rightarrow dW^+)|^2. \quad (177)$$

We can evaluate this parton distribution for each  $W$  polarization state by using the formula (172). The result is

$$\begin{aligned} f_{W-}(z) &= \frac{\alpha_2}{4\pi} \int \frac{dp_T^2 p_T^2}{(p_T^2 + (1-z)m_W^2)^2} \frac{1}{z} \\ f_{W+}(z) &= \frac{\alpha_2}{4\pi} \int \frac{dp_T^2 p_T^2}{(p_T^2 + (1-z)m_W^2)^2} \frac{(1-z)^2}{z} \\ f_{W0}(z) &= \frac{\alpha_2}{8\pi} \int \frac{dp_T^2 m_W^2}{(p_T^2 + (1-z)m_W^2)^2} \frac{(1-z)^2}{z} \end{aligned} \quad (178)$$

For the transverse polarizations, we find a result very similar to the Altarelli-Parisi splitting function for collinear gluon emission,

$$f_{WT}(z) = \frac{\alpha_w}{4\pi} \frac{1 + (1-z)^2}{z} \cdot \log \frac{Q^2}{m_W^2}, \quad (179)$$

where  $Q^2$  is the upper limit of the  $p_T^2$  integral, which is set by the momentum transfer in the hard reaction.

For the longitudinal  $W$  polarization, the story is different. The integral over  $p_T$  is convergent, so that the  $p_T$  is restricted to the region  $p_T \sim m_W$ . In this regime, as we see explicitly, longitudinal  $W$  bosons can be produced with coupling strength  $g$ . Apparently, in this process, the error term in the GBET is actually  $\mathcal{O}(m_W/p_T)$ , which is consistent with (127) but, still, larger than we might expect. The reduction of the longitudinal  $W$  boson to a Higgs boson then is not accurate in the region  $p_T \sim m_W$ , though it does apply—and cuts off the amplitude—when  $p_T \gg m_W$ .

When we perform the convergent integral over  $p_T$ , we find that the parton distribution for  $W_0$  is substantial [43],

$$f_{W0}(z) = \frac{\alpha_w}{8\pi} \frac{1-z}{z}. \quad (180)$$

Then the proton does contain longitudinal  $W$  bosons, which can induce Higgs sector reactions when this proton collides with another proton at high energy. The collinear longitudinal  $W$  bosons have  $p_T \sim m_W$  but not higher, a kinematic feature that can be used to suppress backgrounds from reactions involving transversely-polarized  $W$  bosons.

## 6 The Standard Model theory of Higgs boson decays

There remains one heavy particle of the SM that we have not yet discussed, the Higgs boson. The Higgs boson has a central role in the structure of the weak interactions. Its field is the agent that breaks the  $SU(2) \times U(1)$  symmetry and generates the masses of all quarks, leptons, and vector bosons. This at the same time forms a unified picture of the electroweak interactions as we have studied them so far and also points to new mysteries whose explanations are still to be found.

The best way to enter a discussion of the Higgs boson is to understand thoroughly the predictions for the properties of this particle given by the SM. The Higgs sector involves one more parameter of the SM beyond those we have discussed already, the Higgs field self-coupling  $\lambda$ . However, this coupling is fixed by the measurement of the Higgs boson mass. Thus, the SM makes precise predictions for all of the Higgs boson cross sections and branching fractions. These predictions provide a starting point for any discussion of the properties of the Higgs boson in model that generalize the SM. An excellent reference on the theory of the Higgs boson in the Standard Model is [44]. The best current calculations of the Higgs boson properties are compiled in [45].

Fig. 25: Feynman rules for couplings of the Higgs boson.

### 6.1 Decay modes of the Higgs boson

The basic elements of the SM description of the Higgs boson are extremely simple. A general configuration of the Higgs field can be written in the form of an  $SU(2)$  gauge transformation acting on a simple scalar field

$$\varphi(x) = \exp[-i\alpha^a(x)\sigma^a/2] \begin{pmatrix} 0 \\ (v + h(x))/\sqrt{2} \end{pmatrix}. \quad (181)$$

We can remove the prefactor by a choice of gauge. Then the Higgs field reduces to a vacuum expectation value  $v$  and the dynamical scalar field  $h(x)$ . The values of  $m_W$  and  $g$  give

$$v = 246 \text{ GeV}. \quad (182)$$

The vertices of  $h(x)$  are given by shifting  $v$  everywhere it appears in the SM

$$v \rightarrow v + h(x). \quad (183)$$

This gives rise to the Feynman rules shown in Fig. 25. Within the SM, there is no freedom to change these vertices.

The couplings in Fig. 25 imply that a heavy Higgs boson would decay dominantly into pairs of the other heavy particles of the SM,

$$h \rightarrow W^+W^-, h \rightarrow ZZ, h \rightarrow t\bar{t} \quad (184)$$

However, it has been found at the LHC that there is no heavy resonance that decays to these final states. On the other hand, a narrow resonance with the properties of the Higgs boson has been found at the LHC at a mass of 125 GeV. At this mass value, the otherwise dominant decay modes of the Higgs boson are kinematically forbidden. The actual decay modes of the Higgs are all suppressed in some way, by factors

$$\frac{m_f^2}{m_W^2}, \quad \frac{\alpha_w}{4\pi}, \quad \text{or} \quad \left(\frac{\alpha_s}{4\pi}\right)^2. \quad (185)$$

This means that the decay pattern of the Higgs boson will be more complex than might have been expected, but also that it should be very rich, with a large number of decay modes accessible to observation.

To describe these decays, I begin with the decays to fermions. The matrix element for Higgs decay to a light fermion is

$$i\mathcal{M}(h \rightarrow f_R\bar{f}_R) = -i\frac{m_f}{v}u_R^\dagger v_R = -i\frac{m_f}{v}(2E). \quad (186)$$

and similarly for decay to  $f_L\bar{f}_L$ . The total decay rate is

$$\Gamma(h \rightarrow f\bar{f}) = \frac{1}{2m_h} \frac{1}{8\pi} \frac{m_f^2 m_h^2}{v^2} \cdot 2, \quad (187)$$

or, using  $v^2 = 4m_W^2/g^2$ ,

$$\Gamma(h \rightarrow f\bar{f}) = \frac{\alpha_w}{8} m_h \frac{m_f^2}{m_W^2}. \quad (188)$$

For final-state leptons, we can immediately evaluate this,

$$\Gamma(h \rightarrow \tau^+\tau^-) = 260 \text{ keV} \quad \Gamma(h \rightarrow \mu^+\mu^-) = 9 \text{ keV} \quad (189)$$

for  $m_h = 125 \text{ GeV}$ .

For decays to quarks, a few more details must be added. The quark mass must be defined by some renormalization convention. An appropriate choice that absorbs large logarithms is to set the quark mass in (188) equal to the  $\overline{MS}$  quark mass evaluated at  $Q = m_h$ . This is related to the quark mass as usually quoted by

$$m_f(m_h) = m_f(m_f) \left[ \frac{\alpha_s(m_h)}{\alpha_s(m_f)} \right]^{4/b_0} (1 + \mathcal{O}(\alpha_s)), \quad (190)$$

where  $b_0$  is the first coefficient of the QCD  $\beta$  function, equal to  $23/3$  for 5 light quark flavors. This means that the values of the quark masses appropriate to the calculation of Higgs boson branching ratios are

$$\begin{array}{ccccc} m_u & m_d & m_s & m_c & m_b \\ 1.5 & 3 & 60 & 700 & 2800 \end{array} \quad (191)$$

with all values in MeV. The formula (188) must also be multiplied by the color factor of 3 and a substantial QCD correction

$$3 \cdot \left( 1 + \frac{17}{3\pi} \alpha_s(m_h) + \dots \right) = 3 \cdot 1.24. \quad (192)$$

Then, for example,

$$\Gamma(h \rightarrow b\bar{b}) = \frac{\alpha_w m_h}{8} \left( \frac{2.8}{m_W} \right)^2 \cdot 3 \cdot 1.24 = 2.4 \text{ MeV}. \quad (193)$$

After we compute the other major Higgs boson decay rates, this will correspond to a branching fraction of 58%. Then the total width of the Higgs boson is predicted to be about 4.1 MeV, and the other fermion branching fractions should be

$$\begin{array}{cccc} \tau^+\tau^- & c\bar{c} & s\bar{s} & \mu^+\mu^- \\ 6.3\% & 3\% & 0.03\% & 0.02\% \end{array} \quad (194)$$

It is somewhat surprising that the branching ratio for  $\tau^+\tau^-$  is larger than that for  $c\bar{c}$ , despite the presence of the color factor of 3.

For a heavy Higgs boson that can decay to on-shell  $W$  and  $Z$  bosons, the decay amplitudes would be

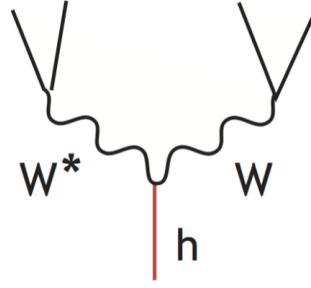
$$\begin{aligned} i\mathcal{M}(h \rightarrow W^+W^-) &= i \frac{2m_W^2}{v} \epsilon^*(+) \cdot \epsilon^*(-) \\ i\mathcal{M}(h \rightarrow ZZ) &= i \frac{2m_Z^2}{v} \epsilon^*(1) \cdot \epsilon^*(2). \end{aligned} \quad (195)$$

For a very heavy Higgs boson, there is a further enhancement for the longitudinal polarization states,

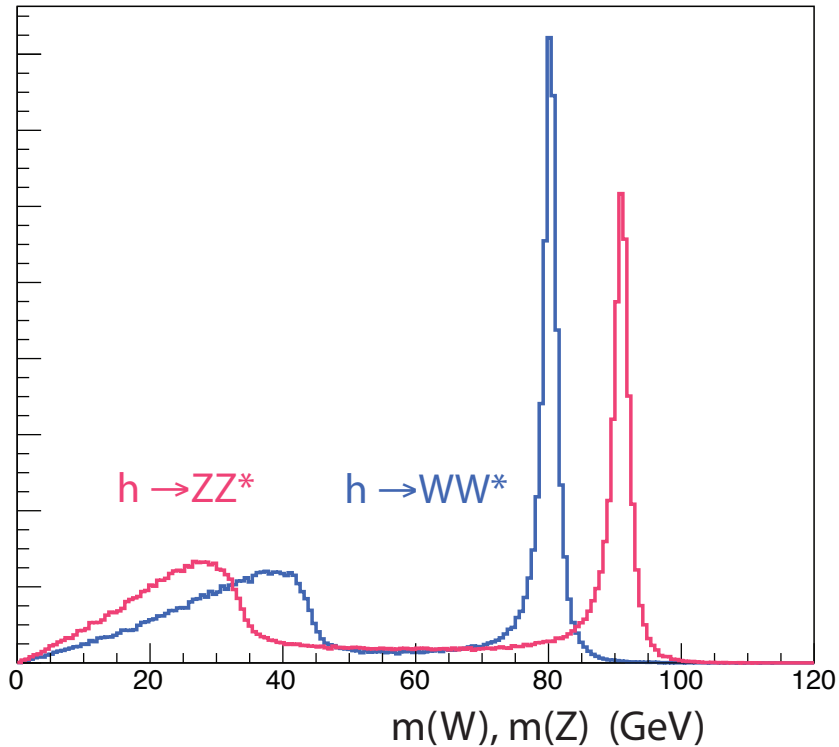
$$\epsilon_0^*(1) \cdot \epsilon_0^*(2) \sim \frac{k_1 \cdot k_2}{m_Z^2} \sim \frac{m_h^2}{2m_Z^2}. \quad (196)$$

This factor is just

$$\frac{\lambda}{(g^2 + g'^2)}. \quad (197)$$



**Fig. 26:** Feynman diagram for  $h \rightarrow WW$  or  $h \rightarrow ZZ$  decay with the vector bosons off-shell.



**Fig. 27:** Mass distributions of the off-shell  $W$  and  $Z$  bosons in the decay of a 125 GeV Higgs boson.

so the longitudinal  $Z$  and  $W$  couple to the Higgs boson as Higgs boson rather than as gauge bosons. This is in accord with the GBET.

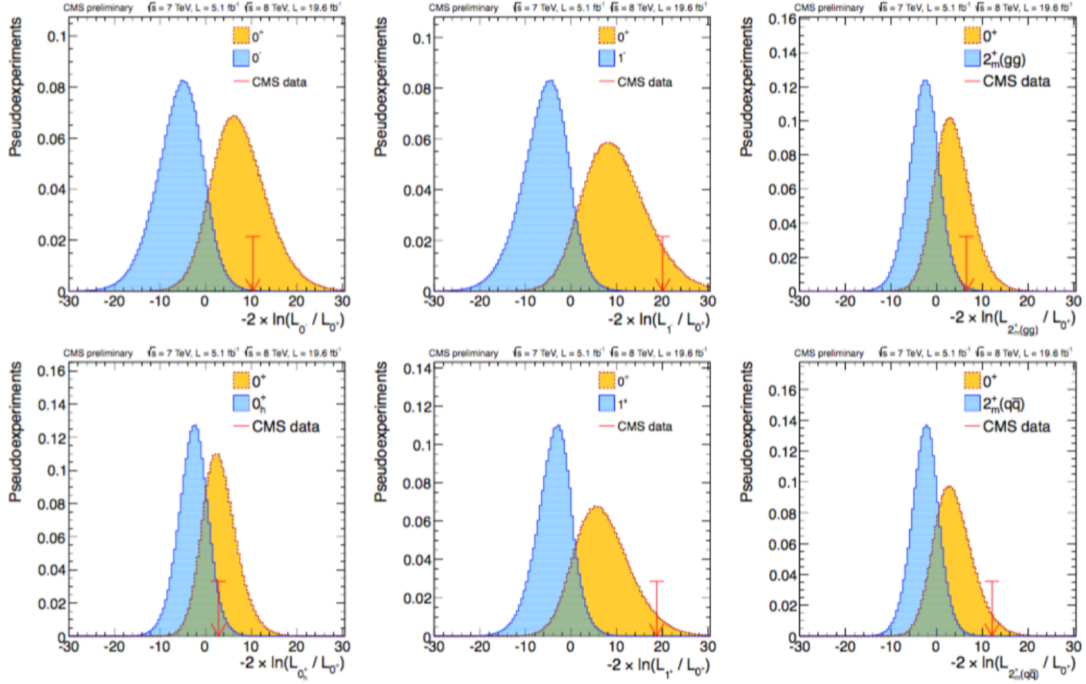
For the actual situation of a 125 GeV Higgs boson, one or both of the  $W$  and  $Z$  bosons must be off-shell. Then the decay is best described as a Higgs decay to 4 fermions, as shown in Fig. 26. The rate is suppressed by a factor of  $\alpha_w$  and by the off-shell  $W$  or  $Z$  propagator. The result is that the rate is competitive with  $b\bar{b}$  for the  $WW$  mode and a factor 10 smaller for  $ZZ$ . The SM branching fractions for these off-shell vector boson modes are

$$BR(h \rightarrow WW^*) = 22\% \quad BR(h \rightarrow ZZ^*) = 2.7\% . \quad (198)$$

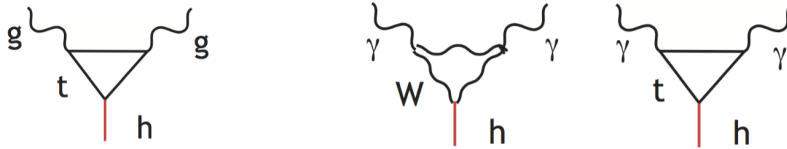
The  $W$  and  $Z$  mass distributions in these decays are shown in Fig. 27.

The Higgs boson decay to  $ZZ^*$  is exceptionally interesting because it is completely reconstructable in LHC events in which both  $Z$ s decay to charged leptons. The angular distribution of the





**Fig. 28:** Likelihood distributions for tests of the spin and parity of the Higgs boson, from [46].



**Fig. 29:** Loop diagrams contributing the the  $h \rightarrow gg$  and  $h \rightarrow \gamma\gamma$  decays.

leptons permits an analysis of the spin and parity of the Higgs resonance. In the SM, where the Higgs boson must have  $J^P = 0^+$ , the two  $Z$  bosons are predicted to be longitudinally polarized with the two decay planes parallel. The polarization of the  $Z$  can be measured from the decay angular distribution, as we have discussed for  $W$  bosons in (136). This prediction contrasts with that for other possible spin 0 assignments, in which the Higgs boson couples to  $ZZ^*$  through the interactions

$$0^- : h\epsilon^{\mu\nu\lambda\sigma} Z_{\mu\nu} Z_{\lambda\sigma} \quad 0_h^+ : hZ_{\mu\nu} Z^{\mu\nu} . \tag{199}$$

For the interactions in (199), the  $Z$  bosons are preferentially transversely polarized; also, with the  $0^-$  type interaction, the two decay planes tend to be orthogonal. The SM prediction was tested even with the relatively small sample of about 15  $Z \rightarrow 4$  lepton events collected by each LHC experiment in run 1 of the LHC. Figure 28 shows the expected distributions of the likelihood for tests of the predicted SM coupling structure against the coupling structures in (199) and 4 other structures for which the resonance has spin 1 or spin 2. The actual value of the likelihood found by CMS experiment is shown by the arrow. In all cases, the results strongly favor the SM hypothesis [46].

Finally, there are loop processes that allow the Higgs boson to decay to a pair of massless vector bosons,  $gg$  or  $\gamma\gamma$ , or to  $Z\gamma$ . The most straightforward of these to analyze is the  $hgg$  vertex. This is generated by loop diagrams that involve quarks, such as the diagram shown on the left in Fig. 29.

If we compute these loop diagrams, we obtain a local operator that gives an effective description of the Higgs boson coupling to  $gg$ . The lowest-dimension operator that is invariant under the  $SU(3)$  gauge symmetry is

$$\Delta\mathcal{L} = \frac{1}{4}AhF_{\mu\nu}^a F^{\mu\nu a}, \quad (200)$$

where  $F_{\mu\nu}^a$  is the QCD field strength. The coefficient  $A$  has the dimensions  $(\text{GeV})^{-1}$ . This operator yields the  $hgg$  vertex

$$-iA\delta^{ab}(k_1 \cdot k_2 g^{\mu\nu} - k_2^\mu k_1^\nu). \quad (201)$$

I will compute the coefficient  $A$  in a moment, but, first I will estimate the order of magnitude of the contribution from a quark of mass  $m_q$ . There is a surprise here. This contribution is proportional to the Higgs Yukawa coupling, so it must be of the form

$$\alpha_s \frac{m_f}{v} \frac{1}{M}, \quad (202)$$

where  $M$  is the momentum that flows in the loop. For  $2m_q \ll m_h$ ,  $M$  will be of order  $m_h$  and so the contribution (202) will be suppressed by a factor  $m_f/m_h$ . On the other hand, if  $2m_q \gg m_h$ ,  $M$  will be of order  $m_q$ . In this case, the factors of  $m_q$  cancel and the diagram is at full strength no matter how large  $m_q$  is. This is bizarre but correct: The  $hgg$  vertex gets only small contributions from quarks to which the Higgs boson can decay and obtains full-strength contributions from quarks to which the Higgs boson *cannot* decay because they are too heavy.

In the SM, the only quark that contributes to the  $hgg$  vertex at full strength is the top quark. If there were a fourth generation of quarks that obtained their masses from the SM Higgs boson, each quark would produce an equal contribution to the  $hgg$  coupling, so that the total decay rate  $\Gamma(h \rightarrow gg)$  would be  $3^2 = 9$  times the SM prediction [47]. Such a large shift is already excluded by the LHC Higgs measurements. This is a much stronger constraint on a fourth generation than the one that we found from precision electroweak measurements at the end of Section 3.

We can compute the contribution to the  $hgg$  vertex from a heavy quark  $t$  from the starting point of the QCD vacuum polarization. The 1-loop quark vacuum polarization diagram has the value

$$\begin{aligned} & i(k^2 g^{\mu\nu} - k^\mu k^\nu) \text{tr}[t^a t^b] \frac{\alpha_s}{3\pi} \log \frac{\Lambda^2}{m_t^2} \\ & i(k^2 g^{\mu\nu} - k^\mu k^\nu) \text{tr}[t^a t^b] \frac{\alpha_s}{3\pi} \log \frac{\Lambda^2}{m_t^2}. \end{aligned} \quad (203)$$

We can produce the top quark loop diagram in Fig. 29, adding a zero-momentum Higgs boson, by shifting  $v \rightarrow v + h$  as in (183). The expression (203) depends on  $v$  through  $m_t = y_t v / \sqrt{2}$ . This yields a contribution to the  $hgg$  vertex that is finite and equal to

$$i(k^2 g^{\mu\nu} - k^\mu k^\nu) \delta^{ab} \frac{\alpha_s}{3\pi} \frac{1}{v}. \quad (204)$$

Comparing to (201), we find

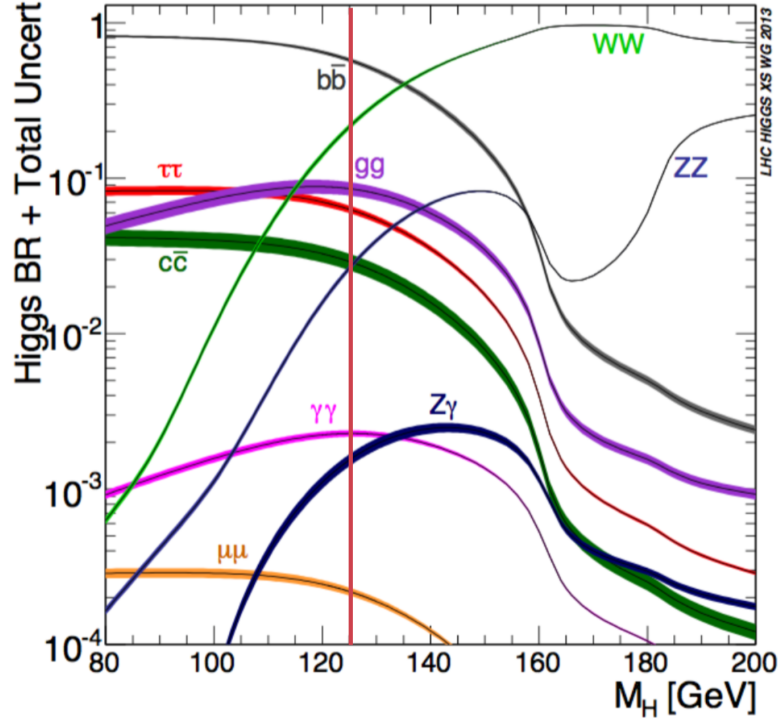
$$A = \frac{\alpha}{3\pi v} = \frac{g\alpha_s}{6\pi m_W}. \quad (205)$$

From this expression, we can compute the partial width  $\Gamma(h \rightarrow gg)$  in the limit  $m_h^2 \ll 4m_t^2$ ,

$$\Gamma(h \rightarrow gg) = \frac{\alpha_w \alpha_s^2}{72\pi^2} \frac{m_h^3}{m_W^2}. \quad (206)$$

The full expression can be shown to be

$$\Gamma(h \rightarrow gg) = \frac{\alpha_w \alpha_s^2}{72\pi^2} \frac{m_h^3}{m_W^2} \cdot \left| \frac{3}{2} \tau (1 - (\tau - 1) (\sin^{-1} \frac{1}{\sqrt{\tau}})^2) \right|^2, \quad (207)$$



**Fig. 30:** Standard Model predictions for the branching ratios of the Higgs boson as a function of the its mass, from [45].

where  $\tau = 4m_t^2/m_h^2$ .

Another way to interpret this argument is that the shift of  $v$  in (183) is a change of scale for the SM. Then the 1-loop Higgs couplings to a gauge boson should be proportional to the 1-loop contribution to the renormalization group  $\beta$  function. The calculation just performed satisfies this, since (203) give the contribution of a quark to the QCD  $\beta$  function. Changing what needs to be changed, we can obtain the coupling of a Higgs boson to  $\gamma\gamma$ . The contribution from the top quark and the  $W$  boson to the QED vacuum polarization is

$$i(k^2 g^{\mu\nu} - k^\mu k^\nu) \frac{\alpha}{4\pi} \left[ -\frac{22}{3} + \frac{1}{3} + \frac{4}{3} \cdot 3 \cdot \left(\frac{2}{3}\right)^2 \right] \log \frac{\Lambda^2}{m_{t,W}^2}. \quad (208)$$

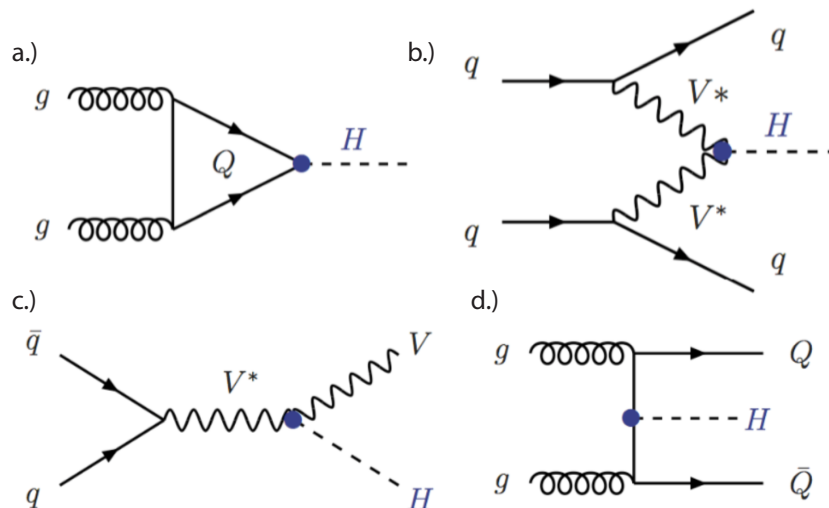
The first term here is contribution from the  $W$ , it is just the standard vector boson contribution to the  $\beta$  function for an  $SU(2)$  gauge theory. The second term comes from the Higgs boson that the  $W$  boson must eat to become massive. The third term comes from the top quark; the last two factors are the top quark color factor and electric charge. In all, we find, for  $m_h \ll 2m_W, 2m_t$ ,

$$\Gamma(h \rightarrow \gamma\gamma) = \frac{\alpha_w \alpha^2}{144\pi^2} \frac{m_h^3}{m_W^2} \left| \frac{21}{4} - \frac{4}{3} \right|^2. \quad (209)$$

Careful evaluation, including all finite mass effects and the QCD corrections to the gluon width, gives

$$BR(h \rightarrow gg) = 8.6\% \quad BR(h \rightarrow \gamma\gamma) = 0.23\%. \quad (210)$$

We are now ready to put all of the pieces together to compile the SM predictions for the various Higgs boson branching ratios. Figure 30 shows the predictions as a function of the Higgs boson mass. It is a useful exercise to understand the shape of the curves based on the physics discussed in this section.



**Fig. 31:** Reactions producing the Higgs boson in  $pp$  collisions

The position of the observed Higgs resonance is shown by the vertical line. At this mass value, there are 10 distinct final states with branching fractions larger than  $10^{-4}$ , including the  $s\bar{s}$  channel not shown on this plot.

## 6.2 Study of the Higgs boson at the LHC

With this understanding of the Higgs boson couplings, I will review very briefly the results for Higgs boson couplings obtained by the ATLAS and CMS experiments. The most important processes for the production of a Higgs boson at the LHC are those shown in Fig. 31: gluon-gluon fusion, vector boson fusion, radiation of the Higgs boson from a  $W$  or  $Z$  (“Higgsstrahlung”), and associate production of a Higgs boson with a pair of top quarks. The cross sections predicted for these processes for a 125 GeV Higgs boson are shown in Fig. 32.

The four reactions have different advantages for the study of Higgs decays. Gluon-gluon fusion has the highest cross section, so it gives access to rare Higgs decays. In vector boson fusion, Higgs events are tagged by the presence of forward quark jets, reducing the background from non-Higgs SM processes. This reaction also has the smallest theoretical error on the predicted cross section. Higgsstrahlung also gives tagged Higgs decays. It also can lead to highly boosted Higgs bosons, which is an advantage for isolating the  $h \rightarrow b\bar{b}$  decay. Finally, the top associated production process gives access to the  $ht\bar{t}$  coupling.

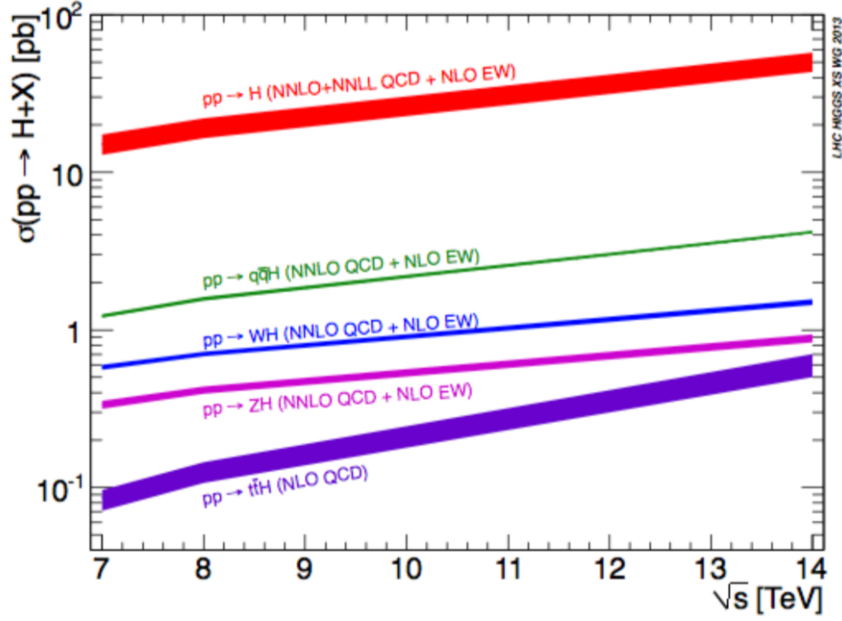
In all cases, what is measured is a combination of the cross section for Higgs production and the branching fraction for Higgs decay into the observed final state. This observable is related to the Higgs couplings through

$$\sigma(pp \rightarrow A\bar{A} \rightarrow h)BR(h \rightarrow B\bar{B}) \sim \frac{\Gamma(h \rightarrow A\bar{A})\Gamma(h \rightarrow B\bar{B})}{\Gamma_h}. \quad (211)$$

In this relation,  $A\bar{A}$  is the parton combination used to produce the Higgs boson— $gg$ ,  $WW$  or  $ZZ$ , and  $tt$ , respectively, for the processes in Fig. 32. The measured rates are quoted in terms of the *signal strength*  $\mu$

$$\mu = \sigma(pp \rightarrow h \rightarrow B\bar{B})/(\text{SM prediction}). \quad (212)$$

Note that, if a departure from the SM value  $\mu = 1$  is seen, this might be due to a nonstandard value of the  $hA\bar{A}$  coupling, the  $hB\bar{B}$  coupling, or the Higgs total width. Multiple measurements would be needed to resolve this ambiguity.



**Fig. 32:** Cross sections for Higgs production in  $pp$  collisions for a 125 GeV Higgs boson, from [48].

The original strategy for observing the Higgs boson at the LHC used the characteristic decay modes in which this particle could be reconstructed as a resonance.

$$h \rightarrow \gamma\gamma, \quad h \rightarrow ZZ^* \rightarrow 4 \text{ leptons} \quad (213)$$

These modes correspond to branching fractions of

$$0.23\% \quad \text{and} \quad 0.012\% \quad (214)$$

With production cross sections of about 20 pb at 7 TeV, these processes have rates corresponding to fractions

$$4 \times 10^{-13} \quad \text{and} \quad 2 \times 10^{-14}, \quad (215)$$

respectively, of the  $pp$  total cross section. The observation of these very tiny components of the total reaction rate at the LHC is quite an achievement! Signals of the Higgs resonance in LHC run 1 data are shown in Fig. 33.

Once we are convinced that the Higgs resonance is actually present at a mass of 125 GeV, we can look for the signatures of this resonance in other decay modes. Higgs decays to these channels give larger total rates than the decays to the discovery modes. But, these channels produce events that are not obviously distinguishable from other SM reactions.

An example is

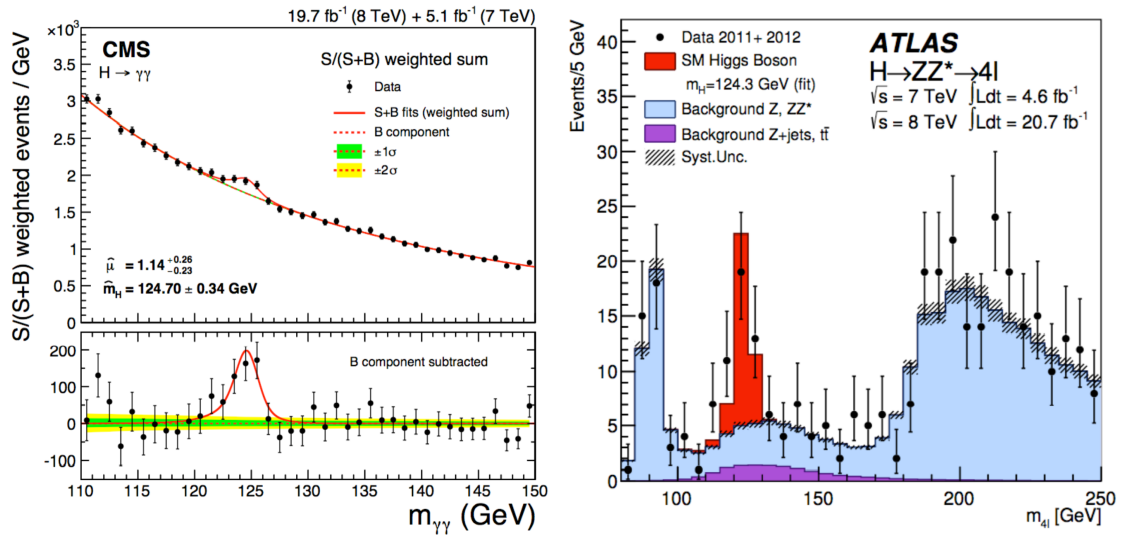
$$pp \rightarrow h \rightarrow W^+W^- \rightarrow \ell^+\ell^-\nu\bar{\nu}. \quad (216)$$

The observable properties of these events overlap strongly with events from

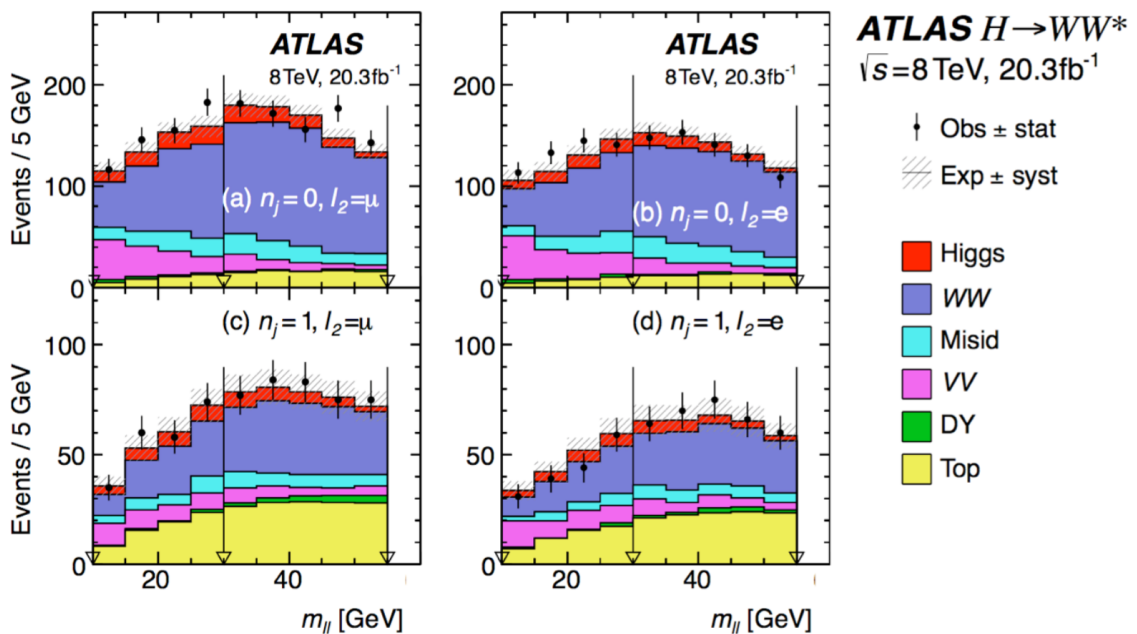
$$pp \rightarrow W^+W^- \rightarrow \ell^+\ell^-\nu\bar{\nu}. \quad (217)$$

The signal to background ratio can be enhanced by selecting the region where  $m(\ell^+\ell^-)$  and the angle between the two leptons are both relatively small. It is also necessary to apply a jet veto (that is, to select events with at most 1 high- $p_T$  jet) in order to avoid background from

$$pp \rightarrow t\bar{t} \rightarrow b\bar{b}\ell^+\ell^-\nu\bar{\nu}. \quad (218)$$



**Fig. 33:** Signals of the Higgs boson resonance at the LHC in run 1: left: Higgs resonance in the  $m(\gamma\gamma)$  distribution, from [49]; right: Higgs resonance in the  $m(4\ell)$  distribution [50].



**Fig. 34:** Evidence for the Higgs boson in its decay to  $WW^*$ , from [51].

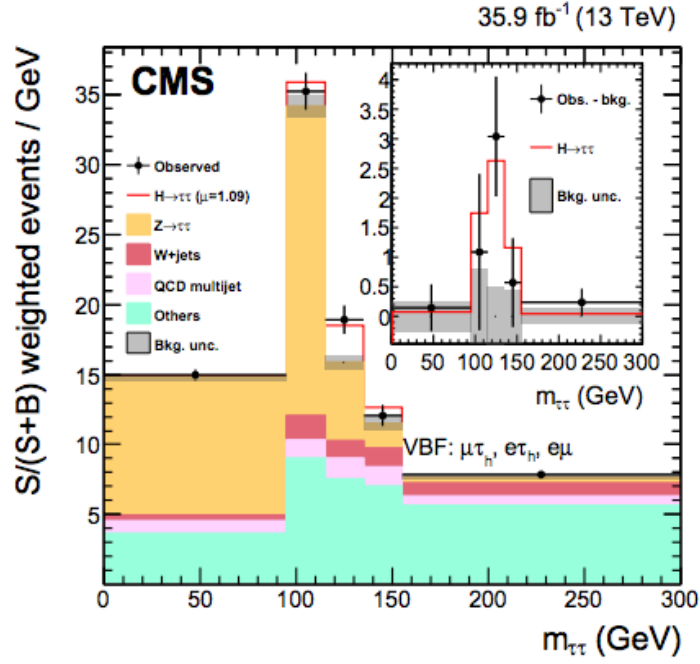


Fig. 35: Evidence for the Higgs boson decay to  $\tau^+\tau^-$ , from [52].

Figure 34 shows the distributions in  $m(\ell^+\ell^-)$  for four event selections from the ATLAS analysis at 8 TeV. The histograms show the SM simulation of this event sample, with the various colored bands indicating the contributions of expected processes. The largest event rates come from  $pp \rightarrow WW$  and, for the 1-jet events shown in the bottom row,  $pp \rightarrow t\bar{t}$ . The data points indicate a 10% excess rate over the SM expectation from processes that do not involve a Higgs boson, which is well accounted for by the expected rate for Higgs production.

Similar analyses support the presence of Higgs boson production and decay to  $\tau^+\tau^-$ . The most important backgrounds are

$$pp \rightarrow Z \rightarrow \tau^+\tau^-, \quad pp \rightarrow W^+W^-, \quad (219)$$

and QCD reactions where two jets in the final state fake the  $\tau$  signatures. The strongest evidence for the reaction comes from vector boson fusion, since the tagging by forward jets helps to minimize the QCD background. Figure 35 shows the very recent CMS run 2 analysis with data from 13 TeV. These events are dominated by the large background from  $Z \rightarrow \tau^+\tau^-$ . However, this background can be understood using the observed distribution of  $Z \rightarrow \mu^+\mu^-$  events. The backgrounds from  $WW$  and QCD are more challenging to estimate. Fig. 36 shows a candidate vector boson fusion  $h \rightarrow \tau^+\tau^-$  event from ATLAS. I use the word “candidate” advisedly; probably this event is a  $Z \rightarrow \tau^+\tau^-$  event produced by vector boson fusion.

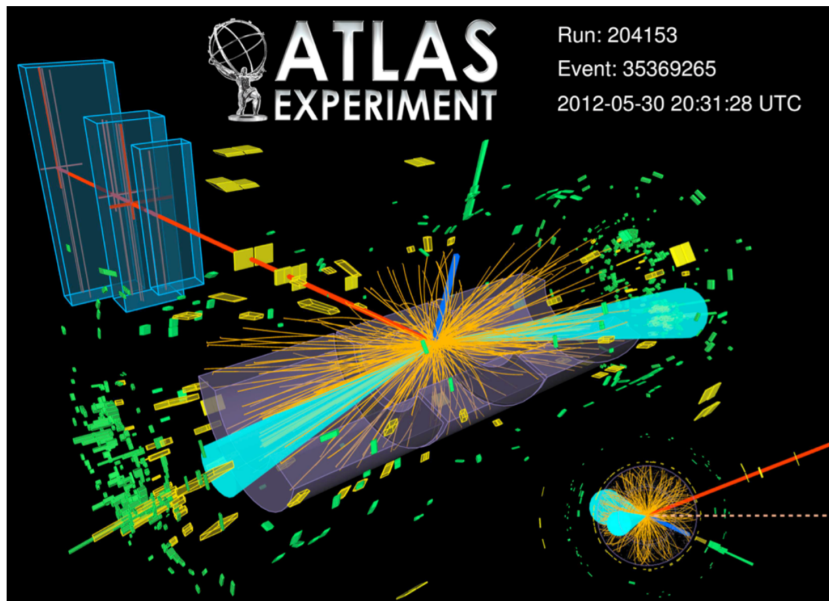
The most challenging of the major modes of Higgs decay is the one with the highest branching ratio,  $h \rightarrow b\bar{b}$ . It is probably hopeless to observe this mode in gluon fusion at low Higgs  $p_T$ , since  $gg \rightarrow b\bar{b}$  with  $m(b\bar{b}) \sim 125$  GeV has a cross section about a million times larger than that of the Higgs process. Current analyses use the Higgsstrahlung process with a tagging  $W$  or  $Z$

$$pp \rightarrow Vh, \quad h \rightarrow b\bar{b} \quad (220)$$

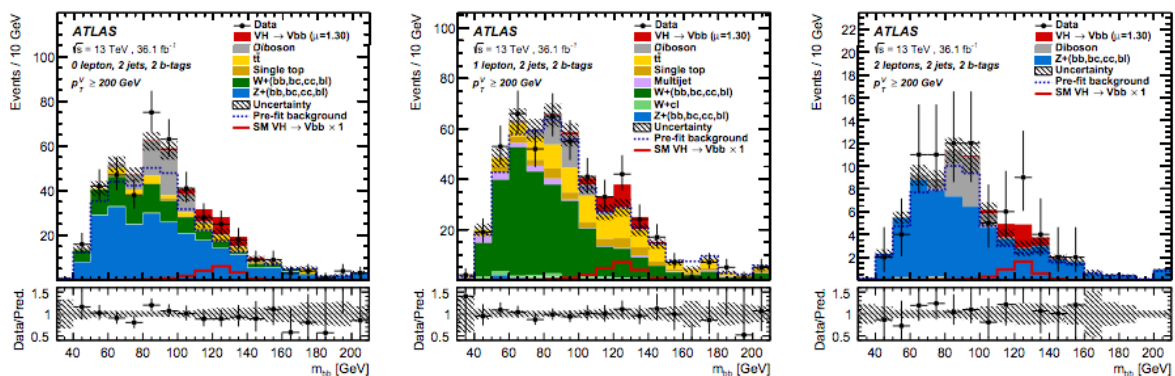
where  $V$  is  $W$  or  $Z$ . However, there are other SM processes with similar signatures that do not involve a Higgs boson,

$$pp \rightarrow VZ, \quad Z \rightarrow b\bar{b}$$





**Fig. 36:** A candidate event for vector boson fusion production of a Higgs boson decaying to  $\tau^+\tau^-$ , from [53].



**Fig. 37:** Evidence for the Higgs boson decay to  $b\bar{b}$ , from [54]. The three distributions show 0, 1, and 2-lepton events. The red (dark) boxes near the mass value of 125 GeV show the expectation from  $pp \rightarrow Vh, h \rightarrow b\bar{b}$ .

$$pp \rightarrow Vg, \quad g \rightarrow b\bar{b}. \quad (221)$$

The second reaction involves an off-shell gluon with a mass near 125 GeV that converts to  $b\bar{b}$ . Convincing evidence for this decay has been obtained only very recently, in the 13 TeV data [54]. The current evidence from the ATLAS run 2 data is shown in Fig. 37. It is expected that discrimination of the three processes (220), (221) can be improved in an event sample in which the state recoiling against the vector boson is highly boosted, using techniques that measure the dijet mass and color flow. A recent analysis by CMS shows a small signal for  $h \rightarrow b\bar{b}$  in a sample of high  $p_T$  jets recoiling against a gluon jet [55].

Figure 38 shows a summary of the measurements of the Higgs boson signal strengths made by ATLAS and CMS in run 1 of the LHC [56]. A signal strength of 0 indicates no presence of the Higgs boson. This hypothesis is excluded by run 1 data for all of the modes considered except  $h \rightarrow b\bar{b}$ . I have discussed above the more significant evidence for  $h \rightarrow \tau^+\tau^-$  and  $h \rightarrow b\bar{b}$  found already in run 2. A signal strength of 1 is the prediction of the SM. The measured rates agree with this prediction within about 30% accuracy. So the quantitative study of the Higgs boson has begun and will be improved as the

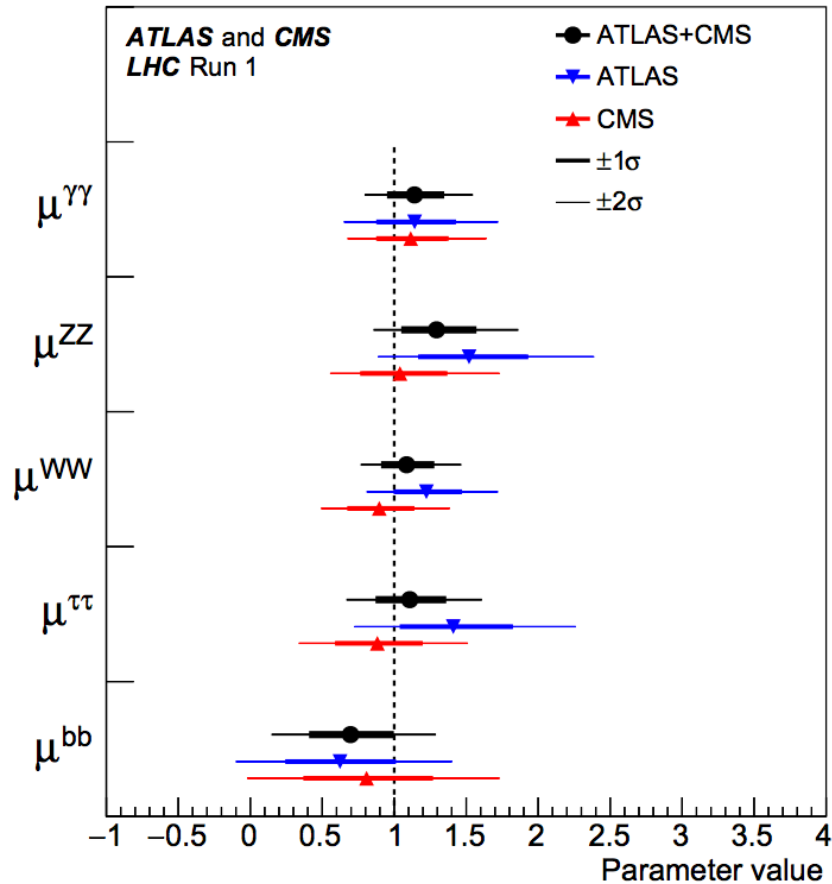


Fig. 38: Summary of Higgs  $\mu$  measurements, from [56].

LHC accumulates data.

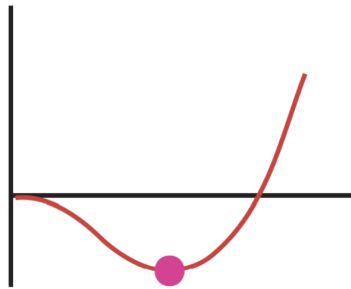
## 7 Precision measurements of the Higgs boson properties

In the last segment of these lectures, I take a step outside the Standard Model. In this section, I will discuss the expectations for the couplings of the Higgs boson in theories beyond the Standard Model. This is an interesting story that motivates a dedicated experimental campaign to measure the couplings of the Higgs boson with high precision. First, though, I will explain why I believe there must be new interactions of physics waiting to be discovered.

### 7.1 The mystery of electroweak symmetry breaking

I have shown in the previous lectures that the SM of weak interactions is an extremely successful theory in its own domain. It is not a complete theory of nature, but we can supplement it by adding gravity, quantum chromodynamics (QCD) as the theory of the strong interactions, and some model of dark matter and dark energy. It is also not difficult to add neutrino masses to the model, either by introducing three generations of right-handed neutrinos or by adding lepton-number-violating Majorana mass terms. Each of these additions accounts for some set of observed phenomena that is outside the range of topics considered in these lectures.

But this is not enough. A key part of the explanation for the structure of the weak interactions and the generation of masses for quarks, leptons, and gauge bosons is the spontaneous symmetry breaking of



**Fig. 39:** The Higgs potential  $V(|\varphi|)$ .

$SU(2) \times U(1)$  and the generation of the Yukawa couplings that link the symmetry-breaking Higgs field to the quarks and leptons. The structure that I have described leads immediately to questions about all of these ingredients:

- Why just quarks and leptons? What is the origin of the quantum number assignments  $(I, Y)$  for the matter particles seen in nature?
- What explains the spectrum of quark and lepton masses? The SM gives the relation

$$m_f = \frac{y_f v}{\sqrt{2}}, \quad (222)$$

where  $v$  is the Higgs field vacuum expectation value. But the  $y_f$  are renormalized parameters that cannot be predicted with the Standard Model. The presence of nonzero CKM angles—and, with neutrinos, PMNS angles—adds further difficulty to this problem.

- What is the origin of the Higgs field? Is there only one such field, or are there multiplets of scalar fields with different quantum numbers? The SM makes the minimal choice of one Higgs multiplet. Is this necessary?
- Why is  $SU(2) \times U(1)$  spontaneously broken? The shape of the Higgs potential energy function is an input for which the SM gives no explanation.

This last question merits more discussion. Here is the explanation for electroweak symmetry breaking given in the SM: The model instructs us to write the most general renormalizable potential for the Higgs field  $\varphi$ ,

$$V(\varphi) = \mu^2 |\varphi|^2 + \lambda |\varphi|^4. \quad (223)$$

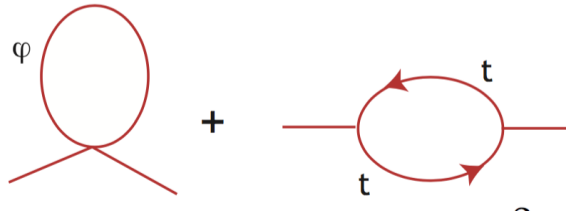
We assume that  $\mu^2 < 0$ . Then the potential has the correct shape, shown in Fig. 39, to drive spontaneous symmetry breaking.

Why must  $\mu^2$  be negative? That question cannot be addressed within the model. It is just a choice, perhaps a random one.

We get into deeper trouble if we try to take this explanation to a higher level of precision by computing the radiative corrections to the parameter  $\mu^2$ . The leading one-loop corrections, from loops containing the Higgs and top quark fields, are shown in Fig 40. They give

$$\mu^2 = \mu_{\text{bare}}^2 + \frac{\lambda}{8\pi^2} \Lambda^2 - \frac{3y_t^2}{8\pi^2} \Lambda^2 + \dots \quad (224)$$

The diagrams are ultraviolet divergent. I have regularized them by cutting off their momentum integrals at a mass scale  $\Lambda$ , arbitrarily chosen to be the same for Higgs and top. The final value of  $\mu^2$  needed to produce the observed Higgs boson mass is  $\mu^2 \approx -(100 \text{ GeV})^2$ . So if  $\Lambda$  is much larger than 1 TeV, this



**Fig. 40:** One-loop corrections to the  $\mu^2$  parameter from the Higgs field coupling to the top quark and from the Higgs field self-coupling.

formula requires large cancellations among the ingredients with no obvious explanation. If we assert that the SM is correct up to the Planck scale, the first 33 significant figures must cancel. It is also apparent that the right-hand side contains both positive and negative contributions, so it is not obvious without invoking a much deeper explanation why the final answer after the cancellation should turn out to be negative.

The simplest resolution of this set of problems would be that there are new particles, not yet known to us, that generate additional diagrams contributing to the calculation of  $\mu^2$ . If these particles have masses of TeV size, they might cancel the divergences seen in (224) and—in the best case—leave over a calculable answer for  $\mu^2$ . However, we have not yet been able to discover these particles in high-energy experiments.

The general problem of the uncalculability of the parameter  $\mu^2$  is not new to high-energy physics. It is encountered in all systems in which a symmetry is spontaneously broken. Condensed matter physics gives many examples.

The most direct analogy to the Higgs theory comes in the phenomenon of superconductivity seen in most metals at cryogenic temperatures. The original papers on the Higgs mechanism by Englert and Brout, Higgs, and Guralnik, Hagen, and Kibble [9–11] all used the analogy to superconductivity to motivate their arguments. However, they used only a piece of the complete theory. Superconductivity was discovered in 1911 by Kamerlingh Onnes and was quickly seen to be associated with a sharp phase transition [57]. However, the explanation for this phase transition was not understood for another 45 years.

In 1950, Landau and Ginzburg proposed a phenomenological theory of superconductivity based on a scalar field with the potential (223) [58]. They assumed that the parameter  $\mu^2$  would be a function of temperature, taking negative values below the phase transition temperature  $T_C$ . Coupling this theory to electromagnetism, they found that the photon acquires a mass by the Higgs mechanism and that the scalar fields in the vacuum can transmit electric current frictionlessly. This theory turned out to be extremely successful in explaining many aspects of superconductivity, including the Meissner effect in which superconductors repel magnetic flux, the existence of Type I and Type II superconductors, and the systematics of the destruction of superconductivity by high currents or high magnetic fields.

However, this theory could not address the most important problem of why superconductivity occurred in the first place. The answer to that question waited until 1957, when Bardeen, Cooper, and Schrieffer discovered the mechanism that causes electrons in a metal to pair up into bound states and form a boson condensate with the properties of the Landau-Ginzburg scalar field [59].

In our understanding of the phase transition to symmetry breaking of  $SU(2) \times U(1)$ , we are now at the Landau-Ginzburg stage.

In the case of superconductivity, physicists knew that there must be a deeper explanation that had to be given in terms of the interactions of electrons and atoms. For the symmetry-breaking of the weak interactions, any analogous explanation must involve new elementary particles outside the SM. We do

not know what these particles are. We only know that we have not discovered them yet.

## 7.2 Expectations for the Higgs boson in theories beyond the Standard Model

Even if we cannot discover new heavy particles responsible for the Higgs potential energy, we can hope to find clues to the nature of these new particles and interactions by looking more deeply into the properties of the Higgs boson itself. In the previous lecture, I emphasized that the SM makes precise predictions for the couplings of the Higgs boson to all particles of the SM in terms of the measured masses of those particles. Any deviation from these predictions must indicate the presence of new interactions beyond the SM. In this and the next two sections, I will trace out the expectations for corrections to the Higgs properties in different classes of models of new physics.

To begin, I will present two sets of expectations for the properties of new physics models. The first is guidance from the concept that these models should solve the problem of the calculability of the Higgs potential. The second comes from a constraint that is well-satisfied in the precision electroweak measurements.

I have already explained that the parameter  $\mu^2$  in the Higgs potential cannot be computed within the SM. To construct a model in which  $\mu^2$  can be computed, that model must satisfy some special properties. In particular, some structure in the theory must require the cancellation of quadratically divergent Feynman diagrams which would otherwise add large, arbitrary terms to the final result for  $\mu^2$ .

There are two strategies to achieve this. The first is to include in the model a symmetry that forbids the appearance of the

$$\mu^2|\varphi|^2 \quad (225)$$

term in the Lagrangian. It is not so obvious how to construct such a symmetry, since the operator (225) seems to be completely neutral. It would be forbidden in a scale-invariant theory, but in quantum field theory scale invariance is usually explicitly broken by the running of coupling constants. Two schemes that do forbid such a term are supersymmetry, the spacetime symmetry that links fermions and bosons, and the identification of  $\varphi$  with a Goldstone boson of some spontaneous symmetry breaking at a very high mass scale. The computation of the Higgs potential in models of supersymmetry is reviewed in [60, 61]. The computation of the Higgs potential in models in which the Higgs boson is a Goldstone boson is reviewed in [62, 63]. There are also other proposed generalizations of the SM Higgs sector in which the Higgs potential is not calculable.

One of the properties of mass generation in the SM is the relation  $m_W = m_Z c_w$ , as we saw in (16). This property can be derived from a symmetry of the Higgs potential assumed in the SM. Since the relation works so well, it is suggested that generalizations of the SM Higgs sector should also have this property.

The origin of the relation (16) can be seen as follows: Look at the form of the vector boson mass matrix acting on the original  $SU(2) \times U(1)$  fields,

$$m^2 = \begin{pmatrix} g^2 & & & \\ & g^2 & & \\ & & g^2 & -gg' \\ & & -gg' & g'^2 \end{pmatrix} \quad \text{on} \quad \begin{pmatrix} A^1 \\ A^2 \\ A^3 \\ B \end{pmatrix}. \quad (226)$$

The form of the matrix is dictated by the requirement that the matrix have a zero eigenvalue, associated with the massless photon, and that the part of the matrix acting on the  $SU(2)$  fields  $(A^1, A^2, A^3)$  should be symmetric among these fields. The requirement for the latter statement is that the theory contains an  $SO(3)$  transformation that rotates the  $SU(2)$  gauge fields into one another and is unbroken even when the  $SU(2)$  gauge symmetry is spontaneously broken. This extra transformation is called *custodial symmetry* [64].

Custodial symmetry is an accidental property of the SM Higgs potential. If we write

$$\varphi = \frac{1}{\sqrt{2}} \begin{pmatrix} \varphi^1 + i\varphi^2 \\ \varphi^0 + i\varphi^3 \end{pmatrix} \quad (227)$$

the Higgs potential depends only on the combination

$$|\varphi|^2 = (\varphi^0)^2 + (\varphi^1)^2 + (\varphi^2)^2 + (\varphi^3)^2. \quad (228)$$

A vacuum expectation value for  $\varphi^0$  preserves the  $SO(3)$  symmetry that acts on  $(\varphi^1, \varphi^2, \varphi^3)$ . From this observation, we understand why the SM satisfies (16).

There are many generalizations of the SM Higgs theory that also satisfy this condition. For example, we could introduce two or more scalar field multiplets with  $(I, Y) = (\frac{1}{2}, \frac{1}{2})$ . In the most general case, a different Higgs boson can be used to give mass to the charged leptons,  $d$  quarks, and  $u$  quarks, by writing the Higgs Yukawa interactions as

$$\mathcal{L} = -y_e L^\dagger \cdot \varphi_1 e_R - y_d Q^\dagger \cdot \varphi_2 d_R - y_u Q_a^\dagger \epsilon_{ab} \varphi_{3b}^\dagger u_R + h.c.. \quad (229)$$

In this equation,  $L$  is the left-handed lepton doublet,  $Q$  is the doublet of left-handed quarks, and all three Higgs multiplets have  $I = \frac{1}{2}$ ,  $Y = \frac{1}{2}$ . The three Higgs fields should have a potential that aligns their vacuum expectation values so that the  $U(1)$  symmetry giving electromagnetism remains unbroken. This structure can be extended to three generations by replacing the three Yukawa couplings by three  $3 \times 3$  matrices. The resulting theory shares with the Standard Model the property that, after a change of variables, the Higgs couplings are all CP even and flavor diagonal.

It can be shown that the Yukawa coupling with a complex conjugated field  $\phi_3^\dagger$  is inconsistent with supersymmetry. Then, in models of supersymmetry, we must introduce at least two Higgs double fields, one with  $I = \frac{1}{2}, Y = +\frac{1}{2}$ , to give mass to the  $d$  quarks and leptons, and a different field with  $I = \frac{1}{2}, Y = -\frac{1}{2}$ , to give mass to the  $u$  quarks.

More complex Higgs field multiplets are also possible. Georgi and Machacek found a way to preserve custodial symmetry with Higgs bosons in higher representations, corresponding to spin  $I$  under the weak interaction  $SU(2)$  symmetry [65, 66]. For example, for  $I = 1$ , we could introduce a  $3 \times 3$  matrix of fields

$$X = \begin{pmatrix} \chi^{0*} & \xi^+ & \chi^{++} \\ -\chi^{+*} & \xi^0 & \chi^+ \\ \chi^{++*} & -\xi^{+*} & \chi^0 \end{pmatrix}, \quad (230)$$

in which the rows are  $SU(2)$  triplets and the columns have  $Y = -1, 0, 1$ , respectively. The potential for this field can be arranged to have  $SU(2) \times SU(2)$  symmetry and a minimum at

$$\langle X \rangle = V \cdot \mathbf{1}_3 \quad (231)$$

that preserves the diagonal  $SU(2)$  as a global symmetry. We need at least one  $I = \frac{1}{2}$  Higgs multiplet to give mass to the quarks and leptons, but we can supplement this with additional Higgs fields with any value of  $I$ .

The criterion of custodial symmetry also provides guidance in constructing models of composite Higgs bosons that satisfy current phenomenological constraints. To provide examples of such models, let me begin by describing the Technicolor model introduced in 1978 by Weinberg and Susskind [67, 68]. These authors introduced a copy of QCD with two massless techni-quark flavors ( $U, D$ ), and with a strong interaction mass scale corresponding to a techni- $\rho$  meson mass at 2 TeV. This model has  $SU(2) \times SU(2)$  chiral symmetry, analogous to that in the known strong interactions. Just as happens there, the theory should have a spontaneous breaking of this symmetry to a diagonal  $SU(2)$  symmetry, dynamically generating masses for the techni-quarks and creating three techni-pions as Goldstone

bosons. The diagonal  $SU(2)$  symmetry remains unbroken, and this plays the role of the custodial symmetry. If this model is coupled to the  $SU(2) \times U(1)$  gauge symmetry of the SM, the  $W$  and  $Z$  bosons eat the Goldstone techni-pions and acquire mass through the Higgs mechanism. The  $W$  and  $Z$  masses obey (16), with

$$m_W = \frac{gF_\pi}{2}, \quad (232)$$

where  $F_\pi$  is the analogue of the pion decay constant in the technicolor interactions. We obtain the observed  $W$  and  $Z$  masses for  $F_\pi = 246$  GeV, the Higgs field expectation value in the SM. In this model, the Higgs boson would be a spin zero, isoscalar bound state of the  $U$  and  $D$  quarks and their antiquarks.

The Weinberg-Susskind technicolor model is now excluded. The model predicts a Higgs boson mass at about 1 TeV, and also too large an  $S$  parameter to be consistent with precision electroweak measurements. However, it points the way to more sophisticated models that also build the Higgs boson as a composite state.

An example is given by the following scenario, which uses the strong interaction chiral symmetry breaking in a different way: Introduce new QCD-like strong interactions at a mass scale of 10 TeV, with 4 associated quarks in real, rather than complex, representations of the gauge group. This theory has a chiral symmetry  $SU(4)$ , which is spontaneously broken to  $SO(4)$  when the quarks dynamically acquire mass.  $SU(4)$  has 15 generators, and  $SO(4)$  has 6, so the symmetry-breaking creates  $15 - 6 = 9$  Goldstone bosons. We might take two of the four quarks to transform as a doublet under the weak interaction  $SU(2)$  and the other two to be weak interaction singlets that form a doublet under another  $SU(2)$ . Then the Goldstone boson multiplet will contain 4 bosons that transform as  $(\frac{1}{2}, \frac{1}{2})$  under this  $SU(2) \times SU(2)$ . We can identify this multiplet with the Higgs boson doublet. This scenario realizes the idea of the Higgs doublet as a set of Goldstone bosons that, by Goldstone's theorem, stay massless while the strong interaction chiral symmetry is broken. In a set of models called *Little Higgs*, it is possible to perturb the strong interaction theory to produce a nonzero, calculable Higgs potential [69, 70].

### 7.3 The Decoupling Theorem

Through the strategies described in the previous section, it is possible to build many models of the Higgs field that are more complex than the SM and yet compatible with all current experimental constraints. One's first instinct is that these models will lead to wildly different predictions for the properties of the Higgs boson that are easily distinguished experimentally. However, this is not correct. To distinguish models of the Higgs sector, it is necessary to make detailed measurements reaching a relatively high degree of precision. This is a consequence of the Decoupling Theorem, enunciated by Howard Haber in [71].

The Decoupling Theorem states: If the spectrum of the Higgs sector contains one Higgs boson of mass  $m_h$ , with all other Higgs particles having masses at least  $M$ , then the influence of these particles on the properties of the light Higgs boson is proportional to

$$m_h^2/M^2. \quad (233)$$

If the Higgs sector contains additional particles, but these particles have masses of 1 TeV, they shift the properties of the known Higgs boson by corrections to the Higgs couplings at the percent level.

The proof of this theorem is quite straightforward. It uses the viewpoint of effective Lagrangians described in Section 5.4. As I have explained above, once we have measured the mass of the Higgs boson, the parameters of the SM relevant to the Higgs field are fixed, and the SM makes precise predictions for the Higgs couplings. On the other hand, I have also explained that the SM Lagrangian is the most general renormalizable Lagrangian with the known quark and lepton fields and the gauge symmetry  $SU(3) \times SU(2) \times U(1)$ . So, in an effective Lagrangian description, any perturbation of the Higgs



couplings away from the SM predictions must be associated with operators of dimension 6. These operators have dimensionful coefficients. If they are generated by particles of mass  $M$ , their coefficients will be of order  $1/M^2$ .

This situation is challenging but not hopeless. It implies that the current level of agreement of the Higgs boson properties with the predictions of the SM—to 20-30%, as described in the previous section—is absolutely to be expected no matter how complex the Higgs sector might be. But, it offers the opportunity that, with measurements of higher precision, a picture of the Higgs boson entirely different from that of the SM might be revealed.

#### 7.4 Effects on the Higgs boson couplings from models of new physics

To amplify this discussion of the effects of new physics on the SM Higgs couplings, I will now review some specific examples of those effects.

To begin, consider models with two Higgs scalar doublets. I remind you that supersymmetric models necessarily contain these effects, since supersymmetry requires two different Higgs doublets  $\varphi_u$ ,  $\varphi_d$  to give mass to the  $u$  and  $d$  quarks.

In a model with two Higgs doublets, there are a total of 8 Higgs degrees of freedom. When the Higgs fields acquire vacuum expectation values, 3 of these bosons are eaten by  $W$  and  $Z$  when these particles obtain mass through the Higgs mechanism. The remaining physical Higgs particles include two CP-even neutral Higgs bosons  $h^0$  and  $H^0$ , a neutral pseudoscalar bosons  $A^0$ , and a pair of charged Higgs bosons  $H^\pm$ . Most of the parameter space for such particles to have masses below 200 GeV has been excluded by searches at the LHC [72, 73].

In general, these particles correspond to mixtures of the fields in the original two Higgs doublets. The mixing angle that defines the CP-even mass eigenstates is called  $\alpha$ . For the CP-odd states, one mixture gives the eaten Goldstone bosons and orthogonal combination gives the physical boson mass eigenstates. The mixing angle that defines these linear combinations is called  $\beta$ , with

$$\tan \beta = \langle \varphi_u \rangle / \langle \varphi_d \rangle . \quad (234)$$

The properties of the observed Higgs boson are then predicted to be modified as a result of these mixings. At the lowest order,

$$g(hdd) = -\frac{\sin \alpha}{\cos \beta} \frac{m_d}{v} \quad g(huu) = \frac{\cos \alpha}{\sin \beta} \frac{m_d}{v} . \quad (235)$$

The first of these modifications applies to the  $b$  quark-Higgs coupling, the second to the  $c$  and  $t$  couplings.

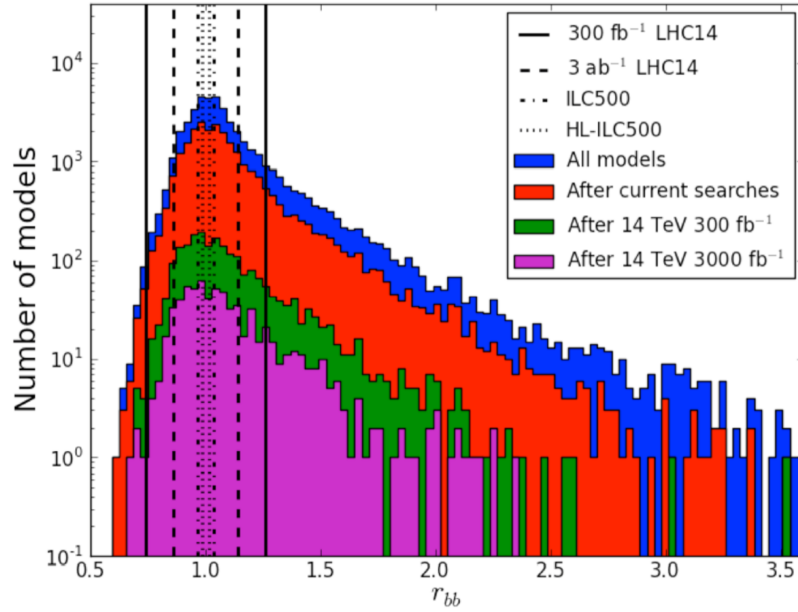
The Decoupling Theorem requires that the angles  $\alpha$ ,  $\beta$  cannot take arbitrary values but rather must be correlated. For example, in the minimal supersymmetric model,

$$-\frac{\sin \alpha}{\cos \beta} = 1 + \mathcal{O}\left(\frac{m_Z^2}{m_A^2}\right) , \quad (236)$$

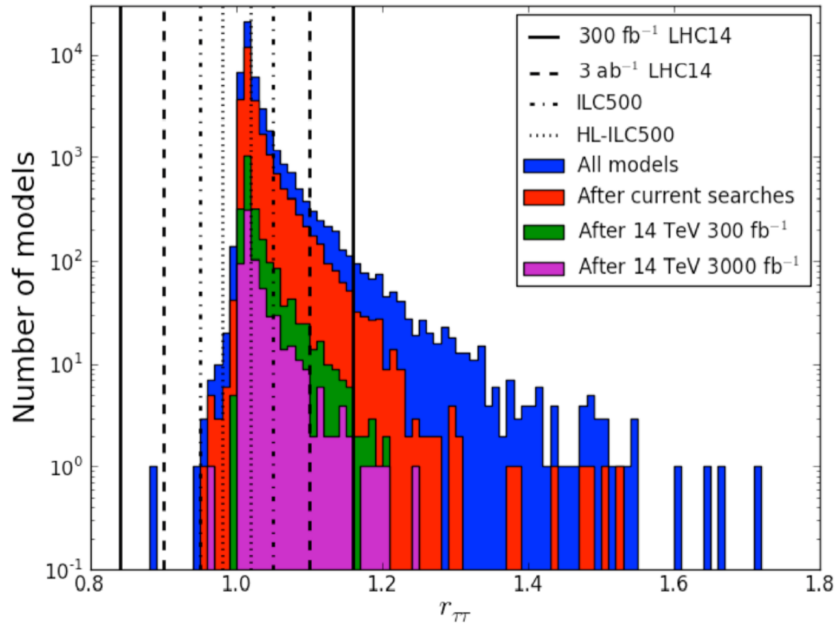
consistent with the expected decoupling.

In supersymmetric models, the Higgs couplings also receive corrections from loop diagrams involving the partners of the quarks and leptons. Typically, the largest effects come from diagrams with the  $b$  squarks and the gluino. These diagrams obey decoupling, but they are enhanced when  $\tan \beta$  is large.

Figure 41 shows the distribution of effects on the Higgs coupling  $g(hbb)$  seen in a large collection of supersymmetric models constructed by Cahill-Rowley, Hewett, Ismail, and Rizzo [74]. The colored panels in the figure show the sensitivity of the models to searches for supersymmetric particles at the LHC. It is interesting that the constraint from a precision measurement of the Higgs coupling to  $b\bar{b}$  is essentially orthogonal to the current and expected constraints from LHC searches. Thus, the precision study of Higgs couplings gives us a new and different way to probe for new physics. Figure 42 shows the comparable distribution for perturbations of the coupling  $g(h\tau\tau)$ .



**Fig. 41:** Values of  $r_{bb} = \Gamma(h \rightarrow b\bar{b})/SM$  in a collection of about 250,000 allowed parameter points of the Minimal Supersymmetric Standard Model, from [74]. The colored bands show models that can be discovered in new particle searches in the various stages of the LHC and the HL-LHC.



**Fig. 42:** Values of  $r_{\tau\tau} = \Gamma(h \rightarrow \tau^+\tau^-)/SM$  in a collection of about 250,000 allowed parameter points of the Minimal Supersymmetric Standard Model, from [74]. The colored bands show models that can be discovered in new particle searches in the various stages of the LHC and the HL-LHC.

It is important to note that, while the presence of multiple Higgs doublets can have significant effects on the Higgs couplings to fermions, it typically has a smaller effect on the Higgs couplings to the  $W$  and  $Z$  bosons. In the minimal supersymmetric model,

$$g(hVV) = \frac{2m_V^2}{v} \cdot \left(1 + \mathcal{O}\left(\frac{m_Z^4}{m_A^4}\right)\right) \quad (237)$$

for  $V = W, Z$ .

However, there are many other scenarios in which the Higgs couplings to  $W$  and  $Z$  are shifted as much as possible consistent with the Decoupling Theorem. If the Higgs boson mixes with a Higgs singlet field of mass  $m_s$  by an angle  $\gamma$ , the whole set of Higgs couplings is shifted by

$$g(hVV) = \frac{2m_V^2}{v} \cdot \cos \gamma \quad (238)$$

where, typically,  $\gamma \sim m_h/m_s$ . A similar effect is produced by loop corrections from any new particles that modify the Higgs boson self-energy diagrams [75, 76].

If the Higgs boson is a composite Goldstone boson, the Higgs couplings are corrected in a similar way by the nonlinear Lagrangian generated by spontaneous symmetry breaking. This gives

$$g(hVV) = \frac{2m_V^2}{v} \cdot (1 - v^2/F^2)^{1/2} \approx \frac{2m_V^2}{v} \cdot \left(1 - \frac{1}{2}v^2/F^2\right), \quad (239)$$

an effect of 1–3%.

We have seen in the previous section that the decays

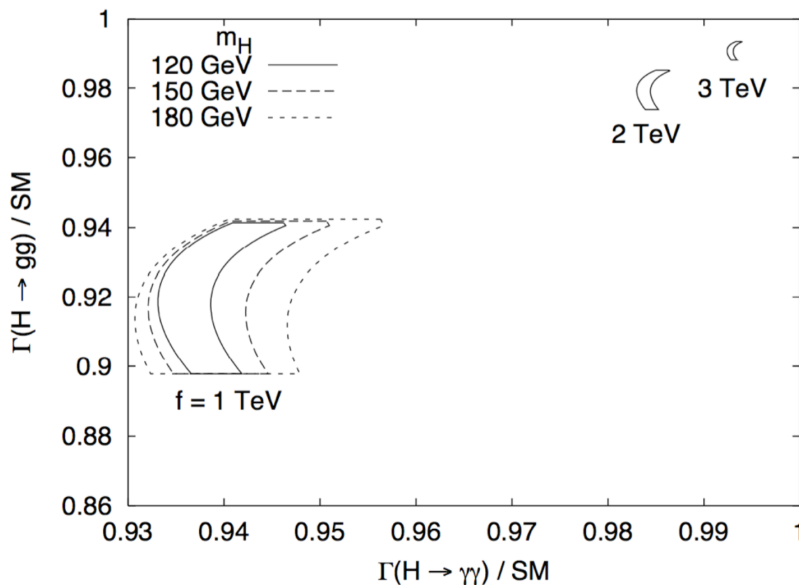
$$h \rightarrow gg, \quad h \rightarrow \gamma\gamma, \quad h \rightarrow \gamma Z \quad (240)$$

proceed through loop diagrams in which the dominant contributions come from particles for which  $2M > m_h$ . This means that new heavy particles have the potential to make large corrections to the rates of these decays. But this would only be true for particles that obtain their full mass from electroweak symmetry breaking.

As we have discussed already, the LHC measurement of  $pp \rightarrow h \rightarrow \gamma\gamma$  already excludes a conventional fourth generation of quarks and lepton, up the mass at which the Yukawa coupling exceeds the unitarity bound. Any fermions that we have not yet discovered must then be vectorlike fermions, with equal electroweak quantum numbers for the left- and right-handed fields. Such fermions can obtain an  $SU(2) \times U(1)$ -invariant mass term that does not require the Higgs field vacuum expectation value. For example, in models with extra space dimensions, excitations in the extra dimensions lead to separate Dirac fermion partners for the left- and right-handed states, which obtain masses  $M \sim \pi/R$ , where  $R$  is the size of the extra dimensions. The Higgs field can mix these states, leading to a small correction  $\delta M$  to the mass matrix that depends on the Higgs vacuum expectation value. The relative shift in the masses due to the Higgs vacuum expectation value is of the order of  $(\delta M)^2/M^2$ , and so the contribution of these particles to loop decays of the Higgs is suppressed by this factor—just as we would expect from the decoupling theorem.

A similar effect is seen in Little Higgs models. These models typically contain several new heavy quarks, which also mix with the top quark. An estimate of the corrections to the loop decays in the “Littlest Higgs” model is shown in Fig. 43 [77]. Mixing with heavy states can also modify the top quark Yukawa coupling. To fully understand the origin of the effects, it is important to measure separately the Higgs-gluon coupling and the Higgs-top coupling. The LHC might provide some complementary information by measuring Higgs boson production from gluon fusion at large  $p_T$  [78].

The Higgs boson also has a self-coupling that determines the shape of the Higgs potential. This is something of a special case in the general story of the Higgs couplings. On one hand, the Higgs



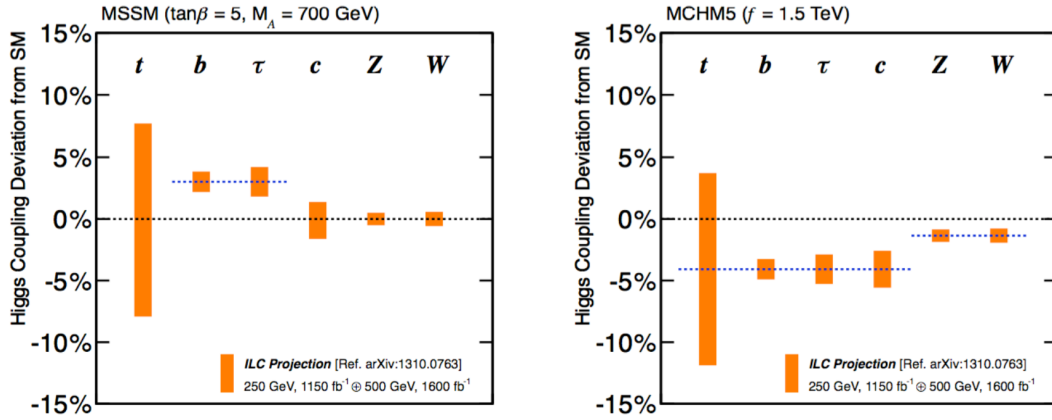
**Fig. 43:** Corrections to  $\Gamma(h \rightarrow \gamma\gamma)$  and  $\Gamma(h \rightarrow gg)$  in the Littlest Higgs model.

self-coupling is more difficult to measure. While there are realistic proposals to measure the other Higgs couplings to the percent level, it will already be difficult to measure the self-coupling to the level of 10–20% accuracy. On the other hand, there are models that require very large deviations of the Higgs self-coupling from its SM value. Theories of baryogenesis, the origin of the matter-antimatter asymmetry of the universe, require a period when the early universe was out of thermal equilibrium. We are confident that the nonzero Higgs field expectation value was established at a phase transition from a hot symmetric phase just after the Big Bang. In the SM, this phase transition is predicted to be second-order and thus too smooth for substantial out-of-equilibrium effects. If the Higgs phase transition were strongly first-order, then it is possible the the universe might have developed a baryon-antibaryon asymmetry through CP- and baryon number violating interactions available at that time [79]. This requires values of the Higgs self-coupling substantially different from that in the SM, a 50% increase or more [80].

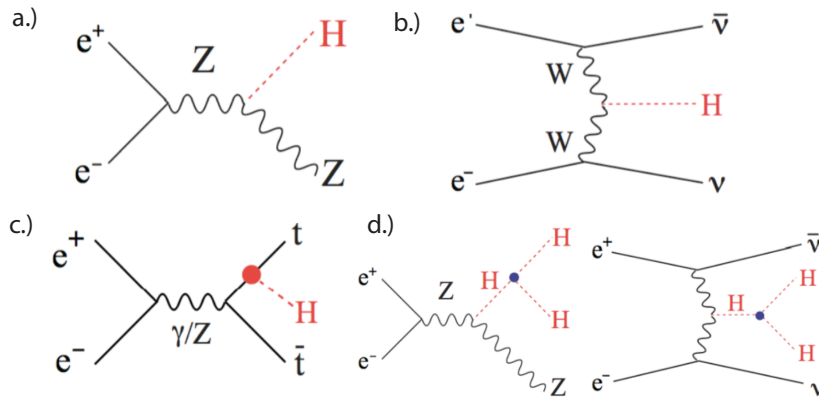
The result of this survey of new physics effects is that each individual Higgs coupling has its own personality and is guided by different types of models. In very broad terms:

- The Higgs couplings to fermions are sensitive to the presence of multiple Higgs doublets.
- The Higgs couplings to  $W$  and  $Z$  are sensitive to the presence of Higgs singlets and to compositeness of the Higgs boson.
- The Higgs couplings to  $gg$  and  $\gamma\gamma$  are sensitive to the presence of new vectorlike fermions.
- The Higgs coupling to  $t\bar{t}$  is sensitive to new heavy fermions that mix with the top quark and to composite structure of the top quark.
- The Higgs self-coupling has large deviations from its SM value in models of baryogenesis at the electroweak scale.

Each model of new physics predicts its own pattern of deviations of the Higgs couplings from the predictions of the SM. Two examples of these patterns, for specific supersymmetric and composite Higgs models, is shown in Fig. 44 [82]. The challenge for us is to measure the full suite of couplings with sufficient accuracy that we can read this pattern and use it to gain information about physics beyond the SM.



**Fig. 44:** Patterns of deviations in Higgs couplings, from [81]. These examples of nonstandard Higgs effects are taken from a broader survey in [82].



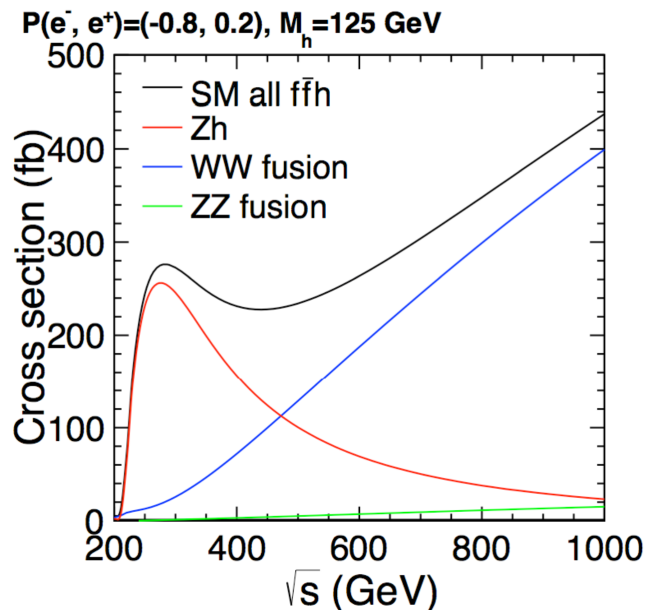
**Fig. 45:** Reactions producing the Higgs boson in  $e^+e^-$  collisions

### 7.5 Measurement of the Higgs boson properties at $e^+e^-$ colliders

Given the interest in obtaining precise knowledge of the couplings of the Higgs boson and the difficulty of reaching a sufficient level of accuracy at the LHC, it is not surprising that there are a number of proposals for new  $e^+e^-$  colliders that would specifically address the measurement of the Higgs couplings. It would be very valuable to study the Higgs boson with precision, in the same way that, in the 1990’s, experiments at  $e^+e^-$  colliders carried out the precision study of the  $Z$  boson that I reviewed in Section 4 of these lectures.

The most important processes for the production of a Higgs boson at  $e^+e^-$  colliders are those shown in Fig. 45. These are analogous to the corresponding processes in hadron-hadron collisions shown in Fig. 31. The most important reaction near the Higgs threshold is radiation of the Higgs boson from a  $W$  or  $Z$  (“Higgsstrahlung”). At higher energies, Higgs bosons are also produced by vector boson fusion, associated production of a Higgs with a pair of top quarks, and the double Higgs production reactions shown in the last line of the figure. The cross sections predicted for the Higgsstrahlung and fusion reactions for a 125 GeV Higgs boson are shown in Fig. 46.

Just as at hadron-hadron colliders, the different reactions available at  $e^+e^-$  colliders have different advantages for the study of Higgs boson decays. Higgsstrahlung is available at the lowest center of mass energy. In this reaction, the Higgs boson is produced in association with a  $Z$  boson at a fixed energy. At

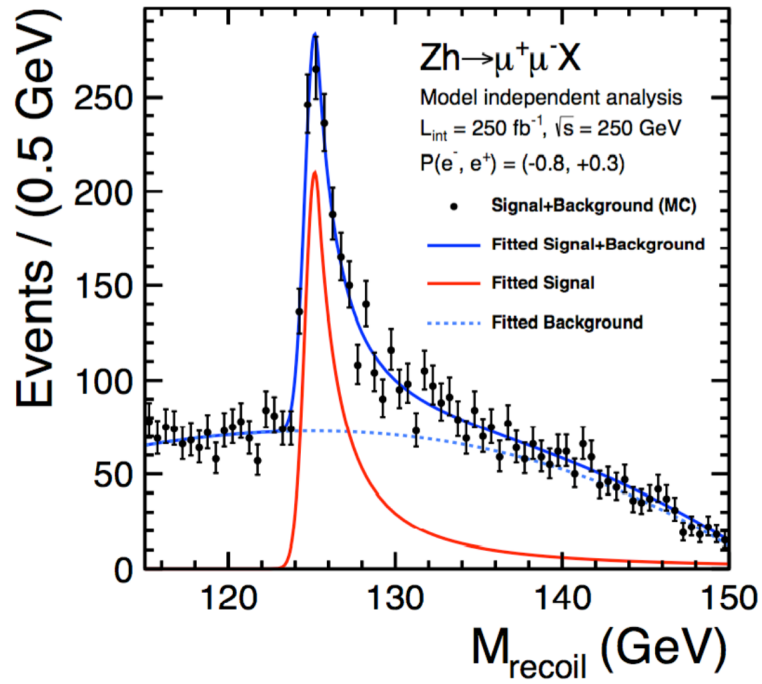


**Fig. 46:** Cross sections for Higgs production in  $e^+e^-$  collisions for a 125 GeV Higgs boson.

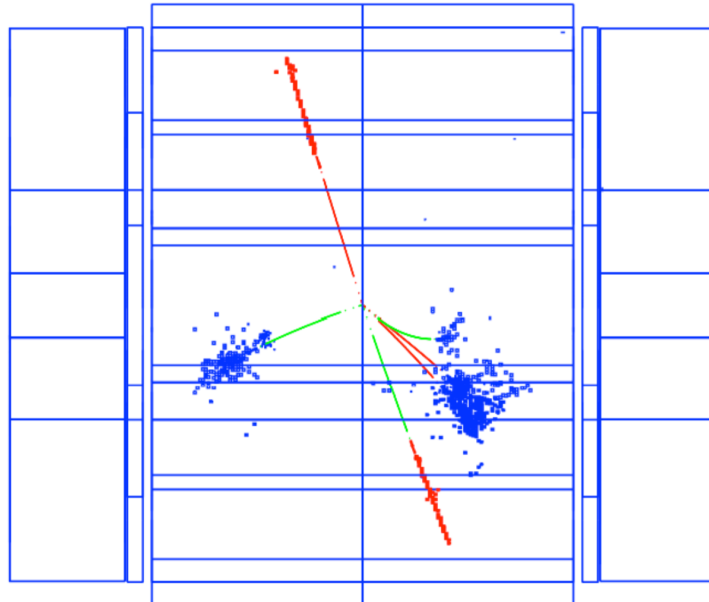
250 GeV in the center of mass, the  $Z$  boson has a lab frame energy of 110 GeV. To a first approximation, any  $Z$  boson observed at this energy arises from the reaction  $e^+e^- \rightarrow Zh$ , and whatever particles are on the other side of the event are the decay products of the Higgs boson. This is an ideal setup for measuring the branching ratios of the Higgs boson and for discovering and identifying Higgs decays into exotic modes not expected in the SM. Also, since  $e^+e^- \rightarrow Zh$  events can be recognized without reconstruction of the Higgs boson, this reaction allows a measurement of the absolute cross section rather than a  $\sigma \cdot BR$  as in (211). Then this reaction can be used to determine the absolute magnitude of the  $Z$ -Higgs coupling.

The remaining reactions have complementary advantages. Using the Higgs branching ratio to  $b\bar{b}$  measured with Higgsstrahlung, the  $WW$  fusion reaction can complement and firm up the measurement of the absolute normalization of Higgs couplings. As we see from Fig. 46, this reaction also gives higher statistics for Higgs decays at energies well above the threshold. The remaining processes allow the measurement of the Higgs coupling to top quarks and the Higgs self-coupling.

A complete description of the program of Higgs studies at  $e^+e^-$  colliders can be found in [83]. Here I will just provide some snapshots of this program. The recoil mass spectrum in the reaction  $e^+e^- \rightarrow Zh$ ,  $Z \rightarrow \mu^+\mu^-$  is shown in Fig. 47. The main background is  $e^+e^- \rightarrow ZZ$  plus initial state radiation, a reaction that is understood to very high accuracy. We estimate that this measurement gives the Higgs boson mass with an accuracy of 15 MeV [85]. The precision Higgs coupling program actually needs a Higgs boson mass with this high accuracy. The partial widths for  $h \rightarrow WW$  and  $h \rightarrow ZZ$  depend strongly on the Higgs mass, so that this accuracy already corresponds to a 0.1% systematic error on the SM predictions. Figure 48 shows a Higgsstrahlung event with Higgs decay to  $\tau^+\tau^-$ . In general, these events are very characteristic of the various  $Zh$  event topologies. Figure 49, from the physics study for the CLIC accelerator, shows the separation of Higgs events  $e^+e^-$  annihilation events at 250 GeV into 4 Higgs categories and one background category by template fitting [87]. The figure shows that the modes  $h \rightarrow gg$  and even  $h \rightarrow c\bar{c}$ , which has a 3% branching ratio in the SM, can be cleanly extracted. Figure 50 shows the recoil mass distribution for events with a  $Z$  boson plus missing momentum. The simulation assumes a high value (10%) for the Higgs branching ratio to invisible decay products, but the figure makes clear that this process is visible at much smaller values of the branching ratio, well below

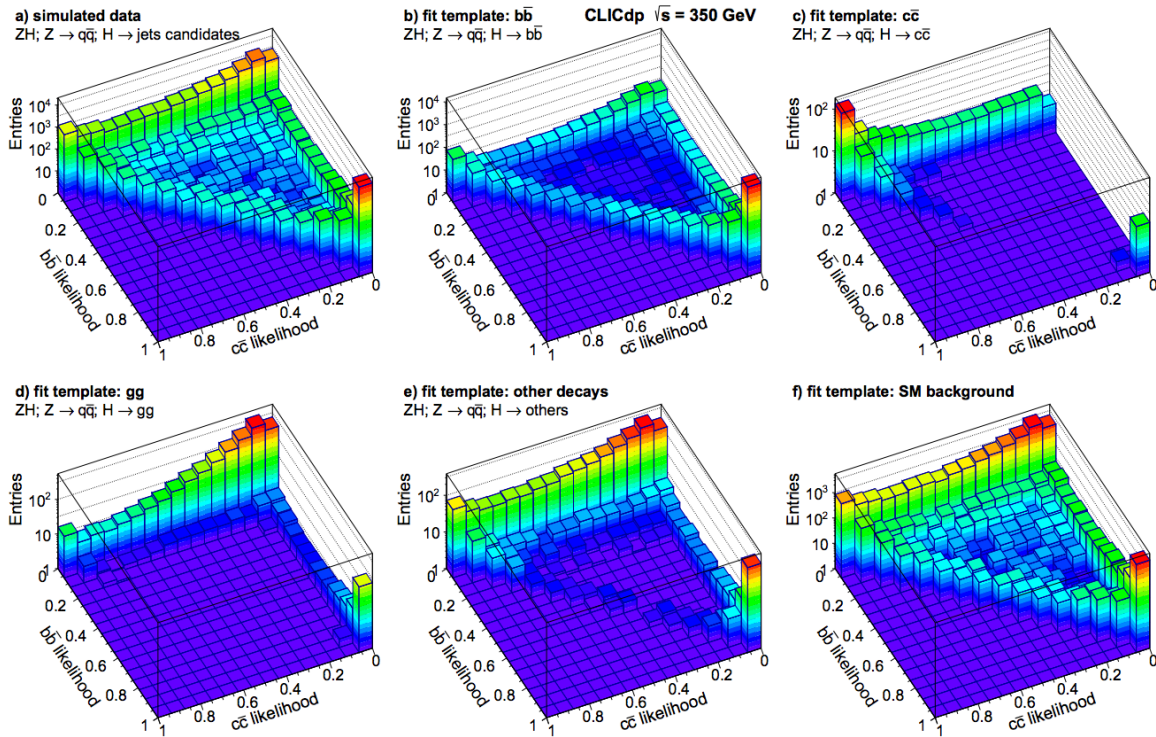


**Fig. 47:** Recoil mass distribution in  $e^+e^- \rightarrow Zh, Z \rightarrow \mu^+\mu^-$ , from [85]



**Fig. 48:** Event display of an  $e^+e^- \rightarrow Zh, h \rightarrow \tau^+\tau^-$  event simulated in the ILD detector [84].





**Fig. 49:** Identification of Higgs boson decays to hadronic final states by template fitting, from [87]. Note in particular the sharp discrimination of the modes  $h \rightarrow b\bar{b}$ ,  $h \rightarrow c\bar{c}$ ,  $h \rightarrow gg$ .

1% [86].

Finally, Figure 51, from [81], shows the accuracies for the determination of Higgs couplings to the full range of SM particles projected for the complete program of the International Linear Collider (ILC). For the Higgs decay to  $\gamma\gamma$ , the blue histograms show the result of combining the ILC data with the LHC measurement of  $BR(h \rightarrow \gamma\gamma)/BR(h \rightarrow ZZ^*)$ . The accuracy of the measurement of the Higgs coupling to the top quark is limited by the fact that this figure considers only ILC running at 500 GeV and below. Even an energy increase to 550 GeV would improve the accuracy of this measurement to 3%.

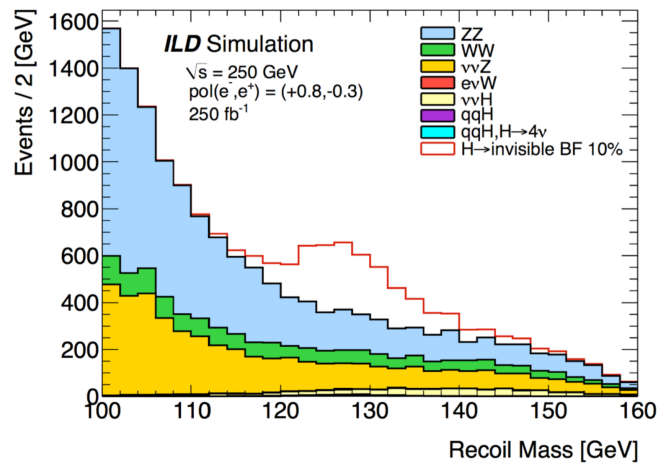
The precision study of Higgs boson couplings at an  $e^+e^-$  collider will then yield a wealth of information about the properties of this particle. Through the logic of the previous section, that information will give us insight not only into the existence of new physics beyond the SM but also into its qualitative nature. I look forward to this program as the next great project in the future of particle physics.

## 8 Conclusions

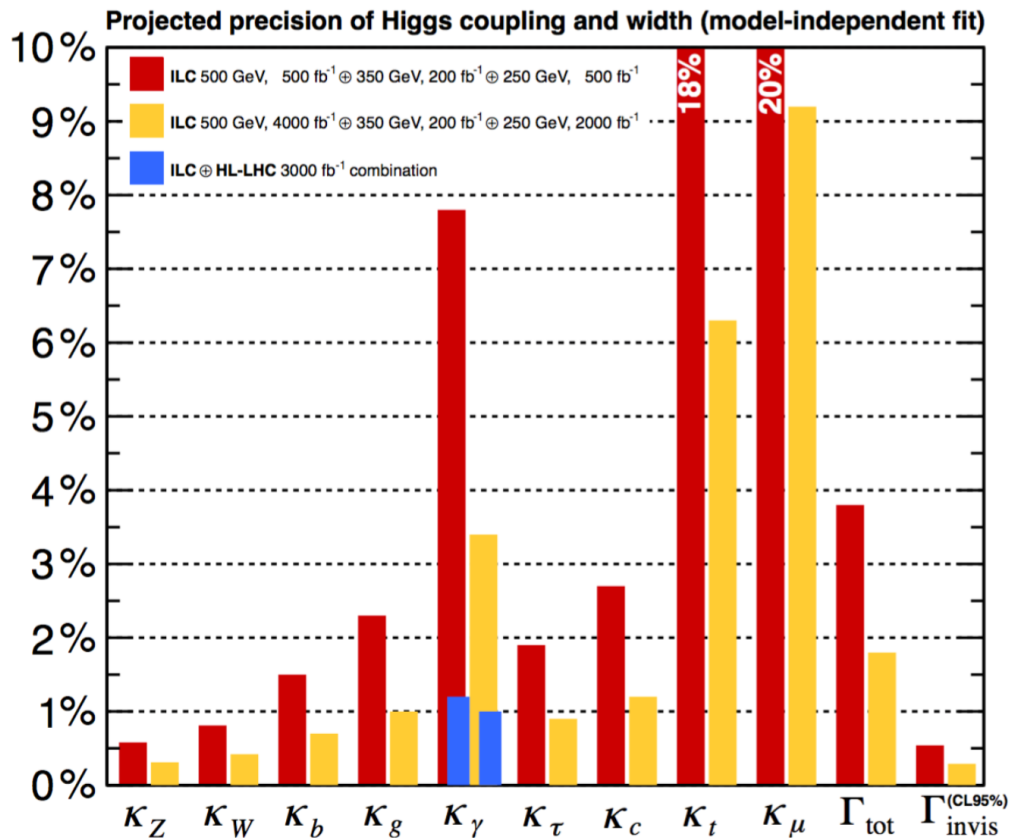
In these lectures, I have developed the theory of the weak interaction from its experimental foundations in the  $V-A$  effective theory, through the precision study of  $SU(2) \times U(1)$  couplings at the  $Z$  resonance, to the present and future study of the couplings of the Higgs boson. We have learned much about this fundamental interaction of nature, but there is much more that we need to learn, and that we can learn from future experiments. The study of the weak interaction is not a closed subject but one that still contains tantalizing questions and promises to open new chapters in our exploration of particle physics.

## References

- [1] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).



**Fig. 50:** Recoil mass distribution for  $e^+e^- \rightarrow Z + \text{missing}$  events, assuming a 10% branching ratio of the Higgs boson into invisible modes, from [86].



**Fig. 51:** Higgs coupling uncertainties projected for the ILC, from [81].

- [2] A. Salam, in *Proceedings of the 8th Nobel Symposium: Elementary Particle Theory*, N. Svartholm, ed. (Almqvist and Wiksell, 1968).
- [3] G. Aad *et al.* [ATLAS Collaboration], Phys. Lett. B **716**, 1 (2012) [arXiv:1207.7214 [hep-ex]].
- [4] S. Chatrchyan *et al.* [CMS Collaboration], Phys. Lett. B **716**, 30 (2012) [arXiv:1207.7235 [hep-ex]].
- [5] J. M. Cornwall, D. N. Levin and G. Tiktopoulos, Phys. Rev. D **10**, 1145 (1974) Erratum: [Phys. Rev. D **11**, 972 (1975)].
- [6] C. E. Vayonakis, Lett. Nuovo Cim. **17**, 383 (1976).
- [7] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. (Westview Press, 1995).
- [8] M. D. Schwartz, *Quantum Field Theory and the Standard Model*. (Cambridge University Press, 2013).
- [9] F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (1964).
- [10] P. W. Higgs, Phys. Lett. **12**, 132 (1964), Phys. Rev. Lett. **13**, 508 (1964).
- [11] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, Phys. Rev. Lett. **13**, 585 (1964).
- [12] C. Patrignani *et al.* (Particle Data Group), Chin. Phys. C **40**, 100001 (2016).
- [13] F. W. J. Koks and J. Van Klinken, Nucl. Phys. A **272**, 61 (1976).
- [14] M. Bardon, P. Norton, J. Peoples, A. M. Sachs and J. Lee-Franzini, Phys. Rev. Lett. **14**, 449 (1965).
- [15] J. G. H. de Groot *et al.*, Z. Phys. C **1**, 143 (1979).
- [16] G. Aad *et al.* [ATLAS Collaboration], Phys. Lett. B **701**, 31 (2011) [arXiv:1103.2929 [hep-ex]].
- [17] S. Chatrchyan *et al.* [CMS Collaboration], JHEP **1104**, 050 (2011) [arXiv:1103.3470 [hep-ex]].
- [18] P. Abreu *et al.* [DELPHI Collaboration], Eur. Phys. J. C **11**, 383 (1999).
- [19] I am grateful to Michael Hildreth for assembling this compilation of preliminary LEP 2 data.
- [20] S. Schael *et al.* [ALEPH and DELPHI and L3 and OPAL and LEP Electroweak Collaborations], Phys. Rept. **532**, 119 (2013) [arXiv:1302.3415 [hep-ex]].
- [21] E. A. Kuraev and V. S. Fadin, Sov. J. Nucl. Phys. **41**, 466 (1985) [Yad. Fiz. **41**, 733 (1985)].
- [22] S. Schael *et al.* [ALEPH and DELPHI and L3 and OPAL and SLD Collaborations and LEP Electroweak Working Group and SLD Electroweak Group and SLD Heavy Flavour Group], Phys. Rept. **427**, 257 (2006) [hep-ex/0509008].
- [23] O. Nicosini and L. Trentadue, Nucl. Phys. B **318**, 1 (1989).
- [24] G. Abbiendi *et al.* [OPAL Collaboration], Eur. Phys. J. C **19**, 587 (2001) [hep-ex/0012018]. I thank Toshinori Mori for the use of this figure.
- [25] D. Decamp *et al.* [ALEPH Collaboration], Z. Phys. C **48**, 365 (1990).
- [26] R. Akers *et al.* [OPAL Collaboration], Z. Phys. C **65**, 17 (1995).
- [27] K. Abe *et al.* [SLD Collaboration], Phys. Rev. Lett. **80**, 660 (1998) [hep-ex/9708015]. Phys. Rev. Lett. **94**, 091801 (2005) [hep-ex/0410042].
- [28] A. Heister *et al.* [ALEPH Collaboration], Eur. Phys. J. C **20**, 401 (2001) [hep-ex/0104038].
- [29] K. Abe *et al.* [SLD Collaboration], Phys. Rev. Lett. **86**, 1162 (2001) [hep-ex/0010015].
- [30] K. Abe *et al.* [SLD Collaboration], Phys. Rev. Lett. **81**, 942 (1998).
- [31] J. Erler and A. Freitas, in [12]
- [32] W. J. Marciano and A. Sirlin, Phys. Rev. D **29**, 945 (1984) Erratum: [Phys. Rev. D **31**, 213 (1985)].
- [33] M. E. Peskin and T. Takeuchi, Phys. Rev. Lett. **65**, 964 (1990), Phys. Rev. D **46**, 381 (1992).
- [34] M. Baak *et al.* [Gfitter Group], Eur. Phys. J. C **74**, 3046 (2014) [arXiv:1407.3792 [hep-ph]].
- [35] M. S. Chanowitz and M. K. Gaillard, Nucl. Phys. B **261**, 379 (1985).
- [36] V. Khachatryan *et al.* [CMS Collaboration], Phys. Lett. B **762**, 512 (2016) [arXiv:1605.09047 [hep-]]

- ex]].
- [37] W. Alles, C. Boyer and A. J. Buras, Nucl. Phys. B **119**, 125 (1977).
  - [38] K. Hagiwara, R. D. Peccei, D. Zeppenfeld and K. Hikasa, Nucl. Phys. B **282**, 253 (1987).
  - [39] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, JHEP **1010**, 085 (2010) [arXiv:1008.4884 [hep-ph]].
  - [40] S. Willenbrock and C. Zhang, Ann. Rev. Nucl. Part. Sci. **64**, 83 (2014) [arXiv:1401.0470 [hep-ph]].
  - [41] B. Henning, X. Lu and H. Murayama, JHEP **1601**, 023 (2016) [arXiv:1412.1837 [hep-ph]].
  - [42] A. Falkowski, M. Gonzalez-Alonso, A. Greljo, D. Marzocca and M. Son, arXiv:1609.06312 [hep-ph].
  - [43] S. Dawson, Nucl. Phys. B **249**, 42 (1985).
  - [44] J. F. Gunion, H. E. Haber, G. L. Kane and S. Dawson, *The Higgs Hunter's Guide*. (Westview Press, 2000).
  - [45] S. Heinemeyer *et al.* [LHC Higgs Cross Section Working Group], *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, arXiv:1307.1347 [hep-ph].
  - [46] CMS Collaboration, CMS-PAS-HIG-13-002 (2013).
  - [47] M. I. Vysotsky, arXiv:1312.0474 [hep-ph].
  - [48] M. Carena, C. Grojean, M. Kado, and V. Sharma, in K. A. Olive *et al.* (Particle Data Group), Chin. Phys. **C38**, 090001 (2014).
  - [49] V. Khachatryan *et al.* [CMS Collaboration], Eur. Phys. J. C **74**, no. 10, 3076 (2014) [arXiv:1407.0558 [hep-ex]].
  - [50] G. Aad *et al.* [ATLAS Collaboration], Phys. Lett. B **726**, 88 (2013) Erratum: [Phys. Lett. B **734**, 406 (2014)] [arXiv:1307.1427 [hep-ex]].
  - [51] G. Aad *et al.* [ATLAS Collaboration], Phys. Rev. D **92**, 012006 (2015) [arXiv:1412.2641 [hep-ex]].
  - [52] A. M. Sirunyan *et al.* [CMS Collaboration], arXiv:1708.00373 [hep-ex].
  - [53] ATLAS Collaboration, ATLAS-CONF-2013-108 (2013).
  - [54] M. Aaboud *et al.* [ATLAS Collaboration], arXiv:1708.03299 [hep-ex].
  - [55] CMS Collaboration, CMS-PAS-HIG-17-010 (2017).
  - [56] M. Carena, C. Grojean, M. Kado, and V. Sharma, in C. Patrignani *et al.* (Particle Data Group), Chin. Phys. **C40**, 100001 (2016).
  - [57] H. K. Onnes, Commun. Phys. Lab. Univ. Leiden. **12**, 120 (1911).
  - [58] V. L. Ginzburg and L. D. Landau, Zh. Eksp. Teor. Fiz. **20**, 1064 (1950).
  - [59] J. Bardeen, L. N. Cooper and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).
  - [60] S. P. Martin, in *Perspectives on Supersymmetry*, G. L. Kane, ed. (World Scientific, 1998) [hep-ph/9709356].
  - [61] M. E. Peskin, in *Exploring New Frontiers Using Collider and Neutrinos (TASI 2006)*, S. Dawson and R. Mohapatra, eds. (World Scientific, 2008) [arXiv:0801.1928 [hep-ph]].
  - [62] R. Contino, Y. Nomura and A. Pomarol, Nucl. Phys. B **671**, 148 (2003) [hep-ph/0306259].
  - [63] T. Gherghetta, in *Physics of the Large and the Small (TASI 09)*, C. Csaki and S. Dodelson, eds. (World Scientific, 2011) [arXiv:1008.2570 [hep-ph]].
  - [64] P. Sikivie, L. Susskind, M. B. Voloshin and V. I. Zakharov, Nucl. Phys. B **173**, 189 (1980).
  - [65] H. Georgi and M. Machacek, Nucl. Phys. B **262**, 463 (1985).
  - [66] H. E. Logan and V. Rantala, Phys. Rev. D **92**, no. 7, 075011 (2015) [arXiv:1502.01275 [hep-ph]].
  - [67] S. Weinberg, Phys. Rev. D **19**, 1277 (1979).
  - [68] L. Susskind, Phys. Rev. D **20**, 2619 (1979).
  - [69] N. Arkani-Hamed, A. G. Cohen, E. Katz and A. E. Nelson, JHEP **0207**, 034 (2002) [hep-

- ph/0206021].
- [70] M. Schmaltz and D. Tucker-Smith, *Ann. Rev. Nucl. Part. Sci.* **55**, 229 (2005) [hep-ph/0502182].
  - [71] H. E. Haber, in *Electroweak Symmetry Breaking*, F. Csikor and G. Pocsik, eds. (World Scientific, 1995) [hep-ph/9501320].
  - [72] R. K. Dewanjee [CMS Collaboration], *PoS DIS* **2016**, 100 (2016).
  - [73] ATLAS Collaboration, *ATLAS-CONF-2017-050* (2017).
  - [74] M. Cahill-Rowley, J. Hewett, A. Ismail and T. Rizzo, *Phys. Rev. D* **90**, 095017 (2014) [arXiv:1407.7021 [hep-ph]].
  - [75] V. Barger, T. Han, P. Langacker, B. McElrath and P. Zerwas, *Phys. Rev. D* **67**, 115001 (2003) [hep-ph/0301097].
  - [76] N. Craig, C. Englert and M. McCullough, *Phys. Rev. Lett.* **111**, 121803 (2013) [arXiv:1305.5251 [hep-ph]].
  - [77] T. Han, H. E. Logan, B. McElrath and L. T. Wang, *Phys. Lett. B* **563**, 191 (2003) Erratum: [*Phys. Lett. B* **603**, 257 (2004)] [hep-ph/0302188].
  - [78] C. Grojean, E. Salvioni, M. Schlaffer and A. Weiler, *JHEP* **1405**, 022 (2014) [arXiv:1312.3317 [hep-ph]].
  - [79] D. E. Morrissey and M. J. Ramsey-Musolf, *New J. Phys.* **14**, 125003 (2012) [arXiv:1206.2942 [hep-ph]].
  - [80] A. Noble and M. Perelstein, *Phys. Rev. D* **78**, 063518 (2008) [arXiv:0711.3018 [hep-ph]].
  - [81] K. Fujii *et al.*, arXiv:1506.05992 [hep-ex].
  - [82] S. Kanemura, K. Tsumura, K. Yagyu and H. Yokoya, *Phys. Rev. D* **90**, 075001 (2014) [arXiv:1406.3294 [hep-ph]].
  - [83] H. Baer *et al.*, *The International Linear Collider Technical Design Report - Volume 2: Physics*, arXiv:1306.6352 [hep-ph].
  - [84] I thank Manqi Ruan for the use of this figure.
  - [85] J. Yan, S. Watanuki, K. Fujii, A. Ishikawa, D. Jeans, J. Strube, J. Tian and H. Yamamoto, *Phys. Rev. D* **94**, 113002 (2016) [arXiv:1604.07524 [hep-ex]].
  - [86] J. Tian and K. Fujii, *Nucl. Part. Phys. Proc.* **273-275**, 826 (2016).
  - [87] H. Abramowicz *et al.*, *Eur. Phys. J. C* **77**, no. 7, 475 (2017) [arXiv:1608.07538 [hep-ex]].

## Introduction to Flavour Physics and CP Violation

*M. Blanke*

Institut für Kernphysik, Karlsruhe Institute of Technology, Eggenstein-Leopoldshafen, Germany  
Institut für Theoretische Teilchenphysik, Karlsruhe Institute of Technology, Karlsruhe, Germany

### Abstract

These lecture notes provide an introduction to the theoretical concepts of flavour physics and  $CP$  violation, based on a series of lectures given at the ESHEP 2016 summer school. In the first lecture we review the basics of flavour and  $CP$  violation in the Standard Model. The second lecture is dedicated to the phenomenology of  $K$  and  $B$  meson decays, where we focus on a few representative observables. In the third lecture we give an introduction to flavour physics beyond the Standard Model, both within the framework of Minimal Flavour Violation and beyond.

### Keywords

Lectures; CKM matrix; CP violation; flavour changing neutral currents.

### Introduction

Flavour physics and  $CP$  violation have played a central role in the development of the Standard Model (SM). It was the underlying  $SU(3)$  flavour symmetry of mesons and baryons that lead Gell-Mann to the introduction of up, down, and strange quarks as the fundamental constituents of hadronic matter [1]. The charm quark was predicted prior to its discovery as an explanation for the smallness of the  $K_L \rightarrow \mu^+ \mu^-$  decay rate [2]. It was then realized that, in order to account for the observed  $CP$  violation in neutral kaon mixing, a third generation of quarks was needed [3]. Furthermore, the heaviness of the top quark was predicted from the size of  $CP$  violation in  $K^0 - \bar{K}^0$  mixing and of the neutral  $B$  meson oscillations prior to its discovery [4].

Subsequently, the role of flavour physics has shifted from the discovery of the building blocks of the SM to the measurement of its parameters. The majority of the SM parameters is related to the flavour sector and can thus be determined in flavour violating decays. With increasing experimental and theoretical accuracy, their determination has by now reached an impressive precision.

Having at hand a good understanding of the SM flavour sector, the measurement of flavour and  $CP$ -violating processes can be used to put constraints on models of New Physics (NP). Due to the strong suppression of flavour violation in the SM, very high energy scales can be probed in this way, well beyond the reach of direct searches for new particles in high energy collisions at the Large Hadron Collider (LHC) (see [5] for a recent review). NP around the TeV scale is therefore required to have a highly non-trivial flavour structure.

The present lecture notes provide a summary of a series of three lectures given at the European School for High Energy Physics (ESHEP) 2016 in Skeikampen, Norway. The topic of these lectures is restricted to quark flavour physics. Flavour in the charged lepton and neutrino sector has been covered by Gabriela Barenboim [6]. In lecture 1 we introduce the quark Yukawa couplings and the Cabibbo-Kobayashi-Maskawa (CKM) matrix as the basic ingredients of flavour and  $CP$  violation in the SM. We then discuss the physics of flavour changing neutral currents and review their description in terms of an effective Hamiltonian and the path from quark level flavour transitions to the decays of mesons. Lecture 2 is devoted to the phenomenology of flavour and  $CP$ -violating decays of kaons and  $B$  mesons. Rather than providing an exhaustive overview, we focus on a number of particularly interesting benchmark processes, like neutral meson mixings, rare  $K$  meson decays, and the recently observed anomalies in

$B$  meson decays based on semileptonic  $b \rightarrow s$  transitions. In lecture 3 we turn our attention to flavour physics beyond the SM. After reviewing the generic constraints on the scale of NP from flavour and  $CP$ -violating decays, we discuss two well-known suppression mechanisms for the size of new flavour violating interactions. We motivate the concept of Minimal Flavour Violation (MFV) from the flavour symmetries of the SM. As an alternative to MFV, we also discuss how models with partially composite fermions can explain the observed flavour hierarchies in the SM, and at the same time suppress flavour changing neutral currents to an acceptable level.

A large number of excellent lecture notes on the physics of flavour and  $CP$  violation can be found on the arXiv. As a few representative examples, let me recommend especially the lectures by Gino Isidori [7], Yuval Grossmann [8], and Andrzej j. Buras [9]. An extensive pedagogical introduction into the technicalities of the theory of flavour physics can be found in [10].

## 1 Flavour Physics in the Standard Model

### 1.1 Quark Yukawa Couplings and the CKM Matrix

In nature, all fundamental matter fields – quarks, charged leptons, and neutrinos – come in three copies, the so-called *flavours*. They can be collected in three fermion generations, with increasing masses, but otherwise identical quantum numbers. The subject of flavour physics is the description of interactions between the various flavours, with the goal to unravel the underlying dynamics of flavour symmetry breaking.

In the SM, the left-handed quarks are arranged in doublets of the  $SU(2)_L$  weak interactions:

$$Q_j = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix}, \quad (1)$$

while the right-handed quarks are introduced as  $SU(2)_L$  singlets:

$$U_j = u_R, c_R, t_R \quad D_j = d_R, s_R, b_R. \quad (2)$$

The quarks' couplings to the gluons, weak gauge bosons  $W^\pm$  and  $Z$ , and the photon are described by the kinetic term in the Lagrangian

$$\mathcal{L}_{\text{fermion}} = \sum_{j=1}^3 \bar{Q}_j i \not{D}_Q Q_j + \bar{U}_j i \not{D}_U U_j + \bar{D}_j i \not{D}_D D_j, \quad (3)$$

with the covariant derivatives

$$D_{Q,\mu} = \partial_\mu + ig_s T^a G_\mu^a + ig \tau^a W_\mu^a + ig' Q_Q^Y B_\mu, \quad (4)$$

$$D_{U,\mu} = \partial_\mu + ig_s T^a G_\mu^a + ig' Q_U^Y B_\mu, \quad (5)$$

$$D_{D,\mu} = \partial_\mu + ig_s T^a G_\mu^a + ig' Q_D^Y B_\mu, \quad (6)$$

and the hypercharges assigned as  $Q_Q^Y = 1/6$ ,  $Q_U^Y = 2/3$ ,  $Q_D^Y = -1/3$ .  $T^a (a = 1, \dots, 8)$  and  $\tau^a (a = 1, 2, 3)$  are the generators of  $SU(3)_c$  and  $SU(2)_L$ , respectively, and the index  $j$  runs over the three generations of quark fields. It is evident that the gauge couplings are universal for all three generations.

Flavour non-universality, on the other hand, is introduced by the quark Yukawa couplings to the Higgs field, responsible for the generation of non-zero quark masses:

$$\mathcal{L}_{\text{Yuk}} = \sum_{i,j=1}^3 (-Y_{U,ij} \bar{Q}_{Li} \tilde{H} U_{Rj} - Y_{D,ij} \bar{Q}_{Li} H D_{Rj} + h.c.), \quad (7)$$



where *h.c.* abbreviates the hermitian conjugate term. The subscripts  $i, j$  are generation indices, and the dual field  $\tilde{H}$  is given as  $\tilde{H} = \epsilon H^* = (H^{0*}, -H^-)^T$ . Replacing the Higgs field  $H$  by its vacuum expectation value  $\langle H \rangle = (0, v)^T$ , we obtain the quark mass terms

$$\sum_{i,j=1}^3 (-m_{U,ij} \bar{u}_{Li} u_{Rj} - m_{D,ij} \bar{d}_{Li} d_{Rj} + h.c.), \quad (8)$$

with the quark mass matrices given by  $m_{U,D} = vY_{U,D}$ .

The quark mass matrices  $m_U$  and  $m_D$  are  $3 \times 3$  complex matrices in flavour space with *a priori* arbitrary entries. They can be diagonalized by making appropriate bi-unitary field redefinitions:

$$u_L = \hat{U}_L u_L^m, \quad u_R = \hat{U}_R u_R^m, \quad d_L = \hat{D}_L d_L^m, \quad d_R = \hat{D}_R d_R^m, \quad (9)$$

with the superscript  $m$  denoting quarks in their mass eigenstate basis.

Is the SM Lagrangian invariant under these transformations? Unitary transformations of the right-handed quark sector are indeed unphysical, as they drop out from the rest of the Lagrangian. However,  $u_{Li}$  and  $d_{Li}$  form the  $SU(2)_L$  doublets  $Q_i$  (with  $i = 1, 2, 3$ ). Their kinetic term gives rise to the interaction

$$\frac{g}{\sqrt{2}} \bar{u}_{Li} \gamma_\mu W^{\mu+} d_{Li}. \quad (10)$$

Transformation of (10) to the mass eigenstate basis yields

$$\frac{g}{\sqrt{2}} \bar{u}_{Li} \hat{U}_{L,ij}^\dagger \hat{D}_{L,jk} \gamma_\mu W^{\mu+} d_{Lk}. \quad (11)$$

We conclude that the combination

$$\hat{V}_{\text{CKM}} = \hat{U}_L^\dagger \hat{D}_L \quad (12)$$

is physical, it is called the *CKM matrix* [3, 11]. It describes the misalignment between left-handed up- and down-type quark mass eigenstates, which leads to flavour violating charged current interactions, mediated by the  $W^\pm$  bosons. It is convenient to label the elements of  $\hat{V}_{\text{CKM}}$  by the quark flavours involved in the respective charged current interaction:

$$\hat{V}_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \quad (13)$$

For example, the element  $V_{ub}$  appears in the coupling of a bottom and an up quark to the  $W$  boson.

## 1.2 Standard Parametrization of the CKM Matrix

Let us now determine the number of physical parameters in the CKM matrix. Being a unitary  $3 \times 3$  matrix, it can be parametrized by three mixing angles and six complex phases in general. However, five of these phases are unphysical, as they can be absorbed as unobservable parameters into the up-type and down-type quarks, respectively. Note that an overall phase rotation of all quarks does not affect the CKM matrix. We are then left with three mixing angles  $\theta_{12}, \theta_{23}, \theta_{13}$  and one complex phase  $\delta$  as the physical parameters of the CKM matrix. Introducing the short-hand notation  $s_{ij} = \sin \theta_{ij}$  and  $c_{ij} = \cos \theta_{ij}$ , the standard parametrization of the CKM matrix reads [12]

$$\hat{V}_{\text{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}. \quad (14)$$

Note that this parametrization is recommended by the Particle Data Group (PDG) [13].

Alternatively, the number of independent flavour parameters in the SM can also be determined from symmetry principles. Ignoring the Yukawa couplings, the SM quark sector has a global

$$G_{\text{flavour}} = U(3)_Q \times U(3)_U \times U(3)_D \quad (15)$$

flavour symmetry. The quark Yukawa couplings  $Y_U, Y_D$  explicitly break  $G_{\text{flavour}}$ , leaving only a single  $U(1)$  factor unbroken, that corresponds to the overall phase of the quark fields. This  $U(1)$  symmetry is associated to baryon number conservation, which is an accidental symmetry of the SM. We can use this symmetry breaking pattern to count the number of flavour parameters.

We start from the Yukawa couplings  $Y_U$  and  $Y_D$ . A priori, these are arbitrary complex  $3 \times 3$  matrices, hence they bring in nine real parameters and nine complex phases each. However, not all of these  $18 + 18$  parameters are physical. In fact, each of the broken generators of the flavour symmetry group  $G_{\text{flavour}}$  corresponds to an unphysical parameter in the Lagrangian which can be removed by making appropriate field redefinitions. A  $3 \times 3$  unitary matrix contains three real parameters and six complex phases. The three  $U(3)$  factors in  $G_{\text{flavour}}$  therefore carry nine real parameters and 18 phases. All but one of them, namely the phase corresponding to the unbroken overall  $U(1)$ , correspond to unphysical parameters that can be removed from  $Y_U$  and  $Y_D$ . We are then left with a total of nine real parameters in the quark flavour sector and one physical complex phase. The nine real parameters are the quark masses  $m_u, m_d, m_c, m_s, m_t, m_b$  and the three mixing angles  $\theta_{12}, \theta_{13}, \theta_{23}$  of the CKM matrix, and the phase is simply the CKM phase  $\delta$ .

Experimentally, it has been found that the CKM matrix exhibits a rather strong hierarchy, with [13]

$$s_{12} \sim 0.2, \quad s_{23} \sim 0.04, \quad s_{13} \sim 4 \cdot 10^{-3}. \quad (16)$$

The CKM matrix hence is close to the unit matrix, with hierarchical off-diagonal elements. Flavour changing transitions are therefore strongly suppressed in the SM. Similarly, also the quark masses are found to follow a hierarchical pattern, spanning five orders of magnitude in size. The lack of a more fundamental theory explaining the origin of this structure is referred to as the flavour hierarchy problem of the SM.

### 1.3 $CP$ violation in the SM

We have seen above that the angles  $\theta_{ij}$  of the CKM matrix parametrize the amount of flavour mixing between the quarks of the generations  $i$  and  $j$ . The amount of flavour violation in the SM is therefore quantified by the values of the CKM mixing angles. But what is the physical meaning of the presence of a complex phase  $\delta$ ?

In order to understand this, let us consider two discrete transformations:

- the parity transformation

$$P : \psi(r, t) \rightarrow \gamma^0 \psi(-r, t) \quad (17)$$

which transforms left-handed fermion fields into right-handed ones and vice versa, and

- the charge conjugation

$$C : \psi \rightarrow i(\bar{\psi} \gamma^0 \gamma^2)^T \quad (18)$$

which transforms left(right)-handed quarks into left(right)-handed antiquarks.

It is evident from equations (3)–(6) that the weak interactions violate both  $C$  and  $P$ , as they treat left- and right-handed quarks differently.

But what about the combination of both transformations,  $CP$ ? A  $CP$  transformation connects left-handed quarks to right-handed antiquarks. It is easy to convince oneself that the neutral current

interactions mediated by gluons, the photon and the  $Z$  boson are indeed invariant under  $CP$ . Let us then look at the charged current interactions mediated by the  $W^\pm$  bosons:

$$\begin{aligned}
 \mathcal{L}_{c.c.} &= \frac{g}{\sqrt{2}} V_{ik} \bar{u}_{Li} \gamma_\mu W^{\mu+} d_{Lk} + h.c. \\
 &= \frac{g}{\sqrt{2}} V_{ik} \bar{u}_{Li} \gamma_\mu W^{\mu+} d_{Lk} + \frac{g}{\sqrt{2}} V_{ik}^* \bar{d}_{Lk} \gamma_\mu W^{\mu-} u_{Li} \\
 \xrightarrow{CP} &\frac{g}{\sqrt{2}} V_{ik} \bar{d}_{Lk} \gamma_\mu W^{\mu-} u_{Li} + \frac{g}{\sqrt{2}} V_{ik}^* \bar{u}_{Li} \gamma_\mu W^{\mu+} d_{Lk} \\
 &= \frac{g}{\sqrt{2}} V_{ik}^* \bar{u}_{Li} \gamma_\mu W^{\mu+} d_{Lk} + h.c.
 \end{aligned} \tag{19}$$

We see that  $CP$  conjugation replaces the CKM element  $V_{ik}$  by its complex conjugate. Hence, the  $CP$  symmetry is violated in the SM by the presence of a non-vanishing complex phase  $\delta \neq 0$  in the CKM matrix.

It is important to note, however, that the phase  $\delta$  is not a physical parameter, as, by means of the aforementioned rephasing of quark fields, it can be shifted to different elements of the CKM matrix. A parametrization-independent and therefore physical measure of  $CP$  violation is instead given by the Jarlskog invariant [14, 15]

$$J_{CP} = \text{Im}(V_{us} V_{cb} V_{ub}^* V_{cs}^*). \tag{20}$$

Experimentally, the Jarlskog invariant is found to be  $J_{CP} \simeq 3 \cdot 10^{-5}$ .

#### 1.4 Flavour Changing Neutral Currents

We have seen above that flavour changing charged currents are present at the tree level in the SM, with the size of the interactions governed by the off-diagonal elements of the CKM matrix. Flavour changing neutral currents (FCNCs), on the other hand, are absent at the tree level in the SM. In order to see this, let us have a closer look at the  $Z$  boson coupling to left-handed down-type quarks, as an illustrative example. Transforming the coupling of the quark flavour eigenstates  $d_{Lj}$  into a coupling of the quark mass eigenstates  $d_{Li}^m$ , we find

$$\begin{aligned}
 \mathcal{L}_{n.c.} &\ni g_Z(d_L) \bar{d}_{Lj} \gamma_\mu Z^\mu d_{Lj} \\
 &= g_Z(d_L) \bar{d}_{Li}^m (\hat{D}_L^\dagger)_{ij} \gamma_\mu Z^\mu (\hat{D}_L)_{jk} d_{Lk}^m \\
 &= g_Z(d_L) \bar{d}_{Li}^m \gamma_\mu Z^\mu \delta_{ik} d_{Lk}^m,
 \end{aligned} \tag{21}$$

where

$$g_Z(d_L) = \frac{g}{\cos \theta_W} \left( -\frac{1}{2} + \frac{1}{3} \sin^2 \theta_W \right) \tag{22}$$

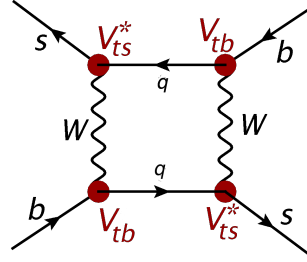
is the  $Z$  boson coupling to left-handed down-type quarks. We can see that due to the unitarity of the flavour rotation matrix  $\hat{D}_L$ , the coupling remains flavour diagonal and flavour universal. The same argument holds for all neutral gauge boson couplings in the SM.

FCNCs are however generated in the SM by loop diagrams with internal  $W^\pm$  bosons. As an example, figure 1 shows the one loop diagram that generates the leading contribution to neutral  $B_s$  meson mixing. The mixing amplitude generated by this contribution is schematically given by

$$\mathcal{M} \propto \sum_{i,j=u,c,t} V_{is}^* V_{ib} V_{js}^* V_{jb} F(x_i, x_j). \tag{23}$$

Here,  $F(x_i, x_j)$  is the relevant one loop function, with  $x_i = (m_i/M_W)^2$ , and the double sum runs over the internal quark flavours. Using the CKM unitarity

$$\sum_{i=u,c,t} V_{is}^* V_{ib} = 0, \tag{24}$$



**Fig. 1:** One loop diagram for  $B_s - \bar{B}_s$  mixing in the SM.

setting  $m_u = 0$  and neglecting contributions proportional to  $V_{cs}^* V_{cb} \ll V_{ts}^* V_{tb}$ , this can be simplified to

$$\mathcal{M} \propto (V_{ts}^* V_{tb})^2 S_0(x_t), \quad (25)$$

with

$$S_0(x_t) = F(x_t, x_t) - 2F(0, x_t) + F(0, 0). \quad (26)$$

We can see that, indeed, FCNCs are generated by loop processes in the SM. However they are suppressed not only by the smallness of the off-diagonal CKM elements, but also by the so-called *GIM mechanism* [2]: All contributions that are independent of the masses of the quarks running in the loop are cancelled by the unitarity of the CKM matrix, and only the differences of mass-dependent terms survive. While above we have seen the GIM mechanism at work for one loop contributions, it in fact holds to all orders.

### 1.5 The Unitarity Triangle

The hierarchical structure of the CKM matrix can be used to derive an alternative parametrization, which turns out to be very useful for estimating the size of flavour violating transitions. In the Wolfenstein parametrization [16]

$$\hat{V}_{\text{CKM}} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \quad (27)$$

$\lambda = |V_{us}| \sim 0.2$  is the only small parameter, while  $A$ ,  $\rho$ , and  $\eta$  are  $\mathcal{O}(1)$ . It is therefore convenient to estimate the size of flavour violating decays by making an expansion in powers of  $\lambda$ . The accuracy of this expansion can be improved by changing the parameters of the Wolfenstein parametrization to [17, 18]

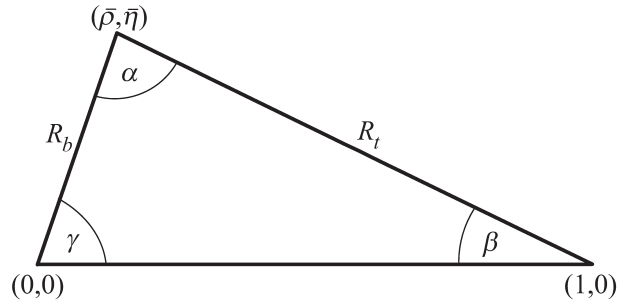
$$\lambda, \quad A, \quad \bar{\rho} = \left(1 - \frac{\lambda^2}{2}\right)\rho, \quad \bar{\eta} = \left(1 - \frac{\lambda^2}{2}\right)\eta. \quad (28)$$

As discussed before, the CKM matrix is a unitary matrix, and not all of its elements are independent parameters. Various relations hold among them, which can be tested experimentally. One of the most popular ones,

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \quad (29)$$

can be displayed as a triangle in the complex plane, the so-called *unitarity triangle* (UT) [19]. With the base of the UT normalized to unity, the apex is simply given by  $(\bar{\rho}, \bar{\eta})$ . The sides  $R_b$  and  $R_t$ , as shown in figure 2, are given by

$$R_b = \left| \frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \right| = \left(1 - \frac{\lambda^2}{2}\right) \frac{1}{\lambda} \frac{|V_{ub}|}{|V_{cb}|}, \quad (30)$$


**Fig. 2:** Unitarity triangle.

$$R_t = \left| \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} \right| = \frac{1}{\lambda} \frac{|V_{td}|}{|V_{cb}|}. \quad (31)$$

For the UT angles, two notations are commonly used in the literature. They are related to each other as follows:

$$\alpha \equiv \phi_2, \quad \beta \equiv \phi_1, \quad \gamma \equiv \phi_3. \quad (32)$$

The UT can be determined experimentally from various measurements of flavour violating decays of  $K$  and  $B$  mesons. A special role in this determination is played by the length of the side  $R_b$  and the angle  $\gamma$ : Being sensitive to the absolute values and  $CP$ -violating phases of the elements  $V_{ub}$  and  $V_{cb}$ , they can be determined from  $B$  decays governed by tree level charged current interactions. It is therefore a good approximation to assume that NP contributions to these measurements are negligible. The measurement of  $|V_{ub}|$ ,  $|V_{cb}|$  and  $\gamma = \arg(-V_{ud}V_{ub}^*/V_{cd}V_{cb}^*)$  then leaves us with the *reference unitarity triangle* [20], which determines the CKM matrix independently of potential NP contributions to rare flavour violating decays.

The length of the side  $R_t$  and the angle  $\beta$ , on the other hand, depend on CKM elements involving the top quark. Hence, they can only be measured in loop-induced flavour changing neutral current (FCNC) processes. Due to their strong suppression in the SM, these observables are sensitive to NP contributions. A model-independent determination of the CKM matrix using these quantities is therefore not possible. NP contributions to the loop induced processes used in the determination of the UT

The strategy to hunt for NP contributions to flavour violating observables is then as follows. First, the CKM matrix and the UT have to be determined from tree level charged current decays as accurately as possible. As this determination is independent of potential NP contributions, the result can be used as input for precise SM predictions of rare, loop-induced FCNC processes. These predictions are then to be compared with the data, which – in case of a discrepancy – would yield an unambiguous sign of a NP contribution to the decay in question. Clearly, in order to be able to claim a NP discovery in flavour violating observables, a solid understanding of the SM contribution and its uncertainties is mandatory.

## 1.6 The effective Hamiltonian

An important theoretical complication arises in the study of quark flavour violating decays. Due to the confinement of QCD at low energies, quarks do not appear as free particles in nature, but are bound in hadrons. Therefore, not only the weak interactions leading to flavour violation, as discussed above, have to be well understood, but also the strong dynamics describing the bound states of QCD. The latter interactions, taking place at the typical hadronic energy scales of a few hundred MeV to a few GeV, are non-perturbative and hence, with our current methods, cannot be calculated analytically.

A convenient theoretical tool to handle these various contributions from processes at different energy scales is provided by the operator product expansion [21]. In this framework, effective flavour violating operators are obtained from integrating out the heavy electroweak (EW) gauge bosons  $W^\pm$  and

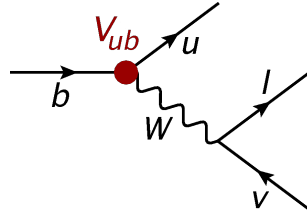
$Z$  and the top quark at the EW scale, and then connecting these operators with the low energy QCD interactions responsible for hadronic interactions. The latter are comprised in matrix elements of the effective operators, involving the initial and final state mesons of the decay in question. These matrix elements, being governed by non-perturbative interactions, cannot be calculated analytically, but have to be evaluated using non-perturbative methods like lattice QCD or QCD sum rules (see e. g. [22] and [23] for pedagogical introductions), unless it is possible to extract them from the data.

To summarize, in order to arrive at a theoretical description of flavour violating meson decays, the following five steps have to be taken:

1. Calculation of the weak interaction process governing the underlying flavour violating quark decay.
2. Construction of the low energy effective Hamiltonian by integrating out the heavy degrees of freedom (i. e.  $W^\pm, Z, t$ ).
3. Renormalization group running from the scale  $\mu \sim M_W$  of weak interactions to the hadronic scale  $\mu \sim \text{GeV}$ .
4. Collection of non-perturbative effects in QCD matrix elements involving initial and final state mesons.
5. Evaluation of matrix elements using non-perturbative methods (lattice QCD, QCD sum rules etc.) or extraction from data.

To better understand how this is done in practice, let us have a look at two simple examples. First we consider the semileptonic charged current decay  $B \rightarrow \pi \ell \nu$  from which the CKM element  $|V_{ub}|$  can be obtained. Then we sketch the SM prediction for neutral  $B_s$  meson mixing.

### Semileptonic charged currents: $B \rightarrow \pi \ell \nu$



**Fig. 3:** Tree level diagram mediating the  $b \rightarrow u \ell \nu$  decay.

The element  $|V_{ub}|$  can be determined from the semileptonic charged current transition  $b \rightarrow u \ell \nu$ , occurring at the tree level in the SM. Evaluating the relevant Feynman diagram shown in figure 3, we find

$$\mathcal{M} = \frac{g}{\sqrt{2}} V_{ub} (\bar{u} \gamma^\mu P_L b) \frac{g_{\mu\nu}}{p^2 - M_W^2} \frac{g}{\sqrt{2}} (\bar{\ell} \gamma^\nu P_L \nu). \quad (33)$$

Neglecting the momentum transfer  $p^2 \ll M_W^2$  in the  $W$  propagator, and using the short hand notation  $2(\bar{f} \gamma^\mu P_L f') = (\bar{f} f')_{V-A}$ ,  $\mathcal{M}$  simplifies to

$$\mathcal{M} = -\frac{g^2}{2M_W^2} V_{ub} (\bar{u} b)_{V-A} (\bar{\ell} \nu)_{V-A}. \quad (34)$$

Introducing the well-known Fermi constant

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2}, \quad (35)$$

we obtain the tree level effective Hamiltonian

$$\mathcal{H}_{\text{eff}} = \frac{4G_F}{\sqrt{2}} V_{ub} (\bar{u} b)_{V-A} (\bar{\ell} \nu)_{V-A} + h.c.. \quad (36)$$

In order to obtain an accurate expression for the  $B \rightarrow \pi \ell \nu$  decay, the known QCD corrections have to be taken into account, including the renormalization group running from the weak scale, where the  $W$  boson is integrated out, to the  $B$  meson scale. The non-perturbative QCD effects describing the  $B \rightarrow \pi$  transition generated by the effective Hamiltonian  $\mathcal{H}_{\text{eff}}$  are collected in the matrix element

$$\langle \pi | (\bar{b}u)_{V-A} | B \rangle, \quad (37)$$

which has been calculated by lattice QCD.

From the measurement of the  $B^0 \rightarrow \pi^- \ell^+ \nu$  branching ratio, one can then extract the value [24] (see also [25, 26])

$$|V_{ub}| = (3.72 \pm 0.16) \cdot 10^{-3}. \quad (38)$$

Note again that as the  $B^0 \rightarrow \pi^- \ell^+ \nu$  is generated at the tree level in the SM, NP contributions are expected to be negligible. Therefore the determination of  $|V_{ub}|$  in (38) holds model-independently. In a similar way, also  $|V_{us}|$ ,  $|V_{cb}|$  and the UT angle  $\gamma$  can be determined from tree level decays, independently of NP in flavour violating decays.

### $B_s - \bar{B}_s$ mixing

As discussed in section 1.4, neutral  $B_s$  meson mixing in the SM is driven by the one loop box diagram in figure 1. We have seen that due to the GIM mechanism and the hierarchical structure of the CKM matrix, only the mass-dependent part of the top quark contribution is relevant. Including a factor  $\eta_B = 0.55 \pm 0.01$  [27, 28] that comprises perturbative QCD corrections and the renormalization group evolution down to the  $B$  meson scale, the effective Hamiltonian can be written as

$$\mathcal{H}_{\text{eff}} = \frac{G_F^2}{16\pi^2} M_W^2 \eta_B (V_{tb}^* V_{ts})^2 S_0(x_t) (\bar{b}s)_{V-A} (\bar{b}s)_{V-A} + h.c., \quad (39)$$

with the one loop function  $S_0(x_t)$  given by

$$S_0(x_t) = \frac{4x_t - 11x_t^2 + x_t^3}{4(1-x_t)^2} - \frac{3x_t^3 \ln x_t}{2(1-x_t)^3}. \quad (40)$$

Sandwiching  $\mathcal{H}_{\text{eff}}$  between the initial and final state mesons, we obtain the mixing matrix element

$$M_{12} = \frac{1}{2m_{B_s}} \langle \bar{B}_s | \mathcal{H}_{\text{eff}} | B_s \rangle^*. \quad (41)$$

Again, the hadronic matrix element  $\langle \bar{B}_s | (\bar{b}s)_{V-A} (\bar{b}s)_{V-A} | B_s \rangle$  comprises the low-energy QCD dynamics of the mesonic process. It has been calculated by lattice QCD with an impressive precision.

Using the CKM matrix determined from tree level decays, the SM predictions for  $B_s - \bar{B}_s$  mixing observables, like the mass difference  $\Delta M_s$  and the  $CP$ -violating phase  $\phi_s$  can then be compared with the data. This comparison yields a good agreement of the measured values with their SM predictions, leaving only little room for NP contributions. In lecture 3 we will have a closer look at the constraints on NP models from flavour violating observables.

## 2 Phenomenology of $K$ and $B$ Meson Decays

### 2.1 Theory of Neutral Kaon Mixing

In the first lecture, we have briefly sketched the theoretical description of  $B_s - \bar{B}_s$  mixing in the SM, paying particular attention to specific details like the GIM mechanism and the construction of the effective Hamiltonian. In this section, we aim at a more thorough derivation of neutral meson mixing. While we focus on the case of kaon mixing in what follows, a generalization to neutral  $B$  and  $D$  mesons is straightforward.

Two neutral pseudoscalar  $K$  mesons exist, the  $|K^0\rangle = |d\bar{s}\rangle$  and the  $|\bar{K}^0\rangle = |s\bar{d}\rangle$ . They are each other's antiparticles and transform under  $CP$  as

$$CP|K^0\rangle = -|\bar{K}^0\rangle, \quad CP|\bar{K}^0\rangle = -|K^0\rangle. \quad (42)$$

As we can deduce from section 1.4,  $K^0$  and  $\bar{K}^0$  can mix by one loop box diagrams in the SM. The time evolution of the  $K^0 - \bar{K}^0$  system is therefore described by the two-component Schrödinger equation

$$i\frac{d\psi(t)}{dt} = \hat{H}\psi(t), \quad \psi(t) = \begin{pmatrix} K^0(t) \\ \bar{K}^0(t) \end{pmatrix} \quad (43)$$

with the  $2 \times 2$  Hamiltonian

$$\hat{H} = \hat{M} - i\frac{\hat{\Gamma}}{2} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} - \frac{i}{2} \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix}. \quad (44)$$

Here,  $\hat{M}$  is the dispersive part of the Hamiltonian, and  $\hat{\Gamma}$  is the absorptive part. Since both  $\hat{M}$  and  $\hat{\Gamma}$  are hermitian, we have

$$M_{21} = M_{12}^*, \quad \Gamma_{21} = \Gamma_{12}^*. \quad (45)$$

In addition,  $CPT$  invariance implies

$$M_{11} = M_{22} \equiv M, \quad \Gamma_{11} = \Gamma_{22} \equiv \Gamma, \quad (46)$$

so that the Hamiltonian simplifies to

$$\hat{H} = \begin{pmatrix} M - \frac{i}{2}\Gamma & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M - \frac{i}{2}\Gamma \end{pmatrix} \quad (47)$$

Diagonalizing  $\hat{H}$ , we obtain the mass eigenstates

$$K_{L,S} = \frac{(1 + \bar{\varepsilon})K^0 \pm (1 - \bar{\varepsilon})\bar{K}^0}{\sqrt{2(1 + |\bar{\varepsilon}|^2)}}, \quad (48)$$

where  $L$  and  $S$  stand for 'long' and 'short', respectively, and refer to the different lifetimes of the two states. The parameter  $\bar{\varepsilon}$  is defined through

$$\frac{1 - \bar{\varepsilon}}{1 + \bar{\varepsilon}} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}} = \frac{\Delta M - \frac{i}{2}\Delta\Gamma}{2M_{12} - i\Gamma_{12}}. \quad (49)$$

Experimentally  $\bar{\varepsilon}$  is found to be of the order  $\mathcal{O}(10^{-3})$ . Therefore the mass difference  $\Delta M$  and the width difference  $\Delta\Gamma$  are well approximated by the simple expressions

$$\Delta M = M_L - M_S = 2 \operatorname{Re}M_{12}, \quad \Delta\Gamma = \Gamma_L - \Gamma_S = 2 \operatorname{Re}\Gamma_{12}. \quad (50)$$

## 2.2 $CP$ Violation in the Neutral Kaon Sector

As we have seen before,  $K^0$  and  $\bar{K}^0$  transform into each other under  $CP$  conjugation. The  $CP$  eigenstates are therefore given by

$$K_1 = \frac{1}{\sqrt{2}}(K^0 - \bar{K}^0), \quad K_2 = \frac{1}{\sqrt{2}}(K^0 + \bar{K}^0). \quad (51)$$



$K_1$  is even under  $CP$  conjugation, while  $K_2$  is  $CP$ -odd. Comparing the  $CP$  eigenstates in (51) to the mass eigenstates in (48), it becomes clear that the mass eigenstates  $K_{L,S}$  are not pure  $CP$  eigenstates. Instead they contain a small admixture of the state with opposite  $CP$  parity:

$$K_S = \frac{K_1 + \bar{\epsilon}K_2}{\sqrt{1 + |\bar{\epsilon}|^2}}, \quad K_L = \frac{K_2 + \bar{\epsilon}K_1}{\sqrt{1 + |\bar{\epsilon}|^2}}. \quad (52)$$

This mixture of states with opposite  $CP$  parities implies that the  $CP$  symmetry is violated in neutral kaon mixing.

The two mass eigenstates,  $K_L$  and  $K_S$  are found experimentally to have very different lifetimes [13]:

$$\tau(K_S) \sim 90 \text{ ps}, \quad \tau(K_L) \sim 5 \cdot 10^3 \text{ ps}. \quad (53)$$

The explanation can be found in the  $CP$  properties of the two states:  $K_S$  is basically the  $CP$ -even state  $K_1$ , with a small  $CP$ -odd admixture.  $K_L$ , on the other hand, is approximately the  $CP$ -odd state  $K_2$ , with a small  $CP$ -even admixture. Now if  $CP$  is conserved in the decay of neutral kaons, then the  $CP$ -even state  $K_1$  will decay into two pions, forming a  $CP$ -even final state.  $K_2$ , however, has to decay into a  $CP$ -odd final state, which contains three pions. As the three pion final state is phase space suppressed with respect to the two pion final state, the  $K_1$  decay rate is much faster, so that its lifetime is much shorter than the one of  $K_2$ . The observed lifetime difference suggests that indeed  $CP$  is, at least approximately, conserved and  $\bar{\epsilon}$  is small.

In 1964, the decay  $K_L \rightarrow \pi^+\pi^-$  has been observed [29], yielding the first experimental confirmation that  $CP$  symmetry is violated. In 1980, the Nobel Prize has been awarded to Cronin and Fitch for this discovery.

The mere discovery of the  $K_L \rightarrow \pi^+\pi^-$  decay however does not tell us where the observed  $CP$  violation originates from.  $CP$  can either be violated in the neutral kaon mixing, if the mass eigenstates  $K_{L,S}$  are not  $CP$  eigenstates. Or  $CP$  can be violated in the decay – in that case the  $CP$ -odd state  $K_2$  can decay into a  $CP$ -even two pion final state. Last but not least, in general,  $CP$  can also be violated in the interference of mixing and decay amplitudes.

How can we distinguish whether  $CP$  is violated in the mixing (also called indirect  $CP$  violation) or in the decay process (direct  $CP$  violation)? The key idea is that the amount of direct  $CP$  violation depends on the decay channel, but indirect  $CP$  violation does not.

Hence, to disentangle the two types of  $CP$  violation, we study the following set of  $K \rightarrow \pi\pi$  decay modes:

$$K_L \rightarrow \pi^0\pi^0, \quad K_L \rightarrow \pi^+\pi^-, \quad (54)$$

$$K_S \rightarrow \pi^0\pi^0, \quad K_S \rightarrow \pi^+\pi^-. \quad (55)$$

As the charged and neutral pions form an isospin triplet, the two pion final state can have either isospin  $I = 0$  or  $I = 2$ . The decay amplitudes into charged and neutral pions can therefore be written as

$$A(K^0 \rightarrow \pi^+\pi^-) = \sqrt{\frac{2}{3}}A_0e^{i\delta_0} + \sqrt{\frac{1}{3}}A_2e^{i\delta_2}, \quad (56)$$

$$A(K^0 \rightarrow \pi^0\pi^0) = \sqrt{\frac{2}{3}}A_0e^{i\delta_0} - 2\sqrt{\frac{1}{3}}A_2e^{i\delta_2}. \quad (57)$$

Here,  $A_0$  and  $A_2$  parametrize the decay amplitudes in the  $I = 0$  and  $I = 2$  final states, respectively. The ‘strong phases’  $\delta_0$  and  $\delta_2$  do not change sign under  $CP$  conjugation.

Defining then the ratios

$$\eta_{00} = \frac{A(K_L \rightarrow \pi^0\pi^0)}{A(K_S \rightarrow \pi^0\pi^0)}, \quad \eta_{+-} = \frac{A(K_L \rightarrow \pi^+\pi^-)}{A(K_S \rightarrow \pi^+\pi^-)}, \quad (58)$$

it is possible to disentangle  $CP$  violation in neutral kaon mixing, parametrized by  $\varepsilon$ , from direct  $CP$  violation in the  $K \rightarrow \pi\pi$  decays, parametrized by  $\varepsilon'$ . One can show that

$$\varepsilon \simeq \frac{1}{3}(\eta_{00} + 2\eta_{+-}), \quad \text{Re}\left(\frac{\varepsilon'}{\varepsilon}\right) \simeq \frac{1}{6}\left(1 - \left|\frac{\eta_{+-}}{\eta_{00}}\right|^2\right). \quad (59)$$

### 2.3 Status of $\varepsilon$ and $\varepsilon'$

Both  $|\varepsilon|$  and  $\text{Re}(\varepsilon'/\varepsilon)$  have been measured with high precision, with the results [13]

$$|\varepsilon| = (2.228 \pm 0.011) \cdot 10^{-3}, \quad (60)$$

$$\text{Re}(\varepsilon'/\varepsilon) = (16.6 \pm 2.3) \cdot 10^{-4}. \quad (61)$$

In the SM,  $CP$  violation in the kaon sector is strongly suppressed. As the presence of three quark generations is needed for  $CP$  violation, both  $\varepsilon$  and  $\varepsilon'$  are generated by top quark contributions. The effect is therefore proportional to the combination of CKM elements

$$\text{Im}(V_{ts}^* V_{td}) \simeq \mathcal{O}(10^{-4}). \quad (62)$$

$CP$  violation in the kaon sector is thus strongly suppressed in the SM. In the presence of NP, however, this strong CKM suppression can be absent, depending on the flavour structure of the model.  $CP$ -violating observables in the kaon sector therefore have an outstanding sensitivity to NP contributions.

For the parameter  $\varepsilon$  a simple yet precise formula can be derived:

$$\varepsilon = \frac{\kappa_\varepsilon e^{i\varphi_\varepsilon}}{\sqrt{2}\Delta M_K} \text{Im}M_{12}. \quad (63)$$

Here, the mass splitting  $\Delta M_K = (0.5292 \pm 0.0009) \cdot 10^{-2} \text{ ps}^{-1}$  and the phase  $\varphi_\varepsilon = 43.51^\circ$  have been measured precisely [13]. The parameter  $\kappa_\varepsilon$  comprises corrections from long-distance dynamics, and has been estimated to be  $\kappa_\varepsilon = 0.94 \pm 0.02$  [30, 31].

The off-diagonal element  $M_{12}$  of the mixing amplitude is, as discussed above, generated by box diagrams in the SM. While its real part receives sizeable long-distance contributions, the  $CP$ -violating imaginary part is driven by short distance dynamics and therefore under good theoretical control. Including the known higher order perturbative corrections and the non-perturbative parameter  $\hat{B}_K$  obtained from lattice QCD calculations, one finds [32]

$$|\varepsilon|_{\text{SM}} = (1.90 \pm 0.26) \cdot 10^{-3}, \quad (64)$$

which is a bit lower albeit still consistent with the data.

Due to its strong suppression in the SM,  $\varepsilon$  is very sensitive to potential NP contributions. The good agreement of the measured value with its SM prediction therefore results in strong constraints on the NP entering  $K^0 - \bar{K}^0$  mixing. We will return to this topic in more detail in the third lecture.

The ratio  $\varepsilon'/\varepsilon$  has recently received a lot of attention. While its measured value in (61) has been available since the late 1990s, until recently no reliable SM prediction was available. The situation changed when the first lattice QCD calculations of the relevant  $K \rightarrow \pi\pi$  hadronic matrix elements became available. The result reads [33]

$$B_6^{(1/2)} = 0.57 \pm 0.19, \quad B_8^{(3/2)} = 0.76 \pm 0.05. \quad (65)$$

While this result, in particular the one for  $B_6^{(1/2)}$ , still carries sizeable uncertainties, it is interesting to note its consistency with the bound

$$B_6^{(1/2)} < B_8^{(3/2)} < 1, \quad (66)$$

that has recently been derived using large  $N_c$  counting and the dual QCD approach [34, 35].

The result for the hadronic matrix elements in (65) can then be plugged into the simple phenomenological expression for  $\varepsilon'/\varepsilon$  in the SM [36, 37]:

$$\text{Re}(\varepsilon'/\varepsilon)_{\text{SM}} = \frac{\text{Im}(V_{ts}^* V_{td})}{1.4 \cdot 10^{-4}} \cdot 10^{-4} \cdot \left( -3.6 + 21.4 B_6^{(1/2)} + 1.2 - 10.4 B_8^{(3/2)} \right), \quad (67)$$

which has been derived using the calculation of perturbative QCD contributions at next-to-leading order (NLO).

The first two terms in the brackets of (67) stem from the  $I = 0$  amplitude which is dominantly generated by QCD penguin contributions. The last two terms, on the other hand, originate in the  $I = 2$  amplitude, caused mainly by EW penguin contributions. The numerical values of these contributions, together with the result for the hadronic matrix elements, leads to a large cancellation between the  $I = 0$  and  $I = 2$  contributions to  $\text{Re}(\varepsilon'/\varepsilon)_{\text{SM}}$ . Due to this cancellation, a precise knowledge of the hadronic matrix elements  $B_6^{(1/2)}$  and  $B_8^{(3/2)}$  is of utmost importance for an accurate prediction of  $\text{Re}(\varepsilon'/\varepsilon)$  in the SM. Using the result in (65), one finds [37]

$$\text{Re}(\varepsilon'/\varepsilon)_{\text{SM}} = (1.4 \pm 4.6) \cdot 10^{-4}. \quad (68)$$

This prediction is significantly lower than the measured value in (61), revealing a  $2.9\sigma$  tension. A consistent result has been obtained in [38]. It is interesting to note that the central value in (68) is much lower than the long-standing result in [39], although consistent due to the large uncertainties in the latter analysis. We are looking forward to future improved lattice QCD calculations by several groups which will clarify the present situation and hopefully strengthen the indicated hint for NP.

Having at hand only a single lattice result for the hadronic matrix elements in question, it would clearly be premature to claim the presence of NP in  $\varepsilon'/\varepsilon$ . The observed tension, however, is intriguing. Due to its strong suppression in the SM, the ratio  $\varepsilon'/\varepsilon$  is extremely sensitive to NP contributions. It is therefore conceivable that NP would first be observed in this observable, even if other flavour data measured so far show little or no discrepancy with their SM predictions.

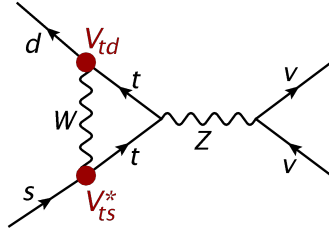
Following the recent progress on the theoretical understanding of  $\varepsilon'/\varepsilon$ , this observable has been revisited in the context of various NP models [40–47]. It turns out that several extensions of the SM can significantly enhance  $\varepsilon'/\varepsilon$  and thereby reconcile the theory prediction with the data. In addition, many NP scenarios predict simultaneous large deviations from the SM prediction of the rare decay  $K_L \rightarrow \pi^0 \nu \bar{\nu}$ , with the sign of the latter effect depending on the structure of the model.

## 2.4 Rare $K$ Decays

Rare and  $CP$ -violating kaon decays, like the aforementioned decay  $K_L \rightarrow \pi^0 \nu \bar{\nu}$ , offer a unique opportunity to look for NP. These decays, mediated by  $s \rightarrow d$  FCNC transitions at the quark level, are strongly suppressed in the SM by the hierarchical structure of the CKM matrix and the GIM mechanism. Consequently, large NP effects are possible even if the NP mass scale is much beyond the TeV scale. Of particular interest are the decay modes  $K^+ \rightarrow \pi^+ \nu \bar{\nu}$  and  $K_L \rightarrow \pi^0 \nu \bar{\nu}$ , as they are not only strongly suppressed in the SM but also theoretically extremely clean. Therefore, an outstanding NP sensitivity is provided by these decays. In fact, it has been shown within a simplified model analysis that a flavour violating  $Z'$  gauge boson can lead to large effects in the  $K \rightarrow \pi \nu \bar{\nu}$  decays even if its mass is in the  $10^3$  TeV range [48].

In the SM, the  $K \rightarrow \pi \nu \bar{\nu}$  decays are governed by  $Z$ -penguin and box diagrams like the ones shown in figure 4. The effective Hamiltonian reads [49]

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}} \frac{\alpha}{2\pi \sin^2 \theta_W} \sum_{\ell=e,\mu,\tau} \left( V_{cs}^* V_{cd} X_{\text{NNL}}^\ell(x_c) + V_{ts}^* V_{td} X(x_t) \right) (\bar{s}d)_{V-A} (\bar{\nu}_\ell \nu_\ell)_{V-A}. \quad (69)$$



**Fig. 4:** Z-penguin diagram contributing to  $K \rightarrow \pi \nu \bar{\nu}$  in the SM.

The first term in the brackets corresponds to the charm quark contribution which is known to NNLO in QCD [50, 51] and NLO in the EW theory [52]. It is relevant only for the  $CP$ -conserving decay  $K^+ \rightarrow \pi^+ \nu \bar{\nu}$ . The second term stems from the top quark contribution which affects both the  $CP$ -conserving mode  $K^+ \rightarrow \pi^+ \nu \bar{\nu}$  and the  $CP$ -violating mode  $K_L \rightarrow \pi^0 \nu \bar{\nu}$ .

The relevant  $\langle \pi | (\bar{s}d)_{V-A} | K \rangle$  matrix elements can be extracted from the data on  $K^+ \rightarrow \pi^0 e^+ \nu$  with high precision, making use of isospin symmetry. The main uncertainty is the SM prediction for the  $K \rightarrow \pi \nu \bar{\nu}$  branching ratios therefore stem from the determination of the relevant CKM elements. Of particular importance is the value of  $|V_{cb}|$  and, for  $K_L \rightarrow \pi^0 \nu \bar{\nu}$ , the UT angle  $\gamma$ . The current SM prediction for the two branching ratios is [53]

$$\mathcal{B}(K^+ \rightarrow \pi^+ \nu \bar{\nu}) = (8.4 \pm 1.0) \cdot 10^{-11}, \quad (70)$$

$$\mathcal{B}(K_L \rightarrow \pi^0 \nu \bar{\nu}) = (3.4 \pm 0.6) \cdot 10^{-11}. \quad (71)$$

A big experimental effort to measure these decays is currently underway at the NA62 experiment [54] at CERN and the KOTO experiment [55] at J-PARC in Japan.

NP contributions to  $\mathcal{H}_{\text{eff}}$  in (69) can be parametrized model-independently by replacing the SM top-loop function  $X(x_t)$  by a general complex function [56]

$$X \equiv |X| e^{i\theta_X}. \quad (72)$$

NP in the  $K \rightarrow \pi \nu \bar{\nu}$  system can therefore be described by two independent parameters  $|X|$  and  $\theta_X$ . Measuring both  $\mathcal{B}(K^+ \rightarrow \pi^+ \nu \bar{\nu})$  and  $\mathcal{B}(K_L \rightarrow \pi^0 \nu \bar{\nu})$  determines both parameters, and observing

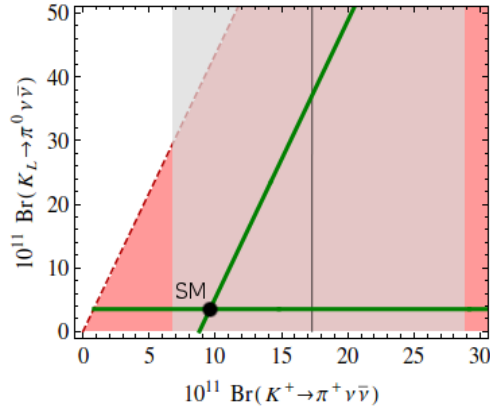
$$|X| \neq X(x_t) \quad \text{and/or} \quad \theta_X \neq 0 \quad (73)$$

would be an unambiguous sign of NP.

Determining both  $|X|$  and  $\theta_X$  not only provides a clean test of the SM, but in case of a non-vanishing NP contribution also allows to draw conclusions about the structure of NP contributions to neutral kaon mixing [57]. The reason is quite simple to understand. If the effective flavour changing  $s \rightarrow d$  transition is, as in the SM, purely left-handed, then the same NP structure is responsible for  $K^0 - \bar{K}^0$  mixing and for the  $K \rightarrow \pi \nu \bar{\nu}$  decays. In particular, the same  $CP$ -violating phase  $\theta_K$  enters, only multiplied by a factor of two for  $K^0 - \bar{K}^0$  mixing. The constraint on NP from  $\Delta M_K$  is much weaker than the one from  $\varepsilon$ , so that any NP contribution to  $K^0 - \bar{K}^0$  mixing must be predominantly real:  $2\theta_K \simeq 0, \pi$ . The phase  $\theta_K$  measured in the  $K \rightarrow \pi \nu \bar{\nu}$  decays is then restricted to the values

$$\theta_K \simeq 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}. \quad (74)$$

In the plane showing the branching ratios  $\mathcal{B}(K^+ \rightarrow \pi^+ \nu \bar{\nu})$  and  $\mathcal{B}(K_L \rightarrow \pi^0 \nu \bar{\nu})$ , these values for  $\theta_K$  correspond to two straight lines: a horizontal one, where  $\mathcal{B}(K_L \rightarrow \pi^0 \nu \bar{\nu})$  remains SM-like as the NP contribution is  $CP$ -conserving, and a slanted one which is parallel to the model-independent Grossman-Nir bound [58]. This pattern is depicted by the green lines in figure 5. If on the other hand, both left-



**Fig. 5:** Model-distinguishing correlation between  $\mathcal{B}(K^+ \rightarrow \pi^+ \nu \bar{\nu})$  and  $\mathcal{B}(K_L \rightarrow \pi^0 \nu \bar{\nu})$  [57].

and right-handed FCNCs are induced by NP, then neutral kaon mixing will usually be dominated by the so-called “left-right” effective operators containing both chiralities. In that case the correlation with the  $K \rightarrow \pi \nu \bar{\nu}$  decays is lost, and no correlation between  $\mathcal{B}(K^+ \rightarrow \pi^+ \nu \bar{\nu})$  and  $\mathcal{B}(K_L \rightarrow \pi^0 \nu \bar{\nu})$  arises. The full range for the two branching ratios, shown in red in figure 5, is then possible.

The discussed correlation between  $\mathcal{B}(K^+ \rightarrow \pi^+ \nu \bar{\nu})$  and  $\mathcal{B}(K_L \rightarrow \pi^0 \nu \bar{\nu})$  has indeed been found in a number of NP models with purely left-handed FCNC transitions [40, 59–62].

## 2.5 Quick Summary of Kaon Physics

Before moving on to  $B$  physics and in particular to a recent set of anomalies, let us recapitulate the unique role of kaon physics. Kaon decays have played an important role in constructing the SM, and they offer unique opportunities to test its extensions.

**The past.** In order to account for the smallness of the  $K_L \rightarrow \mu^+ \mu^-$  branching ratio, a fourth quark, the charm quark, has been predicted prior to its direct discovery. Also  $CP$  violation has first been observed in the kaon system, by measuring a non-zero  $K_L \rightarrow \pi \pi$  decay width. The necessity for a third quark generation had thus been established.

**The present.** Currently, the  $CP$ -violating parameter  $\epsilon$  places one of the most stringent constraints on physics beyond the SM, in particular if a non-trivial flavour structure is involved. In addition, the recent lattice calculations of the hadronic matrix elements entering  $\epsilon'/\epsilon$  seem to hint for a tension between the SM prediction and the data.

**The future.** If future more precise predictions of  $\epsilon'/\epsilon$  confirm this tension, the road will be paved towards spectacular NP discoveries in rare  $K$  decays. A special role is played by the  $K \rightarrow \pi \nu \bar{\nu}$  decays, as thanks to their theoretical cleanliness they offer an extremely sensitive probe of NP.

## 2.6 $B$ Physics

Historically, kaon physics has been the main player in the field of flavour physics. More recently however,  $B$  meson physics has gained significant importance. After the first observation of  $B_d - \bar{B}_d$  oscillations, in the 1990s the two  $B$ -factories BaBar and Belle were built to precisely measure the properties of  $B$  mesons and their decays. The  $B$ -factories delivered a large number of highly relevant results, including the discovery and precise measurement of  $CP$  violation in  $B_d - \bar{B}_d$  oscillations, and measurements of semileptonic  $B$  decays relevant for the extraction of the CKM elements  $|V_{ub}|$  and  $|V_{cb}|$ . Further significant improvements on the physics results of the  $B$ -factories, as well as measurements of a number of so far undetected rare  $B$  meson decays, can be expected from the second generation  $B$ -factory Belle II. First Belle II physics results should become available within a couple of years from now.

A blind spot in the programme of the  $B$ -factories, however, is the physics of  $B_s$  mesons. Due to their larger mass, they are not produced in the decay of the  $\Upsilon(4S)$  resonance, which the  $B$ -factories rely on. Consequently, hadron colliders like the Tevatron and the LHC have an advantage here. Indeed,  $B_s - \bar{B}_s$  oscillations have first been observed by the CDF experiment at Fermilab in 2006 [63], and later confirmed by LHCb [64]. The latter experiment also provides the most stringent constraint on  $CP$  violation in  $B_s - \bar{B}_s$  mixing [65].

LHCb and, to some extent, also CMS and ATLAS have also yielded important data on a number of rare  $B$  and  $B_s$  meson decays, like  $B \rightarrow K^{(*)}\mu^+\mu^-$  and  $B_s \rightarrow \mu^+\mu^-$ . For the latter, a combination of LHCb and CMS data lead to its discovery, with a branching ratio measurement in decent agreement with the SM prediction. The data on the  $B \rightarrow K^{(*)}\mu^+\mu^-$  decays, on the other hand, leaves us with some intriguing anomalies to be discussed in what follows.

## 2.7 Recent Anomalies in $b \rightarrow s$ Transitions

The aforementioned decays  $B \rightarrow K\mu^+\mu^-$ ,  $B \rightarrow K^*\mu^+\mu^-$ , and  $B_s \rightarrow \mu^+\mu^-$  are all governed by the quark level transition  $\bar{b} \rightarrow \bar{s}\mu^+\mu^-$ . In order to understand the physics behind the observed anomalies, we start by writing down the relevant effective Hamiltonian [49]

$$\mathcal{H}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}}V_{tb}V_{ts}^*\frac{e^2}{16\pi^2}\sum_i(C_i\mathcal{O}_i + C'_i\mathcal{O}'_i) + h.c.. \quad (75)$$

In the SM, only the unprimed Wilson coefficients  $C_i$ , corresponding to left-handed FCNC transitions, are relevant, due to the left-handedness of the flavour violating weak interactions. NP contributions, on the other hand, can have either chirality. The operators most sensitive to NP are the dipole operators

$$\mathcal{O}_7^{(\prime)} = \frac{m_b}{e}(\bar{s}\sigma_{\mu\nu}P_{R(L)}b)F^{\mu\nu} \quad (76)$$

and the four fermion operators

$$\mathcal{O}_9^{(\prime)} = (\bar{s}\gamma_\mu P_{L(R)}b)(\bar{\mu}\gamma^\mu\mu), \quad (77)$$

$$\mathcal{O}_{10}^{(\prime)} = (\bar{s}\gamma_\mu P_{L(R)}b)(\bar{\mu}\gamma^\mu\gamma_5\mu), \quad (78)$$

that are not affected by tree-level contributions in the SM.

The dipole operators  $\mathcal{O}_7^{(\prime)}$  are constrained by the well-measured  $B \rightarrow X_s\gamma$  transition, whose branching ratio is in good agreement with the SM prediction. The four-fermion operators  $\mathcal{O}_{10}^{(\prime)}$  mediate the decay  $B_s \rightarrow \mu^+\mu^-$ . While the data do not show a significant deviation from the SM prediction in this case, the experimental uncertainties are still sizeable, allowing for a relevant NP contribution to their Wilson coefficients  $C_{10}^{(\prime)}$ . The scalar and pseudoscalar four-fermion operators, on the other hand, are strongly constrained by the  $B_s \rightarrow \mu^+\mu^-$  branching ratio measurement. We therefore neglect them in this discussion.

The observation that different decays and observables are sensitive to different Wilson coefficients in the effective Hamiltonian (75) is crucial for the theoretical interpretation of the data. Of particular interest is the decay  $B \rightarrow K^*\mu^+\mu^-$ , where  $K^*$  further decays into a kaon and a pion. The four-body final state can be described in terms of three angles and the invariant mass square of the muon pair,  $q^2 = (p_{\mu^+} + p_{\mu^-})^2$ . The differential decay rates can then be decomposed into a sum of contributions with specific angular dependence. We note that different parametrizations have been proposed in the literature, with the goal to minimize the theoretical uncertainties in the observables in question [66–68].

In one parametrization, a set of ‘optimized’ observables has been derived with the goal to cancel the  $B \rightarrow K^*$  form factor dependence at leading order [68]. One of these observables which has attracted a lot of attention over the past few years is  $P'_5$ . A few years ago, the LHCb collaboration reported an

anomaly in the low  $q^2$  region in this observable, which is by now established at the  $3.4\sigma$  level [69]. Also more recent data from ATLAS [70] and Belle [71] hint in the same direction, albeit with much smaller significance. The recent measurement of  $P'_5$  by CMS [72], on the other hand, is consistent with the SM.

While the physical meaning of  $P'_5$  can be understood in terms of the transversity amplitudes depending on the spin of the muon pair, its interpretation is not very intuitive and we do not go into the details here. Further information can for example be found in [73]. In what follows we focus instead on possible interpretations of the observed anomaly.

Global fits of the Wilson coefficients in the effective Hamiltonian (75) reveal that a sizeable non-standard contribution to  $C_9$  is required to solve the  $P'_5$  anomaly [74–76]. Interestingly, at the same time also other, smaller tensions in the data are softened, such as the  $B^+ \rightarrow K^+\mu^+\mu^-$  and  $B_s \rightarrow \phi\mu^+\mu^-$  branching ratios. Further, if the NP is assumed to contribute only to the muon channel, i. e. to violate lepton flavour universality, also the  $R_K$  anomaly can be explained. Here,  $R_K$  is defined as the ratio of  $B^+ \rightarrow K^+\mu^+\mu^-$  and  $B^+ \rightarrow K^+e^+e^-$  branching ratios,

$$R_K = \frac{\mathcal{B}(B^+ \rightarrow K^+\mu^+\mu^-)}{\mathcal{B}(B^+ \rightarrow K^+e^+e^-)}. \quad (79)$$

The LHCb measurement of  $R_K$  [77] in the low  $q^2$  region is 2.6 standard deviations below the very accurate SM prediction  $R_K \simeq 1$  [78]. Similar hints for a violation of lepton flavour universality have recently also been found in the  $B \rightarrow K^*\ell^+\ell^-$  ( $\ell = \mu, e$ ) decays [71, 79].

The question is now how to interpret this result theoretically. Given the loop suppression of FCNCs in the SM, it is conceivable that the shift in  $C_9$  is induced by NP. Popular and well-motivated NP models, such as supersymmetric theories or models with partial compositeness, can, however, not account for this deviation [80]. It is however possible to induce a large contribution to  $C_9$  in phenomenologically viable NP models: two known examples are models with a flavour violating  $Z'$  gauge boson [81–85], and leptoquark scenarios [86–88]. Interestingly the latter can also address the tension in  $B \rightarrow D^{(*)}\tau\nu$  data.

Before claiming the presence of NP in  $b \rightarrow s\mu^+\mu^-$  transitions, it is however necessary to investigate the SM prediction for potentially underestimated theoretical uncertainties, see [76, 89, 90] for recent discussions. The main theoretical uncertainties lie, on the one hand, in the  $B \rightarrow K^*$  form factors that describe the hadronic physics of the  $B \rightarrow K^*$  transition in the factorization limit. On the other hand, sizeable uncertainties stem from non-factorizable corrections that arise at  $\mathcal{O}(\Lambda_{\text{QCD}}/m_b)$ .

The hadronic form factors can be computed at large  $q^2$  by lattice QCD, and at low  $q^2$  by light-cone sum rule techniques. Their extrapolation yields consistent results, so that the form factors are unlikely the source of the observed anomaly. Systematic improvements of the form factor calculations can be expected over the coming years, further reducing the associated uncertainties.

The non-factorizable corrections, however, are difficult to assess theoretically, and the associated uncertainties can only be estimated. The dominant contributions arise from long-distance charm loops coupling to photons and, in turn, the final state muon pair. In the effective Hamiltonian description of (75), these contributions would mimic a NP contribution to the operator  $\mathcal{O}_9$ , due to the vector coupling of the photon.

There are however two crucial differences that can distinguish non-factorizable corrections from NP contributions to  $C_9$ . First, as the photon couples universally to all lepton flavours, a lepton flavour non-universal signal would be a clear sign of NP. If the violation of lepton flavour universality is confirmed by future data, and analogous ratios in other channels show the same pattern, then we would have an unambiguous sign of NP in semileptonic  $b \rightarrow s$  transitions. Second, NP contributions are in general independent of the dimuon momentum  $q^2$ , while non-factorizable charm loop contributions are expected to be enhanced near the  $c\bar{c}$  threshold. While the current data are consistent with a  $q^2$ -independent  $C_9$ , future more accurate measurements could reveal a  $q^2$ -dependence of the required new contribution, hence clearly disfavouring the NP interpretation.

We have thus seen that even though the understanding of the observed  $P'_5$  anomaly is currently limited by theoretical uncertainties, future more accurate experimental data will provide a significant contribution to its resolution.

### 3 Flavour Physics beyond the Standard Model

#### 3.1 The SM Flavour Problem

As we have seen in lecture 1, flavour violation in the SM is generated by the Yukawa couplings, generating the fermion masses and flavour mixings. Most of the free parameters of the SM are related to the flavour sector, calling for a more fundamental theory that explains their origin. Moreover, the SM flavour sector is found experimentally to obey a very hierarchical pattern, with quark masses spanning five orders of magnitude, and a CKM mixing matrix close to the unit matrix. This structure seems to suggest the presence of an approximate flavour symmetry in the fundamental theory of flavour.

Experimentally, the SM quark flavour sector has been well tested by precise measurements of a large number of flavour violating  $K$ ,  $B$ , and  $D$  meson decays. Despite a few anomalies, overall the SM with its simple CKM picture of flavour violation has been extremely successful at describing the data. Consequently, as we will see in section 3.2, strong constraints on the scale of NP with generic flavour violating interactions can be derived. NP at the TeV scale must then have a very non-generic flavour structure, with an efficient suppression of FCNC transitions. The most widely known and employed example is the concept of Minimal Flavour Violation (MFV), which we discuss in section 3.3. However it is also possible to avoid dangerously large FCNCs without imposing MFV. Different flavour symmetries and symmetry breaking patterns can be employed. A complementary approach to flavour is provided by models with partially composite fermions, where the observed flavour structure has a dynamical origin. We will briefly review this idea in section 3.4.

#### 3.2 Constraints on the Scale of New Physics

In the SM, FCNC processes receive various strong suppression factors that make them highly sensitive to NP contributions. Firstly, as FCNC couplings are generated only at the loop level, they are suppressed by a loop factor  $g^2/(16\pi)^2$ , where  $g$  is the weak  $SU(2)_L$  coupling constant. The GIM mechanism further reduces the size of FCNC transitions, in particular in the kaon system. FCNC transitions in the  $K$ ,  $B_d$  and  $B_s$  systems, respectively, are then governed by the following CKM factors:

$$\underbrace{V_{ts}^* V_{td}}_{K \text{ system}} \sim 5 \cdot 10^{-4}, \quad \underbrace{V_{tb}^* V_{td}}_{B_d \text{ system}} \sim 10^{-2}, \quad \underbrace{V_{tb}^* V_{ts}}_{B_s \text{ system}} \sim 4 \cdot 10^{-2}. \quad (80)$$

We observe that the CKM hierarchy yields the strongest suppression in the kaon system, while  $b \rightarrow d$  and in particular  $b \rightarrow s$  transitions are much less rare. Lastly, due to the left-handedness of weak interactions, FCNC processes in the SM are purely left-handed. As we will see below, the purely left-handed effective operators are much less affected by renormalization group effects than the left-right ones that are generated in many NP models.

All of these suppression mechanisms can in principle be circumvented by NP, so that FCNC transitions provide an excellent sensitivity to NP even much beyond the TeV scale. To explore the NP reach of flavour physics in a model-independent way, it is useful to study it in terms of the effective field theory language. In this framework, the renormalizable SM Lagrangian is extended by including all higher-dimensional effective operators that are consistent with the gauge symmetries of the SM:

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{C_i}{\Lambda} \mathcal{O}_i^{\text{dim } 5} + \sum_i \frac{C_i}{\Lambda^2} \mathcal{O}_i^{\text{dim } 6} + \dots \quad (81)$$

The SM then constitutes the low energy limit valid at energy scales much below  $\Lambda$ , where the higher-dimensional operator contributions are irrelevant. The scale  $\Lambda$  is the cut-off scale of the effective theory.



It generally arises from integrating out new particles with masses  $M \sim \mathcal{O}(\Lambda)$ . Hence, at energy scales above  $\Lambda$  it has to be replaced by the full theory in which the new particles are physical degrees of freedom.

The only operator of dimension five in (81) is the Weinberg operator [91] that is relevant for the generation of neutrino masses. For details, see the lectures of Gabriela Barenboim [6]. The leading operators mediating FCNCs arise at dimension six and are therefore suppressed by two powers of the inverse of the cut-off scale  $\Lambda$ .

In a general NP scenario, FCNC transitions are not related to the CKM matrix and the respective CKM suppression factors in (80) are absent. We can instead parametrize the strength of flavour violating transitions by a parameter  $\delta$  that can in general depend on the meson system in question. Then the operators contributing to neutral meson mixings ( $\Delta F = 2$ ) are proportional to  $\delta^2$ , while the operators contributing to rare decays like  $K \rightarrow \pi\nu\bar{\nu}$  violate flavour only by one unit ( $\Delta F = 1$ ) and are therefore proportional to  $\delta$ . Hence if flavour violation is suppressed in the NP sector, i. e.  $\delta \ll 1$ , then rare decays are in general more sensitive to the contributions from NP than  $\Delta F = 2$  observables. If, on the other hand, FCNC effects are suppressed by a large NP scale,  $\Lambda \gg M_W$  and  $\delta \sim \mathcal{O}(1)$ , then  $\Delta F = 2$  observables are typically more sensitive to NP effects than  $\Delta F = 1$  ones. The sensitivity to large NP scales increases with increasing flavour violation  $\delta$  in the NP sector, and for sizeable values of  $\delta$  extends far beyond the reach of the LHC.

To investigate the NP reach of flavour physics, specifically of  $\Delta F = 2$  transitions, more explicitly, let us consider the general dimension six effective Hamiltonian:

$$\mathcal{H}_{\text{eff}}^{\Delta F=2} = \frac{1}{\Lambda^2} \left[ \sum_{i=1}^5 C_i \mathcal{O}_i + \sum_{i=1}^3 \tilde{C}_i \tilde{\mathcal{O}}_i \right]. \quad (82)$$

Here, the four fermion operators mediating  $B_{d,s} - \bar{B}_{d,s}$  mixing are defined as ( $q = d, s$ )

$$\mathcal{O}_1 = (\bar{q}^\alpha \gamma_\mu P_L b^\alpha) (\bar{q}^\beta \gamma^\mu P_L b^\beta), \quad (83)$$

$$\mathcal{O}_2 = (\bar{q}^\alpha P_L b^\alpha) (\bar{q}^\beta P_L b^\beta), \quad \mathcal{O}_3 = (\bar{q}^\alpha P_L b^\beta) (\bar{q}^\beta P_L b^\alpha), \quad (84)$$

$$\mathcal{O}_4 = (\bar{q}^\alpha P_L b^\alpha) (\bar{q}^\beta P_R b^\beta), \quad \mathcal{O}_5 = (\bar{q}^\alpha P_L b^\beta) (\bar{q}^\beta P_R b^\alpha), \quad (85)$$

$$\tilde{\mathcal{O}}_1 = (\bar{q}^\alpha \gamma_\mu P_R b^\alpha) (\bar{q}^\beta \gamma^\mu P_R b^\beta), \quad (86)$$

$$\tilde{\mathcal{O}}_2 = (\bar{q}^\alpha P_R b^\alpha) (\bar{q}^\beta P_R b^\beta), \quad \tilde{\mathcal{O}}_3 = (\bar{q}^\alpha P_R b^\beta) (\bar{q}^\beta P_R b^\alpha), \quad (87)$$

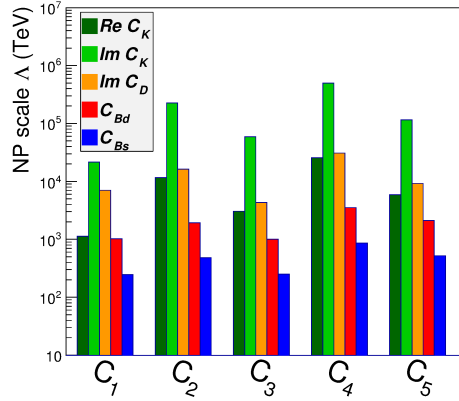
where summation over the colour indices  $\alpha, \beta$  is understood. Analogous expressions hold for the operators mediating  $K^0 - \bar{K}^0$  and  $D^0 - \bar{D}^0$  mixing. In the SM, only the operator  $\mathcal{O}_1$  is present.

Assuming a generic NP flavour structure, i. e.

$$|C_i| \sim \mathcal{O}(1), \quad \arg C_i \sim \mathcal{O}(1), \quad (88)$$

it is possible to derive lower bounds on the NP scale by comparing the data on meson mixing observables with their theory predictions [92]. The result of this exercise is shown in figure 6. The constraints from different meson systems are shown in different colours. We find that, as expected, the strongest constraints arise from neutral kaon mixing, and in particular from the  $CP$ -violating parameter  $\varepsilon$ . Most constraining are the operators  $\mathcal{O}_2$  and  $\mathcal{O}_4$ , which constrain the scale of generic NP to be above  $10^5$  TeV. Stringent bounds are also obtained from the non-observation of  $CP$ -violation in neutral  $D$  meson mixing. The  $B_d$  and  $B_s$  systems, although less constraining, still push the NP scale above 100 TeV.

We conclude that NP at or near the TeV scale, required for a natural origin of EW symmetry breaking, must have a very non-trivial flavour structure. Clearly, flavour can not be a conserved quantum number in the NP model, as the flavour symmetry is already broken in the SM. Yet it is possible to achieve an approximate conservation of flavour, as we will see in what follows.



**Fig. 6:** Model-independent constraints on the NP scale  $\Lambda$ , obtained from the various operators of the  $\Delta F = 2$  effective Hamiltonian. Figure taken from [92] with kind permission of the authors.

### 3.3 Minimal Flavour Violation

In lecture 1 we have seen that the SM quark sector possesses a global flavour symmetry

$$G_{\text{flavour}} = U(3)_Q \times U(3)_U \times U(3)_D \quad (89)$$

that is explicitly broken by the quark Yukawa couplings  $Y_U$  and  $Y_D$ . Due to the hierarchical structure of the Yukawa matrices, FCNC processes in the SM are strongly suppressed. The idea is then to extend this mechanism to the NP sector in order to suppress flavour violation also beyond the SM and reconcile TeV scale NP with the stringent experimental flavour constraints.

This leads us to formulate to the *Minimal Flavour Violation (MFV)* hypothesis [93–95]: In MFV, also in the NP sector the quark Yukawa couplings  $Y_U$  and  $Y_D$  constitute the only source of breaking of the flavour symmetry  $G_{\text{flavour}}$ . For phenomenological reasons, it is usually also assumed that no new  $CP$ -violating phases arise, so that the only source of  $CP$  violation remains the CKM phase. The MFV hypothesis ensures that all flavour and  $CP$ -violating NP effects are suppressed by the same CKM factors as the corresponding SM ones.

To impose the MFV hypothesis, it is useful to think of the Yukawa couplings as so-called spurion fields. That means, we can formally restore the flavour symmetry  $G_{\text{flavour}}$  by treating the Yukawa couplings  $Y_U, Y_D$  as dimensionless auxiliary (i. e. non-dynamical) fields. Assigning the  $G_{\text{flavour}}$  transformation properties

$$Y_U \sim (3, \bar{3}, 1), \quad Y_D \sim (3, 1, \bar{3}) \quad (90)$$

restores the flavour symmetry of the SM Lagrangian. In particular, the Yukawa coupling Lagrangian (7) becomes (formally) invariant under the flavour symmetry.

In order to extend this concept to the NP sector, we consider again the effective Lagrangian in (81). The MFV ansatz requires us to render the higher-dimensional effective NP operators invariant, by expanding the Wilson coefficients  $C_i$  in terms of appropriate combinations of  $Y_U, Y_D$ .

As we are mainly interested in FCNC processes in the down quark sector, it is convenient to work in the down quark mass basis. In this basis, the Yukawa matrices simplify to

$$Y_D = \text{diag}(y_d, y_s, y_b), \quad Y_U = \hat{V}_{\text{CKM}}^\dagger \text{diag}(y_u, y_c, y_t), \quad (91)$$

where  $y_i = m_i/v$ .

Let us now consider the operator  $\mathcal{O}_1$  in the  $\Delta F = 2$  effective Hamiltonian (82):

$$\frac{C_1^{ij}}{\Lambda^2} (\bar{Q}_{Li} \gamma_\mu Q_{Lj}) (\bar{Q}_{Li} \gamma^\mu Q_{Lj}), \quad (92)$$

where  $i \neq j$  are flavour indices. Restoring the flavour symmetry  $G_{\text{flavour}}$ , we find the following expression for the Wilson coefficient  $C_1^{ij}$  in terms of the Yukawa couplings  $Y_U, Y_D$ :

$$\begin{aligned} C_1^{ij} &= (a \cdot \mathbb{1} + b \cdot Y_D Y_D^\dagger + c \cdot Y_U Y_U^\dagger + \dots)_{ij}^2 \\ &= (c \cdot \hat{V}_{\text{CKM}}^\dagger \text{diag}(y_u^2, y_c^2, y_t^2) \hat{V}_{\text{CKM}})_{ij}^2 \\ &\simeq c^2 \cdot y_t^4 (V_{ti}^* V_{tj})^2. \end{aligned} \quad (93)$$

Here  $a, b, c$  are real expansion parameters that are assumed to be  $\mathcal{O}(1)$ . We confirm that the NP contribution to  $C_1$  is suppressed by the same CKM factors (80) as the SM contribution.

Next, let us have a look at the operator  $\mathcal{O}_4$ ,

$$\frac{C_4^{ij}}{\Lambda^2} (\bar{D}_{Ri} Q_{Lj}) (\bar{Q}_{Li} D_{Rj}), \quad (94)$$

which was found to generate the strongest constraints on the scale of generic NP contributions. Again using the MFV hypothesis, we find

$$\begin{aligned} C_4^{ij} &= \left[ Y_D^\dagger (a \cdot \mathbb{1} + b \cdot Y_D Y_D^\dagger + c \cdot Y_U Y_U^\dagger + \dots) \right]_{ij} \\ &\quad \times \left[ (d \cdot \mathbb{1} + e \cdot Y_D Y_D^\dagger + f \cdot Y_U Y_U^\dagger + \dots) Y_D \right]_{ij} \\ &= c \cdot y_i y_t^2 (V_{ti}^* V_{tj}) \times f \cdot y_j y_t^2 (V_{ti}^* V_{tj}). \end{aligned} \quad (95)$$

Again  $a, b, c, d, e, f$  are real and  $\mathcal{O}(1)$  expansion parameters. We observe that in MFV the Wilson coefficient  $C_4$  is strongly suppressed not only by CKM elements, but in addition also by the masses of the external quarks  $i$  and  $j$ . The stringent constraints can therefore be evaded.

This observation is actually quite general. Whenever a right-handed down quark appears in a higher-dimensional operator, in the MFV framework it must necessarily be accompanied by at least one power of the Yukawa coupling  $Y_D$ . As the top mass is found to be much larger than the bottom mass, in the SM as well as in many concrete NP models, the hierarchy  $y_t \gg y_b$  holds. In this case, operators that involve additional  $Y_D$  factors become negligible. Therefore, only those operators are relevant for  $K$  and  $B$  physics that involve flavour transitions of left-handed quark fields  $Q_L$ . These are the ones that are already present in the SM effective Hamiltonian, due to the left-handedness of the weak interactions. Note however that not in all NP models the relation  $y_t \gg y_b$  holds. It can be violated in models with extended Higgs sectors, such as the MSSM at large  $\tan \beta$ .

These considerations lead us to the definition of a slightly more restrictive version of MFV, the framework of *Constrained Minimal Flavour Violation (CMFV)* [93, 94, 96]. Again, in CMFV the global quark flavour symmetry  $G_{\text{flavour}}$  is broken only by the SM Yukawa couplings  $Y_U$  and  $Y_D$ , and no new  $CP$ -violating phases are present. In addition to these MFV assumptions, in CMFV only those effective operators are relevant that are present already in the SM.

In the  $\Delta F = 2$  effective Hamiltonian then only the operator  $\mathcal{O}_1$  remains. As shown in (93), in MFV models it has the same CKM dependence  $(V_{ti}^* V_{tj})^2$  as the SM contribution. Consequently, in the CMFV scenario we can parametrise the NP contributions to meson mixing observables by a real and flavour-universal shift in the loop function

$$S_0(x_t) \rightarrow S(p), \quad (96)$$

where  $p$  collectively denotes the parameters of a given CMFV model. Hence in CMFV models the  $\Delta F = 2$  sector can be described by a single new parameter, and the NP contributions to neutral meson mixings in the various meson systems are correlated.

We can make use of this feature to construct the unitarity triangle in figure 2 from  $\Delta F = 2$  observables in such a way that the result holds for all CMFV models, independent of the value of the function  $S(p)$  [93]. As always, the value of  $|V_{us}|$  has to be determined from tree level decays, and is precisely known. The angle  $\beta$  can be measured in the time-dependent  $CP$  asymmetry of the  $B_d \rightarrow J/\psi K_S$  that measures  $CP$  violation in  $B_d - \bar{B}_d$  mixing. Finally, the length of the side  $R_t$  is proportional to the square root of the ratio  $\Delta M_d / \Delta M_s$ , where the NP contribution cancels out in the ratio.

The thus determined *universal unitarity triangle* is currently much more precisely known [97] than the one determined solely from tree level decays. While the two determinations currently show a good agreement, a potential mismatch observed with future more precise tree level data, would be an unambiguous sign of physics beyond the CMFV hypothesis.

We have thus seen that the MFV ansatz allows us to lower the scale of NP to the TeV range, without inducing large NP contributions to FCNC observables. In particular the CMFV hypothesis provides a very predictive framework, inducing many correlations between FCNC observables. A pedagogical review can be found in [94]. However, in the next section we will see that MFV is not the only option to reconcile TeV-scale NP with the flavour data.

### 3.4 Flavour Hierarchies from Partial Compositeness

It is also possible to suppress flavour violating processes without introducing flavour symmetries. Models with partially composite fermions are a well-known example where flavour hierarchies have a dynamical origin. In this section we only provide a brief and superficial overview on the flavour structure of composite models. More detailed, excellent lectures on composite Higgs models and their dual 5D description in the Randall-Sundrum framework can be found e. g. in [98–103].

The basic idea of composite Higgs models is the realization of the Higgs boson as a light composite state of a strongly coupled sector in analogy to the pions of QCD. The naturalness problem of fundamental scalars is thereby avoided. The EW symmetry is assumed to be broken by a condensate of the composite sector. Further composite resonances are then expected at the TeV scale, similar to the  $\rho, \omega \dots$  mesons of QCD with masses an order of magnitude larger than the pion mass. In order to avoid constraints from precision tests of the SM, the SM particles – except for the Higgs boson and the top quark – have to be mostly elementary.

In traditional composite models, like Technicolour [104, 105], the generation of realistic fermion masses turned out to be a major problem. As a solution, more recently the concept of partially composite fermions has been put forward [106]. In this setup, the known SM fermions and gauge bosons form an elementary sector. The composite sector, responsible for EW symmetry breaking, gives rise to operators describing composite resonances that have quantum numbers identical to the ones of the SM fields. The elementary fermions  $Q, U, D$  of (1) are coupled to the composite sector by a linear mixing with the composite operators  $\mathcal{O}_Q, \mathcal{O}_U, \mathcal{O}_D$ :

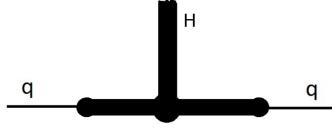
$$\mathcal{L}_{\text{mixing}} = \epsilon_Q \bar{Q}_L \mathcal{O}_Q + \epsilon_U \bar{U}_R \mathcal{O}_U + \epsilon_D \bar{D}_R \mathcal{O}_D. \quad (97)$$

The observed Yukawa couplings  $Y_U, Y_D$  are then a combination of the strong sector coupling responsible for the interactions  $\lambda_U, \lambda_D$  among the composite resonances and the elementary-composite fermion mixings  $\epsilon_Q, \epsilon_U, \epsilon_D$ ,

$$Y_{U,D} = \epsilon_Q \epsilon_{U,D} \lambda_{U,D}. \quad (98)$$

Figure 7 displays the generation of the effective Yukawa couplings diagrammatically. As the strong sector couplings are in general expected to be structureless (“anarchic”),

$$|\lambda_{U,D}^{ij}| \sim \mathcal{O}(1) \quad \arg \lambda_{U,D}^{ij} \sim \mathcal{O}(1), \quad (99)$$



**Fig. 7:** Effective Yukawa coupling in models with partially composite fermions.

the flavour hierarchies in the effective SM Yukawa couplings has to be induced by hierarchies in the elementary-composite mixing.

The stringent experimental constraints on non-SM interactions of the first two quark generations tell us that the latter must be mostly elementary, i. e. their mixing with the composite sector is small:

$$\epsilon_Q^{1,2}, \epsilon_U^{1,2}, \epsilon_D^{1,2} \ll 1. \quad (100)$$

The third generation is less constrained, and indeed the large top quark Yukawa coupling requires that

$$\epsilon_Q^3 \sim \epsilon_U^3 \sim \mathcal{O}(1). \quad (101)$$

Overall, the observed quark masses and the experimental precision constraints imply the pattern

$$\epsilon_Q^1 \ll \epsilon_Q^2 \ll \epsilon_Q^3 \sim \mathcal{O}(1), \quad \epsilon_U^1 \ll \epsilon_U^2 \ll \epsilon_U^3 \sim \mathcal{O}(1), \quad \epsilon_D^1 \ll \epsilon_D^2 \ll \epsilon_D^3 \ll 1. \quad (102)$$

Inserting this pattern into (98) and assuming an anarchic composite sector, we find the following hierarchical structure for the observed Yukawa couplings

$$Y_U \sim \begin{pmatrix} \epsilon_Q^1 \epsilon_U^1 & \epsilon_Q^1 \epsilon_U^2 & \epsilon_Q^1 \epsilon_U^3 \\ \epsilon_Q^2 \epsilon_U^1 & \epsilon_Q^2 \epsilon_U^2 & \epsilon_Q^2 \epsilon_U^3 \\ \epsilon_Q^3 \epsilon_U^1 & \epsilon_Q^3 \epsilon_U^2 & \epsilon_Q^3 \epsilon_U^3 \end{pmatrix}, \quad Y_D \sim \begin{pmatrix} \epsilon_Q^1 \epsilon_D^1 & \epsilon_Q^1 \epsilon_D^2 & \epsilon_Q^1 \epsilon_D^3 \\ \epsilon_Q^2 \epsilon_D^1 & \epsilon_Q^2 \epsilon_D^2 & \epsilon_Q^2 \epsilon_D^3 \\ \epsilon_Q^3 \epsilon_D^1 & \epsilon_Q^3 \epsilon_D^2 & \epsilon_Q^3 \epsilon_D^3 \end{pmatrix}. \quad (103)$$

Note that the hierarchical structure of the effective Yukawa couplings  $Y_U, Y_D$  is analogous to the one obtained in models with a Froggatt-Nielsen [107] flavour symmetry [108–110].

Diagonalizing these matrices, we not only recover the observed quark mass hierarchy, but we also find predictions for the off-diagonal elements of the CKM matrix [111, 112]:

$$|V_{us}| \sim \frac{\epsilon_Q^1}{\epsilon_Q^2} \ll 1, \quad |V_{cb}| \sim \frac{\epsilon_Q^2}{\epsilon_Q^3} \ll 1, \quad |V_{ub}| \sim \frac{\epsilon_Q^1}{\epsilon_Q^3} \ll 1. \quad (104)$$

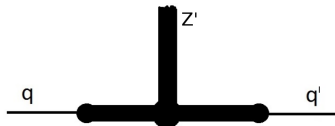
The measurements of  $|V_{us}|$  and  $|V_{cb}|$  fix the hierarchies among the  $\epsilon_Q^i$  parameters. We then obtain a prediction for  $|V_{ub}|$ :

$$|V_{ub}| \sim |V_{us}| \cdot |V_{cb}| \sim 0.2 \cdot 4 \cdot 10^{-2} = 8 \cdot 10^{-2}. \quad (105)$$

This number is larger than the measured value of  $|V_{ub}|$  by a factor of two. Keeping in mind that in the derivation of (105) we dropped  $\mathcal{O}(1)$  factors, this result is quite remarkable.

In addition to providing a dynamical origin for the observed pattern of the SM quark masses and CKM mixings, the hierarchies in the elementary-composite fermion mixing also efficiently suppress tree level FCNC couplings [113] that are generated in the composite sector and mediated to the SM fermions by the elementary-composite mixing. Assuming again that the composite sector couplings are anarchic, the FCNC couplings of the SM quarks are suppressed by the same hierarchical pattern as the effective Yukawa couplings  $Y_U, Y_D$ , as can be seen from comparing the FCNC coupling in figure 8 to the Yukawa coupling in figure 7.

With the masses of the composite resonances at a few TeV, as required by EW precision data and direct LHC searches, the suppression of FCNC couplings is sufficient to comply with most of the



**Fig. 8:** Tree level flavour changing coupling of a composite  $Z'$  gauge boson in models with partially composite fermions.

available flavour data. Some tension arises however with  $CP$ -violating observables in the kaon sector, specifically  $\varepsilon$  [114] and  $\varepsilon'$  [115]. The latter constraints require the masses of the strong sector resonances to be above 10 TeV, unless some additional structure is imposed on the composite sector to evade the constraints.

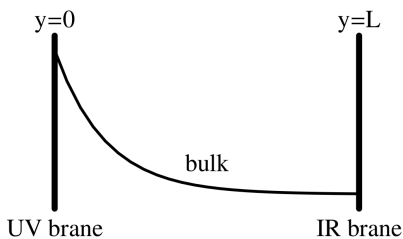
In models with partially composite fermions, the SM flavour hierarchies is thus traced back to the exponentially small mixing of elementary fermions with the composite sector containing the Higgs. But where does this exponential suppression come from?

The origin of the flavour hierarchies in partially composite models can be best understood by considering the *holographic dual* of these theories. In classical optics, holograms are known as 2D images containing the information of a 3D object. In that sense, the 3D object and its 2D holographic image are dual to each other: they contain the same information. In a similar way, it has been proposed that a 4D composite model is, under certain conditions, dual to a 5D weakly coupled model. By studying the properties of the dual 5D model, we can then obtain a better understanding of the 4D composite model.

The foundation for the construction of the 5D holographic dual is laid by the AdS/CFT-correspondence [116]. It has been conjectured that a strongly coupled 4D conformal theory is dual to a weakly coupled 5D theory in the Randall-Sundrum (RS) [117] background. Conformal symmetry is a particular internal symmetry of the strongly coupled sector – for our purposes it is sufficient to know that conformality implies scale invariance. The same symmetry group,  $SO(4, 2)$  can also be implemented as a space-time symmetry. In order to achieve this, our 4D space-time has to be extended by one extra dimension. The symmetry of the 5D space-time is then described by the RS metric [117]

$$ds^2 = e^{-2ky} \eta_{\mu\nu} dx^\mu dx^\nu - dy^2, \quad (106)$$

displayed in figure 9. The extra dimensional coordinate  $y$  is confined to an interval  $0 \leq y \leq L$ . In the dual strongly coupled theory, the coordinate  $y$  corresponds to the energy scale  $\Lambda$  of the theory. The endpoint  $y = 0$ , called UV brane, corresponds to the Planck scale where the conformal symmetry is explicitly broken. The other endpoint  $y = L$ , the IR brane, corresponds to the TeV scale, where a strong sector condensate spontaneously breaks conformality.



**Fig. 9:** RS space-time.

The change of energy scales along the fifth coordinate  $y$  can be understood by having another look at the RS metric (106). The “warp factor”  $e^{-2ky}$  induces a  $y$  dependence of the effective energy scale

$$\Lambda(y) = e^{-ky} \Lambda_0 \quad (107)$$

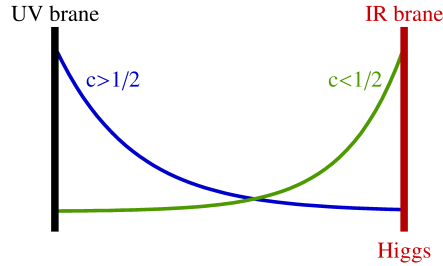
in terms of the fundamental energy scale  $\Lambda_0$  of the theory. The large hierarchy between the Planck scale and the scale of EW symmetry breaking can then be explained by localizing the Higgs boson on the IR brane where the effective cut-off scale is warped down to the TeV scale.

Due to the confinement of the coordinate  $y$  to an interval, the fifth dimension is too small to be directly observed at low energies, and our world effectively appears four-dimensional. However the extra dimension leaves observable traces in terms of its 4D remnants, the Kaluza-Klein (KK) modes [118]. To understand their origin, let us recall the description of a potential well in quantum mechanics. Due to the wave function confinement to an interval, an infinite tower of discrete modes appears, with quantized energy levels. The same concept applies when considering 5D fields, when the fifth dimension is confined to an interval: Integrating out the dynamics of the extra dimension, an infinite tower of 4D fields arises, with identical quantum numbers and increasing masses. The lowest-lying mode is massless and identified with the corresponding SM field. The lowest excited modes have masses in the TeV range. They are the dual states of the composite resonances in the strongly coupled 4D theory.

Of particular interest is the fermion sector of RS models. The localization of the fermionic zero modes, corresponding to the SM fermions, is described by the wave function [118–120]

$$f^{(0)}(y, c) \propto e^{(\frac{1}{2}-c)ky}, \quad (108)$$

where  $c$  is the bulk mass parameter of the respective fermion, a fundamental parameter of the 5D Lagrangian. As for all dimensionless parameters, naturally  $c \sim \mathcal{O}(1)$ . The localization of a given fermion zero mode and its overlap with the Higgs boson wave function localized on the IR brane hence exponentially depends on its bulk mass parameter  $c$ , as shown in figure 10.



**Fig. 10:** Fermion zero mode localisation in the RS background.

The effective 4D Yukawa couplings then have the following structure:

$$(Y_{U,D})_{ij} \sim f_i^Q f_j^{U,D} (\lambda_{U,D})_{ij}. \quad (109)$$

Here,  $\lambda_{U,D}$  are the fundamental Yukawa couplings of the full 5D theory that are assumed to be anarchic, and  $f_i^Q, f_j^{U,D}$  are the relevant fermion zero mode wave functions evaluated on the IR brane.

Comparing (109) to the effective Yukawa couplings in partially composite fermion models, (98), we can straightforwardly deduce the dual interpretation of the fermion zero mode localization. Fermion zero modes with  $c < 1/2$  are localized close to the IR brane and couple strongly to the Higgs boson. These are the heavy fermions that in the 4D theory have a large elementary-composite mixing. The light, mostly elementary fermions, on the other hand, have  $c > 1/2$  and are therefore localized near the UV brane. Consequently their coupling to the Higgs boson is exponentially suppressed. We conclude that the localization of fields along the 5D bulk corresponds to their degree of compositeness in the 4D dual theory.

Making use of the AdS/CFT correspondence, we have thus found a dynamical origin for the exponential suppression of the elementary-composite mixing of the light fermions in partially composite

models. In the 5D dual of these theories, the exponential flavour hierarchies arise from different values for their bulk mass parameters, all naturally of  $\mathcal{O}(1)$ . These models therefore provide an appealing alternative to models with approximate flavour symmetries. The flavour phenomenology of RS models with bulk fermions has been the subject of many detailed studies, see e. g. [109, 114, 121–124].

## Summary and Outlook

This lecture series provided a basic introduction to flavour physics and  $CP$  violation, as well as an overview over some current hot topics.

In lecture 1 we reviewed the basics of flavour physics in the SM. The flavour symmetry of the SM is violated by the Yukawa couplings, which give rise to the quark masses and the CKM mixing matrix. Their observed very hierarchical structure constitutes the SM flavour problem and calls for a more fundamental theory of flavour. We also outlined the theoretical description of flavour violating processes in the SM and beyond, using the effective Hamiltonian and the operator product expansion.

Lecture 2 was devoted to the discussion of some basic phenomenological concepts in  $K$  and  $B$  meson physics. As flavour physics in the quark sector has a very rich phenomenology, we restricted our attention to a few but very important representative examples here. We introduced the physics of neutral kaon mixing and  $CP$  violation in  $K \rightarrow \pi\pi$  decays, described by the parameters  $\varepsilon$  and  $\varepsilon'$ . We then turned our attention to the very rare decays  $K^+ \rightarrow \pi^+\nu\bar{\nu}$  and  $K_L \rightarrow \pi^0\nu\bar{\nu}$ , which, due to their theoretical cleanliness, offer an excellent probe of NP even at large energy scales. In the  $B$  system, we focused on the semileptonic  $b \rightarrow s\mu^+\mu^-$  transitions, for which several anomalies have been found in recent data.

In lecture 3 we gave an introduction to flavour physics in theories beyond the SM. We started by identifying the stringent constraints on the scale of generic NP contributions obtained from the neutral meson mixing observables. We then introduced two concepts that suppress large non-standard contributions to FCNC observables and thereby reconcile TeV-scale NP with flavour data. In Minimal Flavour Violation, the SM Yukawa couplings are assumed to be the only source of flavour and  $CP$  violation also in the NP sector. Consequently, the new flavour violating effects are suppressed by the same hierarchical structures as in the SM, and a predictive pattern of correlations arises. A dynamical origin of flavour hierarchies, on the other hand, is provided by models with partially composite fermions, which are dual to 5D theories in the Randall-Sundrum background.

Flavour physics has played an essential role in the construction of the SM. More recently, its importance has shifted from measuring the parameters of the SM to hunting for possible NP contributions. With the lack of direct discoveries of new particles at the LHC, indirect searches for NP are becoming increasingly relevant. Indeed, some of the most convincing anomalies of today's particle physics are related to the flavour sector. In addition, over the coming years, a lot of experimental progress will be made, with the potential for striking NP discoveries.

## Acknowledgements

I am grateful to the organizers of the ESHEP 2016 school for inviting me to give the series of lectures presented here, and for putting so much time and effort into making this school a very successful and enjoyable one. I also thank all participants for the pleasant and fruitful atmosphere in Skeikampen. Last but not least I am much obliged to the Jotunheimen mountain guides for not leaving me behind on the Besseggen ridge.

## References

- [1] M. Gell-Mann, Phys. Lett. **8** (1964) 214.
- [2] S. L. Glashow, J. Iliopoulos and L. Maiani, Phys. Rev. D **2** (1970) 1285.
- [3] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49** (1973) 652.



- [4] A. J. Buras, Phys. Lett. B **317** (1993) 449 [hep-ph/9307318].
- [5] A. J. Buras and J. Girrbach, Rept. Prog. Phys. **77** (2014) 086201 [arXiv:1306.3775 [hep-ph]].
- [6] G. Barenboim, arXiv:1610.09835 [hep-ph].
- [7] G. Isidori, arXiv:1302.0661 [hep-ph].
- [8] Y. Grossman, arXiv:1006.3534 [hep-ph].
- [9] A. J. Buras, hep-ph/0505175.
- [10] A. J. Buras, hep-ph/9806471.
- [11] N. Cabibbo, Phys. Rev. Lett. **10** (1963) 531.
- [12] L. L. Chau and W. Y. Keung, Phys. Rev. Lett. **53** (1984) 1802.
- [13] C. Patrignani *et al.* [Particle Data Group], Chin. Phys. C **40** (2016) no.10, 100001.
- [14] C. Jarlskog, Phys. Rev. Lett. **55** (1985) 1039.
- [15] C. Jarlskog, Z. Phys. C **29** (1985) 491.
- [16] L. Wolfenstein, Phys. Rev. Lett. **51** (1983) 1945.
- [17] A. J. Buras, M. E. Lautenbacher and G. Ostermaier, Phys. Rev. D **50** (1994) 3433 [hep-ph/9403384].
- [18] M. Schmidtler and K. R. Schubert, Z. Phys. C **53** (1992) 347.
- [19] C. Jarlskog and R. Stora, Phys. Lett. B **208** (1988) 268.
- [20] Y. Grossman, Y. Nir and M. P. Worah, Phys. Lett. B **407** (1997) 307 [hep-ph/9704287].
- [21] K. G. Wilson, Phys. Rev. **179** (1969) 1499.
- [22] L. Lellouch, arXiv:1104.5484 [hep-lat].
- [23] A. V. Radyushkin, hep-ph/0101227.
- [24] J. A. Bailey *et al.* [Fermilab Lattice and MILC Collaborations], Phys. Rev. D **92** (2015) no.1, 014024 [arXiv:1503.07839 [hep-lat]].
- [25] E. Dalgic, A. Gray, M. Wingate, C. T. H. Davies, G. P. Lepage and J. Shigemitsu, Phys. Rev. D **73** (2006) 074502 Erratum: [Phys. Rev. D **75** (2007) 119906] [hep-lat/0601021].
- [26] J. M. Flynn, T. Izubuchi, T. Kawanai, C. Lehner, A. Soni, R. S. Van de Water and O. Witzel, Phys. Rev. D **91** (2015) no.7, 074510 [arXiv:1501.05373 [hep-lat]].
- [27] A. J. Buras, M. Jamin and P. H. Weisz, Nucl. Phys. B **347** (1990) 491.
- [28] J. Urban, F. Krauss, U. Jentschura and G. Soff, Nucl. Phys. B **523** (1998) 40 [hep-ph/9710245].
- [29] J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay, Phys. Rev. Lett. **13** (1964) 138.
- [30] A. J. Buras and D. Guadagnoli, Phys. Rev. D **78** (2008) 033005 [arXiv:0805.3887 [hep-ph]].
- [31] A. J. Buras, D. Guadagnoli and G. Isidori, Phys. Lett. B **688** (2010) 309 [arXiv:1002.3612 [hep-ph]].
- [32] J. Brod and M. Gorbahn, Phys. Rev. Lett. **108** (2012) 121801 [arXiv:1108.2036 [hep-ph]].
- [33] Z. Bai *et al.* [RBC and UKQCD Collaborations], Phys. Rev. Lett. **115** (2015) no.21, 212001 [arXiv:1505.07863 [hep-lat]].
- [34] A. J. Buras and J. M. Gerard, JHEP **1512** (2015) 008 [arXiv:1507.06326 [hep-ph]].
- [35] A. J. Buras and J. M. Gerard, Eur. Phys. J. C **77** (2017) no.1, 10 [arXiv:1603.05686 [hep-ph]].
- [36] A. J. Buras and M. Jamin, JHEP **0401** (2004) 048 [hep-ph/0306217].
- [37] A. J. Buras, M. Gorbahn, S. Jäger and M. Jamin, JHEP **1511** (2015) 202 [arXiv:1507.06345 [hep-ph]].
- [38] T. Kitahara, U. Nierste and P. Tremper, JHEP **1612** (2016) 078 [arXiv:1607.06727 [hep-ph]].
- [39] E. Pallante, A. Pich and I. Scimemi, Nucl. Phys. B **617** (2001) 441 [hep-ph/0105011].
- [40] M. Blanke, A. J. Buras and S. Recksiegel, Eur. Phys. J. C **76** (2016) no.4, 182 [arXiv:1507.06316 [hep-ph]].

- [41] A. J. Buras, D. Buttazzo and R. Knegjens, JHEP **1511** (2015) 166 [arXiv:1507.08672 [hep-ph]].
- [42] A. J. Buras, JHEP **1604** (2016) 071 [arXiv:1601.00005 [hep-ph]].
- [43] A. J. Buras and F. De Fazio, JHEP **1603** (2016) 010 [arXiv:1512.02869 [hep-ph]].
- [44] A. J. Buras and F. De Fazio, JHEP **1608** (2016) 115 [arXiv:1604.02344 [hep-ph]].
- [45] M. Tanimoto and K. Yamamoto, PTEP **2016** (2016) no.12, 123B02 [arXiv:1603.07960 [hep-ph]].
- [46] T. Kitahara, U. Nierste and P. Tremper, Phys. Rev. Lett. **117** (2016) no.9, 091802 [arXiv:1604.07400 [hep-ph]].
- [47] C. Bobeth, A. J. Buras, A. Celis and M. Jung, arXiv:1609.04783 [hep-ph].
- [48] A. J. Buras, D. Buttazzo, J. Girrbach-Noe and R. Knegjens, JHEP **1411** (2014) 121 [arXiv:1408.0728 [hep-ph]].
- [49] G. Buchalla, A. J. Buras and M. E. Lautenbacher, Rev. Mod. Phys. **68** (1996) 1125 [hep-ph/9512380].
- [50] A. J. Buras, M. Gorbahn, U. Haisch and U. Nierste, Phys. Rev. Lett. **95** (2005) 261805 [hep-ph/0508165].
- [51] A. J. Buras, M. Gorbahn, U. Haisch and U. Nierste, JHEP **0611** (2006) 002 Erratum: [JHEP **1211** (2012) 167] [hep-ph/0603079].
- [52] J. Brod and M. Gorbahn, Phys. Rev. D **78** (2008) 034006 [arXiv:0805.4119 [hep-ph]].
- [53] A. J. Buras, D. Buttazzo, J. Girrbach-Noe and R. Knegjens, JHEP **1511** (2015) 033 [arXiv:1503.02693 [hep-ph]].
- [54] F. Newson *et al.*, arXiv:1411.0109 [hep-ex].
- [55] T. Yamanaka [KOTO Collaboration], PTEP **2012** (2012) 02B006.
- [56] A. J. Buras, R. Fleischer, S. Recksiegel and F. Schwab, Nucl. Phys. B **697** (2004) 133 [hep-ph/0402112].
- [57] M. Blanke, Acta Phys. Polon. B **41** (2010) 127 [arXiv:0904.2528 [hep-ph]].
- [58] Y. Grossman and Y. Nir, Phys. Lett. B **398** (1997) 163 [hep-ph/9701313].
- [59] M. Blanke, A. J. Buras, A. Poschenrieder, S. Recksiegel, C. Tarantino, S. Uhlig and A. Weiler, JHEP **0701** (2007) 066 [hep-ph/0610298].
- [60] A. J. Buras, F. De Fazio, J. Girrbach and M. V. Carlucci, JHEP **1302** (2013) 023 [arXiv:1211.1237 [hep-ph]].
- [61] A. J. Buras, F. De Fazio and J. Girrbach, JHEP **1302** (2013) 116 [arXiv:1211.1896 [hep-ph]].
- [62] A. Crivellin, G. D'Ambrosio, T. Kitahara and U. Nierste, arXiv:1703.05786 [hep-ph].
- [63] A. Abulencia *et al.* [CDF Collaboration], Phys. Rev. Lett. **97** (2006) 242003 [hep-ex/0609040].
- [64] R. Aaij *et al.* [LHCb Collaboration], Phys. Lett. B **709** (2012) 177 [arXiv:1112.4311 [hep-ex]].
- [65] R. Aaij *et al.* [LHCb Collaboration], Phys. Rev. Lett. **114** (2015) no.4, 041801 [arXiv:1411.3104 [hep-ex]].
- [66] W. Altmannshofer, P. Ball, A. Bharucha, A. J. Buras, D. M. Straub and M. Wick, JHEP **0901** (2009) 019 [arXiv:0811.1214 [hep-ph]].
- [67] U. Egede, T. Hurth, J. Matias, M. Ramon and W. Reece, Acta Phys. Polon. BProceed **.Suppl.3** (2010) 151 [arXiv:0912.1339 [hep-ph]].
- [68] S. Descotes-Genon, T. Hurth, J. Matias and J. Virto, JHEP **1305** (2013) 137 [arXiv:1303.5794 [hep-ph]].
- [69] R. Aaij *et al.* [LHCb Collaboration], JHEP **1602** (2016) 104 [arXiv:1512.04442 [hep-ex]].
- [70] The ATLAS collaboration [ATLAS Collaboration], ATLAS-CONF-2017-023.
- [71] S. Wehle *et al.* [Belle Collaboration], Phys. Rev. Lett. **118** (2017) no.11, 111801 [arXiv:1612.05014 [hep-ex]].

- [72] CMS Collaboration [CMS Collaboration], CMS-PAS-BPH-15-008.
- [73] S. Descotes-Genon, L. Hofer, J. Matias and J. Virto, JHEP **1606** (2016) 092 [arXiv:1510.04239 [hep-ph]].
- [74] B. Capdevila, S. Descotes-Genon, L. Hofer, J. Matias and J. Virto, PoS LHCP **2016** (2016) 073 [arXiv:1609.01355 [hep-ph]].
- [75] F. Beaujean, C. Bobeth and S. Jahn, Eur. Phys. J. C **75** (2015) no.9, 456 [arXiv:1508.01526 [hep-ph]].
- [76] W. Altmannshofer, C. Niehoff, P. Stangl and D. M. Straub, arXiv:1703.09189 [hep-ph].
- [77] R. Aaij *et al.* [LHCb Collaboration], Phys. Rev. Lett. **113** (2014) 151601 [arXiv:1406.6482 [hep-ex]].
- [78] G. Hiller and F. Kruger, Phys. Rev. D **69** (2004) 074020 [hep-ph/0310219].
- [79] S. Bifani [LHCb Collaboration], CERN LHC seminar April 18, 2017. Slides available on <https://indico.cern.ch/event/580620/>.
- [80] W. Altmannshofer and D. M. Straub, Eur. Phys. J. C **73** (2013) 2646 [arXiv:1308.1501 [hep-ph]].
- [81] R. Gauld, F. Goertz and U. Haisch, Phys. Rev. D **89** (2014) 015005 [arXiv:1308.1959 [hep-ph]].
- [82] A. J. Buras and J. Girrbach, JHEP **1312** (2013) 009 [arXiv:1309.2466 [hep-ph]].
- [83] W. Altmannshofer, S. Gori, M. Pospelov and I. Yavin, Phys. Rev. D **89** (2014) 095033 [arXiv:1403.1269 [hep-ph]].
- [84] C. W. Chiang, X. G. He and G. Valencia, Phys. Rev. D **93** (2016) no.7, 074003 [arXiv:1601.07328 [hep-ph]].
- [85] A. Crivellin, J. Fuentes-Martin, A. Greljo and G. Isidori, Phys. Lett. B **766** (2017) 77 [arXiv:1611.02703 [hep-ph]].
- [86] G. Hiller and M. Schmaltz, Phys. Rev. D **90** (2014) 054014 [arXiv:1408.1627 [hep-ph]].
- [87] M. Bauer and M. Neubert, Phys. Rev. Lett. **116** (2016) no.14, 141802 [arXiv:1511.01900 [hep-ph]].
- [88] S. Fajfer and N. Kosnik, Phys. Lett. B **755** (2016) 270 [arXiv:1511.06024 [hep-ph]].
- [89] S. Jäger and J. Martin Camalich, Phys. Rev. D **93** (2016) no.1, 014028 [arXiv:1412.3183 [hep-ph]].
- [90] B. Capdevila, S. Descotes-Genon, L. Hofer and J. Matias, JHEP **1704** (2017) 016 [arXiv:1701.08672 [hep-ph]].
- [91] S. Weinberg, Phys. Rev. Lett. **43** (1979) 1566.
- [92] M. Bona *et al.* [UTfit Collaboration], PoS ICHEP **2016** (2016) 149.
- [93] A. J. Buras, P. Gambino, M. Gorbahn, S. Jager and L. Silvestrini, Phys. Lett. B **500** (2001) 161 [hep-ph/0007085].
- [94] A. J. Buras, Acta Phys. Polon. B **34** (2003) 5615 [hep-ph/0310208].
- [95] G. D'Ambrosio, G. F. Giudice, G. Isidori and A. Strumia, Nucl. Phys. B **645** (2002) 155 [hep-ph/0207036].
- [96] M. Blanke, A. J. Buras, D. Guadagnoli and C. Tarantino, JHEP **0610** (2006) 003 [hep-ph/0604057].
- [97] M. Blanke and A. J. Buras, Eur. Phys. J. C **76** (2016) no.4, 197 [arXiv:1602.04020 [hep-ph]].
- [98] C. Csaki, [hep-ph/0404096].
- [99] R. Sundrum, hep-th/0508134.
- [100] C. Csaki, J. Hubisz and P. Meade, hep-ph/0510275.
- [101] R. Contino, arXiv:1005.4269 [hep-ph].
- [102] T. Gherghetta, arXiv:1008.2570 [hep-ph].
- [103] E. Ponton, arXiv:1207.3827 [hep-ph].
- [104] R. K. Kaul, Rev. Mod. Phys. **55** (1983) 449.
- [105] K. Lane, hep-ph/0202255.

- [106] D. B. Kaplan, Nucl. Phys. B **365** (1991) 259.
- [107] C. D. Froggatt and H. B. Nielsen, Nucl. Phys. B **147** (1979) 277.
- [108] S. Casagrande, F. Goertz, U. Haisch, M. Neubert and T. Pfoh, JHEP **0810** (2008) 094 [arXiv:0807.4937 [hep-ph]].
- [109] M. Blanke, A. J. Buras, B. Duling, S. Gori and A. Weiler, JHEP **0903** (2009) 001 [arXiv:0809.1073 [hep-ph]].
- [110] M. E. Albrecht, M. Blanke, A. J. Buras, B. Duling and K. Gemmler, JHEP **0909** (2009) 064 [arXiv:0903.2415 [hep-ph]].
- [111] S. J. Huber and Q. Shafi, Phys. Lett. B **498** (2001) 256 [hep-ph/0010195].
- [112] S. J. Huber, Nucl. Phys. B **666** (2003) 269 [hep-ph/0303183].
- [113] K. Agashe, G. Perez and A. Soni, Phys. Rev. D **71** (2005) 016002 [hep-ph/0408134].
- [114] C. Csaki, A. Falkowski and A. Weiler, JHEP **0809**, 008 (2008) [arXiv:0804.1954 [hep-ph]].
- [115] O. Gedalia, G. Isidori and G. Perez, Phys. Lett. B **682** (2009) 200 [arXiv:0905.3264 [hep-ph]].
- [116] J. M. Maldacena, Int. J. Theor. Phys. **38** (1999) 1113 [Adv. Theor. Math. Phys. **2** (1998) 231] [hep-th/9711200].
- [117] L. Randall and R. Sundrum, Phys. Rev. Lett. **83** (1999) 3370 [hep-ph/9905221].
- [118] T. Gherghetta and A. Pomarol, Nucl. Phys. B **586** (2000) 141 [hep-ph/0003129].
- [119] Y. Grossman and M. Neubert, Phys. Lett. B **474** (2000) 361 [hep-ph/9912408].
- [120] G. Burdman, Phys. Rev. D **66** (2002) 076003 [hep-ph/0205329].
- [121] K. Agashe, G. Perez and A. Soni, Phys. Rev. Lett. **93**, 201804 (2004) [hep-ph/0406101].
- [122] M. Blanke, A. J. Buras, B. Duling, K. Gemmler and S. Gori, JHEP **0903** (2009) 108 [arXiv:0812.3803 [hep-ph]].
- [123] M. Bauer, S. Casagrande, U. Haisch and M. Neubert, JHEP **1009** (2010) 017 [arXiv:0912.1625 [hep-ph]].
- [124] M. Blanke, B. Shakya, P. Tanedo and Y. Tsai, JHEP **1208** (2012) 038 [arXiv:1203.6650 [hep-ph]].

# Neutrinos: Fast & Curious

*G. Barenboim*

Departament de Física Teòrica and IFIC, Universitat de València- CSIC, València, Spain

## Abstract

The Standard Model has been effective way beyond expectations in foreseeing the result of almost all the experimental tests done up so far. In it, neutrinos are massless. Nonetheless, in recent years we have collected solid proofs indicating little but non zero masses for the neutrinos (when contrasted with those of the charged leptons). These masses permit neutrinos to change their flavor and oscillate, indeed a unique treat. In these lectures, I discuss the properties and the amazing potential of neutrinos in and beyond the Standard Model.

## Keywords

Lectures; neutrino; neutrino oscillations; mass; flavor; majorana.

## 1 Introduction

Last decade witnessed a brutal transformation in neutrino physics. It has been experimentally observed that neutrinos have nonzero masses, implying that leptons blend. This fact was demonstrated by the experimental evidence that neutrinos can change from one state, or “flavour”, to another. All the information we have accumulated about neutrinos, is quite recent. Less than twenty years old. Neutrino physics as a solid science is in its teenage years and therefore as any adolescence, in a wild and very exciting (and excited) state.

However, before jumping into the late “news” about neutrinos, lets understand how and why neutrinos were conceived.

The '20s saw the death of numerous sacred cows, and physics was no exemption. One of physic’s most holly principles, energy conservation, apparently showed up not to hold inside the subatomic world.

For some radioactive nuclei, it appeared that a non-negligible fraction of its energy simply vanished, leaving no trace of its presence.

In 1920, in a (by now famous) letter to a meeting [1], Pauli quasi apologetically wrote, “Dear radioactive Ladies and Gentlemen, ... as a desperate remedy to save the principle of energy conservation in beta decay, ... I propose the idea of a neutral particle of spin half”. Pauli hypothesised that the missing energy was taken off by another particle, whose properties were such that made it invisible and impossible to detect: it had no electric charge, no mass and only very rarely interacted with matter. Along these lines, the neutrino was naturally introduced to the universe of particle physics.

Before long, Fermi postulated the four-Fermi Hamiltonian in order to describe beta decay utilising the neutrino, electron, neutron and proton. Another field was born: weak interactions took the stage to never leave it.

Closing the loop, twenty years after Pauli’s letter, Cowan and Reines got the experimental signature of anti-neutrinos emitted by a nuclear power plant.

As more particles who participated in weak interactions were found in the years following neutrino discovery, weak interactions got credibility as an authentic new force of nature and the neutrino got to be a key element of it.

Further experimental tests through the span of the following 30 years demonstrated that there were not one but three sort, or “flavours” of neutrinos (electron neutrinos ( $\nu_e$ ), muon neutrinos ( $\nu_\mu$ ) and tau neutrinos ( $\nu_\tau$ )) and that, to the extent we could test, had no mass (and no charge) whatsoever.

The neutrino adventure could have easily finish there, however new analyses in neutrinos coming from the sun shown us that the neutrino saga was just beginning....

In the canonical Standard Model, neutrinos are completely massless and as a consequence are flavour eigenstates,

$$\begin{aligned} W^+ &\longrightarrow e^+ + \nu_e & ; & & Z &\longrightarrow \nu_e + \bar{\nu}_e \\ W^+ &\longrightarrow \mu^+ + \nu_\mu & ; & & Z &\longrightarrow \nu_\mu + \bar{\nu}_\mu \\ W^+ &\longrightarrow \tau^+ + \nu_\tau & ; & & Z &\longrightarrow \nu_\tau + \bar{\nu}_\tau \end{aligned} \quad (1)$$

Precisely because they are massless, they travel at the speed of light and accordingly their flavour does not change from generation up to detection. It is evident then, that as flavour is concerned, zero mass neutrinos are not an attractive object to study, specially when contrasted with quarks.

However, if neutrinos were massive, and these masses where not degenerate (degenerate masses flavour-wise is identical to the zero mass case) would mean that neutrino mass eigenstates exist  $\nu_i, i = 1, 2, \dots$ , each with a mass  $m_i$ . The impact of leptonic mixing becomes apparent by looking at the leptonic decays,  $W^+ \longrightarrow \nu_i + \bar{\ell}_\alpha$  of the charged vector boson  $W$ . Where,  $\alpha = e, \mu$ , or  $\tau$ , and  $\ell_e$  refers to the electron,  $\ell_\mu$  the muon, or  $\ell_\tau$  the tau.

We call particle  $\ell_\alpha$  as the charged lepton of flavour  $\alpha$ . Mixing basically implies that when the charged boson  $W^+$  decays to a given kind of charged lepton  $\bar{\ell}_\alpha$ , the neutrino that goes along is not generally the same mass eigenstate  $\nu_i$ . Any of the different  $\nu_i$  can appear.

The amplitude for the decay of a vector boson  $W^+$  to a particular mix  $\bar{\ell}_\alpha + \nu_i$  is given by  $U_{\alpha i}^*$ . The neutrino that is radiated in this decay alongside the given charged lepton  $\bar{\ell}_\alpha$  is then

$$|\nu_\alpha \rangle = \sum_i U_{\alpha i}^* |\nu_i \rangle . \quad (2)$$

This specific mixture of mass eigenstates yields the neutrino of flavour  $\alpha$ .

The different  $U_{\alpha i}$  can be gathered in a unitary matrix (in the same way they were collected in the CKM matrix in the quark sector) that receives the name of the leptonic mixing matrix or  $U_{PNMS}$  [2]. The unitarity of  $U$  ensures that each time a neutrino of flavour  $\alpha$  through its interaction produces a charged lepton, the produced charged lepton will always be  $\ell_\alpha$ , the charged lepton of flavour  $\alpha$ . That is, a  $\nu_e$  produces exclusively an  $e$ , a  $\nu_\mu$  exclusively a  $\mu$ , and in a similar way  $\nu_\tau$  a  $\tau$ .

The expression (2), portraying each neutrino of a given flavour as a linear combination of the three mass eigenstates, can be easily inverted to depict every mass eigenstate  $\nu_i$  as an analogous linear combination of the three flavours:

$$|\nu_i \rangle = \sum_\alpha U_{\alpha i} |\nu_\alpha \rangle . \quad (3)$$

The amount of  $\alpha$ -flavour (or the  $\alpha$ -fraction) of  $\nu_i$  is obviously  $|U_{\alpha i}|^2$ . When a  $\nu_i$  interacts and creates a charged lepton, this  $\alpha$ -content (or fraction) expresses the probability that the created charged lepton be of flavour  $\alpha$ .

## 2 Neutrino Oscillations basics

The phenomenon of neutrino morphing, flavour transition or in short oscillation, can be understood in the following form.

A neutrino is created or emitted by a source along with a charged lepton  $\bar{\ell}_\alpha$  of flavour  $\alpha$ . In this way, at the emission point, the neutrino does have a definite flavour. It is a  $\nu_\alpha$ . After that point, after production, the neutrino covers some length (propagates thorough a distance)  $L$  until it is absorbed.

At this point, when it has already covered the distance to the target, the neutrino interacts and these interactions create another charged lepton  $\ell_\beta$  of flavour  $\beta$ , which we can detect. In this way, at the target, we can know that the neutrino is again a neutrino of definite flavour, a  $\nu_\beta$ . Of course there is a chance that  $\beta \neq \alpha$  (for instance, if  $\ell_\alpha$  is a  $\mu$  however  $\ell_\beta$  is a  $\tau$ ), then, all along his journey from the source to the identification point, the neutrino has morphed or transformed from a  $\nu_\alpha$  into a  $\nu_\beta$ .

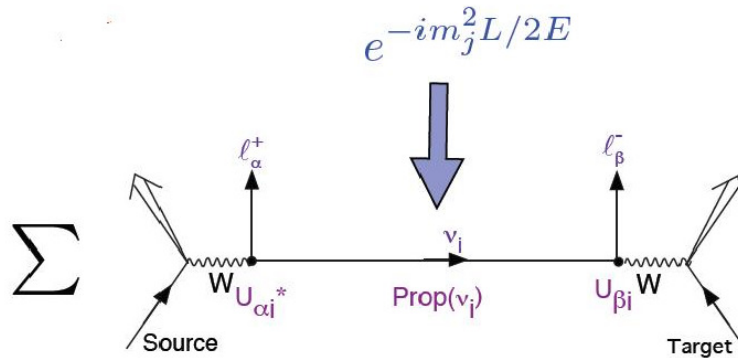
This transition from one flavour to the other,  $\nu_\alpha \rightarrow \nu_\beta$ , is a canonical case of the widely known quantum-mechanical effect present in a variety of two state systems and not a particular property of neutrinos.

Since, as shown clearly by Eq. (2), a  $\nu_\alpha$  is truly a coherent superposition of the three mass eigenstates  $\nu_i$ , the neutrino that travels since it is born until it is detected, can be any of the three  $\nu_i$ 's. Because of that, we should include the contributions of each of the  $\nu_i$  in a coherent way. As a consequence, the transition amplitude,  $\text{Amp}(\nu_\alpha \rightarrow \nu_\beta)$  receives a contribution of each  $\nu_i$  and turns out to be the product of three pieces. The first factor is the amplitude for the neutrino created at the generation point along with a charged lepton  $\bar{\ell}_\alpha$  to be, particularly, a  $\nu_i$ . And as we have said already, it is given by  $U_{\alpha i}^*$ .

The second component of our product is the amplitude for the  $\nu_i$  made by the source to cover the distance up to the detector. We will name this element  $\text{Prop}(\nu_i)$  for the time being and will postpone the calculation of its value until later. The last (third) piece is the amplitude for the charged lepton born out of the interaction of the neutrino  $\nu_i$  with the target to be, particularly, a  $\ell_\beta$ .

Being the Hamiltonian that describes the interactions between neutrinos, charged leptons and charged bosons  $W$  bosons hermitian (otherwise probability won't be conserved), it follows that if  $\text{Amp}(W \rightarrow \bar{\ell}_\alpha \nu_i) = U_{\alpha i}^*$ , then  $\text{Amp}(\nu_i \rightarrow \ell_\beta W) = U_{\beta i}$ . In this way, the third and last component of the product the  $\nu_i$  contribution is given by  $U_{\beta i}$ , and

$$\text{Amp}(\nu_\alpha \rightarrow \nu_\beta) = \sum_i U_{\alpha i}^* \text{Prop}(\nu_i) U_{\beta i} . \quad (4)$$



**Fig. 1:** Neutrino flavour change (oscillation) in vacuum

It still remains to be established the value of  $\text{Prop}(\nu_i)$ . To determine it, we'd better study the  $\nu_i$  in its rest frame. We will label the time in that system  $\tau_i$ . If  $\nu_i$  does have a rest mass  $m_i$ , then in this frame of reference its state vector satisfies the good old Schrödinger equation

$$i \frac{\partial}{\partial \tau_i} |\nu_i(\tau_i)\rangle = m_i |\nu_i(\tau_i)\rangle . \quad (5)$$

whose solution is given clearly by

$$|\nu_i(\tau_i)\rangle = e^{-im_i \tau_i} |\nu_i(0)\rangle . \quad (6)$$

Then, the amplitude for a given mass eigenstate  $\nu_i$  to travel freely during a time  $\tau_i$ , is simply the amplitude  $\langle \nu_i(0) | \nu_i(\tau_i) \rangle$  for observing the initial state  $\nu_i, |\nu_i(0)\rangle$  after some time as the evolved state  $|\nu_i(\tau_i)\rangle$ , *ie.*  $\exp[-im_i\tau_i]$ . Thus  $\text{Prop}(\nu_i)$  is only this amplitude where we have used that the time taken by  $\nu_i$  to cover the distance from the source to the detector is just  $\tau_i$ , the proper time.

Nevertheless, if we want  $\text{Prop}(\nu_i)$  to be of any use to us, we must write it first in terms of variables we can measure, this means to express it, in variables in the lab frame. The natural choice is obviously the distance,  $L$ , that the neutrino covers between the source and the detector as seen in the lab frame, and the time,  $t$ , that slips away during the journey, again in the lab frame. The distance  $L$  is set by the experimentalists through the selection of the place of settlement of the source and that of the detector and is unique to each experimental setting. Likewise, the value of  $t$  is selected by the experimentalists through their election for the time at which the neutrino is made and that when it dies (or gets detected). Therefore,  $L$  and  $t$  are determined (hopefully carefully enough) by the experiment design, and are the same for all the  $\nu_i$  in the beam. The different  $\nu_i$  do travel through an identical distance  $L$ , in an identical time  $t$ .

We still have two additional lab frame variables to determine, the energy  $E_i$  and three momentum  $p_i$  of the neutrino mass eigenstate  $\nu_i$ . By using the Lorentz invariance of the four component internal product (scalar product), we can obtain the expression for the  $m_i\tau_i$  appearing in the  $\nu_i$  propagator  $\text{Prop}(\nu_i)$  in terms of the (easy to measure) lab frame variable we have been looking for, which is given by

$$m_i\tau_i = E_it - p_iL . \quad (7)$$

At this point however one may argue that, in real life, neutrino sources are basically constant in time, and that the time  $t$  that slips away since the neutrino is produced till it dies in the detector is actually not measured. This argument is absolutely right. In reality, an experiment smears over the time  $t$  used by the neutrino to complete its route. However, lets consider that two constituents of the neutrino beam, the first one with energy  $E_1$  and the second one with energy  $E_2$  (both measured in the lab frame), add up coherently to the neutrino signal produced in the detector. Now, let us call  $t$  to the the time used by the neutrino to cover the distance separating the production and detection points. Then by the time the constituent whose energy is  $E_j$  ( $j = 1, 2$ ) arrives to the detector, it has raised a phase factor  $\exp[-iE_jt]$ . Therefore, we will have an interference between the  $E_1$  and  $E_2$  beam participants that will include a phase factor  $\exp[-i(E_1 - E_2)t]$ . When smeared over the non-observed travel time  $t$ , this factor goes away, *except when*  $E_2 = E_1$ . Therefore, only those constituents of the neutrino beam that share the same energy contribute coherently to the neutrino oscillation signal [3,4]. Specifically, only the different mass eigenstates constituents of the beams that have the same energy weight in. The rest gets averaged out.

Courtesy to its dispersion relation, a mass eigenstate  $\nu_i$ , with mass  $m_i$ , and energy  $E$ , has a three momentum  $p_i$  whose absolute value is given by

$$p_i = \sqrt{E^2 - m_i^2} \cong E - \frac{m_i^2}{2E} . \quad (8)$$

Where, we have utilised that as the masses of the neutrinos are miserably small,  $m_i^2 \ll E^2$  for a typical energy  $E$  attainable at any experiment (the lowest energy neutrinos have MeV energies and sub-eV masses). From Eqs. (7) and (8), it is easy to see that at a given energy  $E$  the phase  $m_i\tau_i$  appearing in  $\text{Prop}(\nu_i)$  takes the value

$$m_i\tau_i \cong E(t - L) + \frac{m_i^2}{2E}L . \quad (9)$$

As the phase  $E(t - L)$  appears in all the interfering terms it will eventually disappear when calculating the transition amplitude. After all is a common phase factor (its absolute value is one). Thus, we can get rid of it already now and use

$$\text{Prop}(\nu_i) = \exp[-im_i^2 \frac{L}{2E}] . \quad (10)$$



Plugging this into Eq. (4), we can obtain that the amplitude for a neutrino born as a  $\nu_\alpha$  to be detected as a  $\nu_\beta$  after covering a distance  $L$  with energy  $E$  yields

$$\text{Amp}(\nu_\alpha \longrightarrow \nu_\beta) = \sum_i U_{\alpha i}^* e^{-im_i^2 \frac{L}{2E}} U_{\beta i} . \quad (11)$$

The expression above is valid for an arbitrary number of neutrino flavours and an identical number of mass eigenstates, as far as they travel through vacuum. The probability  $P(\nu_\alpha \longrightarrow \nu_\beta)$  for  $\nu_\alpha \longrightarrow \nu_\beta$  can be found by squaring it, giving

$$\begin{aligned} P(\nu_\alpha \longrightarrow \nu_\beta) &= |\text{Amp}(\nu_\alpha \longrightarrow \nu_\beta)|^2 \\ &= \delta_{\alpha\beta} - 4 \sum_{i>j} \Re(U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^*) \sin^2 \left( \Delta m_{ij}^2 \frac{L}{4E} \right) \\ &\quad + 2 \sum_{i>j} \Im(U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^*) \sin \left( \Delta m_{ij}^2 \frac{L}{2E} \right) , \end{aligned} \quad (12)$$

with

$$\Delta m_{ij}^2 \equiv m_i^2 - m_j^2 . \quad (13)$$

In order to get Eq. (12) we have used that the mixing matrix  $U$  is unitary.

The oscillation probability  $P(\nu_\alpha \longrightarrow \nu_\beta)$  we have just obtained corresponds to that of a *neutrino*, and not to an *antineutrino*, as we have used that the oscillating neutrino was produced along with a charged *antilepton*  $\bar{\ell}$ , and gives birth to a charged *lepton*  $\ell$  once it reaches the detector. The corresponding probability  $P(\bar{\nu}_\alpha \longrightarrow \bar{\nu}_\beta)$  for an antineutrino oscillation can be obtained from  $P(\nu_\alpha \longrightarrow \nu_\beta)$  taking advantage of the fact that the two transitions  $\bar{\nu}_\alpha \longrightarrow \bar{\nu}_\beta$  and  $\nu_\beta \longrightarrow \nu_\alpha$  are CPT conjugated processes. Thus, assuming that neutrino interactions respect CPT [5],

$$P(\bar{\nu}_\alpha \longrightarrow \bar{\nu}_\beta) = P(\nu_\beta \longrightarrow \nu_\alpha) . \quad (14)$$

Then, from Eq. (12) we obtain that

$$P(\nu_\beta \longrightarrow \nu_\alpha; U) = P(\nu_\alpha \longrightarrow \nu_\beta; U^*) . \quad (15)$$

Therefore, if CPT is a good symmetry (as far as neutrino interactions are concerned), Eq. (12) tells us that

$$\begin{aligned} P(\bar{\nu}_\alpha \longrightarrow \bar{\nu}_\beta) &= \delta_{\alpha\beta} - 4 \sum_{i>j} \Re(U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^*) \sin^2 \left( \Delta m_{ij}^2 \frac{L}{4E} \right) \\ &\quad + (-) 2 \sum_{i>j} \Im(U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^*) \sin \left( \Delta m_{ij}^2 \frac{L}{2E} \right) . \end{aligned} \quad (16)$$

These expressions make it clear that if the mixing matrix  $U$  is complex,  $P(\bar{\nu}_\alpha \longrightarrow \bar{\nu}_\beta)$  and  $P(\nu_\alpha \longrightarrow \nu_\beta)$  will not be identical, in general. As  $\bar{\nu}_\alpha \longrightarrow \bar{\nu}_\beta$  and  $\nu_\alpha \longrightarrow \nu_\beta$  are CP conjugated processes,  $P(\bar{\nu}_\alpha \longrightarrow \bar{\nu}_\beta) \neq P(\nu_\alpha \longrightarrow \nu_\beta)$  would provide evidence of CP violation in neutrino oscillations (if Nature has chosen its mixing parameters so that the mixing matrix is indeed complex). Until now, CP violation has been observed only in the quark sector, so its measurement in neutrino physics would be quite exciting.

So far, we have been working in natural units. A fact that becomes transparent by looking at the dispersion relation Eq. (9). If we restore now the  $\hbar$ 's and  $c$  factors (we have happily set to one) into the oscillation probability we find that

$$\sin^2 \left( \Delta m_{ij}^2 \frac{L}{4E} \right) \longrightarrow \sin^2 \left( \Delta m_{ij}^2 c^4 \frac{L}{4\hbar c E} \right) \quad (17)$$

Having done that, it is easy and instructive to explore the semi-classical limit,  $\hbar \rightarrow 0$ . In this limit the oscillation length goes to zero (the oscillation phase goes to infinity) and the oscillations are averaged to 1/2. The interference pattern is lost. A similar situation appears if we let the mass difference  $\Delta m^2$  become large. This is exactly what happens in the quark sector (and the reason why we never study quark oscillations despite knowing that mass eigenstates do not coincide with flavour eigenstates).

In terms of real life units (which are not "natural" units), the oscillation phase is given by

$$\Delta m_{ij}^2 \frac{L}{4E} = 1.27 \Delta m_{ij}^2 (\text{eV}^2) \frac{L (\text{km})}{E (\text{GeV})} . \quad (18)$$

then, since  $\sin^2[1.27 \Delta m_{ij}^2 (\text{eV}^2) L (\text{km}) / E (\text{GeV})]$  can be experimentally observed (*ie.* not smeared out) only if its argument is in a ballpark around one, an experimental set-up with a baseline  $L$  (km) and an energy  $E$  (GeV) is sensitive to neutrino mass squared differences  $\Delta m_{ij}^2 (\text{eV}^2)$  of order  $\sim [L (\text{km}) / E (\text{GeV})]^{-1}$ . For example, an experiment with a baseline of  $L \sim 10^4$  km, roughly the size of Earth's diameter, and  $E \sim 1$  GeV would explore mass differences  $\Delta m_{ij}^2$  down to  $\sim 10^{-4}$  eV<sup>2</sup>. This fact makes it clear that neutrino long-baseline experiments can test even miserably small neutrino mass differences. It does so by exploiting the quantum mechanical interference between amplitudes whose relative phases are given precisely by these super tiny neutrino mass differences, which can be transformed into sizeable effects by choosing  $L/E$  appropriately.

But let's keep analysing the oscillation probability and see whether we can learn more about neutrino oscillations by studying its expression.

It is clear from  $P(\bar{\nu}_\alpha \rightarrow \bar{\nu}_\beta)$  that if neutrinos have zero mass, in such a way that all  $\Delta m_{ij}^2 = 0$ , then,  $P(\bar{\nu}_\alpha \rightarrow \bar{\nu}_\beta) = \delta_{\alpha\beta}$ . Therefore, the experimental observation that neutrinos can morph from one flavour to a different one indicates that neutrinos are not only massive but also that their masses are not degenerate. Actually, it was precisely this evidence the one that proved beyond any reasonable doubt that neutrinos are massive.

However, every neutrino oscillation seen so far has involved at some point neutrinos that travel through matter. But the expression we derived is valid only for flavour change in vacuum, and does not take into account any interaction between the neutrinos and the matter traversed between their source and their detector. Thus, the question remains whether it may be that some unknown flavour changing interactions between neutrinos and matter are indeed responsible of the observed flavour transitions, and not neutrino masses. Regarding this question, a couple of things should be said. First, although it is true that the Standard Model of elementary particle physics contains only massless neutrinos, it provides an amazingly well corroborated description of weak interactions, and therefore of all the ways a neutrino interacts. Such a description does not include flavour change. Second, for some of the processes experimentally observed where neutrinos do change flavour, matter effects are expected to be miserably small, and on those cases the evidence points towards a dependence on  $L$  and  $E$  in the flavour transition probability through the combination  $L/E$ , as anticipated by the oscillation hypothesis. Modulo a constant,  $L/E$  is precisely the proper time that goes by in the rest frame of the neutrino as it covers a distance  $L$  possessing an energy  $E$ . Therefore, these flavour transitions behave as if they were a true progression of the neutrino itself over time, and not a result of an interaction with matter.

Now, lets explore the case where the leptonic mixing were trivial. This would imply that in the charged boson decay  $W^+ \rightarrow \bar{\ell}_\alpha + \nu_i$ , which as we established has an amplitude  $U_{\alpha i}^*$ , the emerging charged antilepton  $\bar{\ell}_\alpha$  of flavour  $\alpha$  comes along always with the *same* neutrino mass eigenstate  $\nu_i$ . That is, if  $U_{\alpha i}^* \neq 0$ , then due to unitarity,  $U_{\alpha j}$  becomes zero for all  $j \neq i$ . Therefore, from Eq. (16) it is clear that,  $P(\bar{\nu}_\alpha \rightarrow \bar{\nu}_\beta) = \delta_{\alpha\beta}$ . Thus, the observation that neutrinos morph indicates non trivial a mixing matrix.

Then, we are left with basically two ways to detect neutrino flavour change. The first one is to observe, in a beam of neutrinos which are all created with the same flavour, say  $\alpha$ , some amount of neutrinos of a new flavour  $\beta$  that is different from the flavour  $\alpha$  we started with. This goes under the

name of appearance experiments. The second way is to start with a beam of identical  $\nu_\alpha$ s, whose flux is either measured or known, and observe that after travelling some distance this flux is depleted. Such experiments are called disappearance experiments.

As Eq. (16) shows, the transition probability in vacuum does not only depend on  $L/E$  but also oscillates with it. It is because of this fact that neutrino flavour transitions are named “neutrino oscillations”. Now notice also that neutrino transition probabilities do not depend on the individual neutrino masses (or masses squared) but on the squared-mass *differences*. Thus, oscillation experiments can only measure the neutrino mass squared spectrum. Not its absolute scale. Experiments can test the pattern but cannot determine the distance above zero the whole spectra lies.

It is clear that neutrino transitions cannot modify the total flux in a neutrino beam, but simply alter its distribution between the different flavours. Actually, from Eq. (16) and the unitarity of the  $U$  matrix, it is obvious that

$$\sum_{\beta} P(\bar{\nu}_\alpha \longrightarrow \bar{\nu}_\beta) = 1 \quad , \quad (19)$$

where the sum runs over all flavours  $\beta$ , including the original one  $\alpha$ . Eq. (19) makes it transparent that the probability that a neutrino morphs its flavour, added to the probability that it keeps the flavour it had at birth, is one. Ergo, flavour transitions do not modify the total flux. Nevertheless, some of the flavours  $\beta \neq \alpha$  into which a neutrino can oscillate into may be *sterile* flavours; that is, flavours that do not take part in weak interactions and therefore escape detection. If any of the original (active) neutrino flux turns into sterile, then an experiment able to measure the total *active* neutrino flux—that is, the flux associated to those neutrinos that couple to the weak gauge bosons:  $\nu_e$ ,  $\nu_\mu$ , and  $\nu_\tau$ —will observe it to be not exactly the original one, but smaller than it. In the experiments performed up today, no flux was ever missed.

In the literature, description of neutrino oscillations normally assume that the different mass eigenstates  $\nu_i$  that contribute coherently to a beam share the same *momentum*, rather than the same *energy* as we have argued they must have. While the supposition of equal momentum is technically wrong, it is an inoffensive mistake, since, as can easily be shown, it conveys to the same oscillation probabilities as the ones we have obtained.

A relevant and interesting case of the (not that simple) formula for  $P(\bar{\nu}_\alpha \longrightarrow \bar{\nu}_\beta)$  is the case where only two flavours participate in the oscillation. The only-two-neutrino scenario is a rather rigorous description of a vast number of experiments. In fact only recently (and in few experiments) a more sophisticated (three neutrino description) was needed to fit observations. Lets assume then, that only two mass eigenstates, which we will name  $\nu_1$  and  $\nu_2$ , and two reciprocal flavour states, which we will name  $\nu_\mu$  and  $\nu_\tau$ , are relevant, in such a way that only one squared-mass difference,  $m_2^2 - m_1^2 \equiv \Delta m^2$  arises. Even more, neglecting phase factors that can be proven to have no impact on oscillation probabilities, the mixing matrix  $U$  can be written as

$$\begin{pmatrix} \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} \quad (20)$$

The unitary mixing matrix  $U$  of Eq. (20) is just a  $2 \times 2$  rotation matrix, and as such, parameterized by a single rotation angle  $\theta$  which is named (in neutrino physics) as the mixing angle. Plugging the  $U$  of Eq. (20) and the unique  $\Delta m^2$  into the general formula of the transition probability  $P(\bar{\nu}_\alpha \longrightarrow \bar{\nu}_\beta)$ , Eq. (16), we can readily see that, for  $\beta \neq \alpha$ , when only two neutrinos are relevant,

$$P(\bar{\nu}_\alpha \longrightarrow \bar{\nu}_\beta) = \sin^2 2\theta \sin^2 \left( \frac{\Delta m^2 L}{4E} \right) \quad . \quad (21)$$

Moreover, the survival probability, *ie.* the probability that the neutrino remains with the same flavour its was created with is, as expected, one minus the probability that it changes flavour.

### 3 Neutrino Oscillations in a medium

When we create a beam of neutrinos on earth through an accelerator and send it up to thousand kilometres away to a meet detector, the beam does not move through vacuum, but through matter, earth matter. The beam of neutrinos then scatters from the particles it meets along the way. Such a coherent forward scattering can have a large effect on the transition probabilities. We will assume for the time being that neutrino interactions with matter are flavour conserving, as described by the Standard Model, and comment on the possibility of flavour changing interactions later. Then as there are only two types of weak interactions (mediated by charged and neutral currents) the would be accordingly only two possibilities for this coherent forward scattering from matter particles to take place. Charged current mediated weak interactions will occur only if and only if the incoming neutrino is an electron neutrino. As only the  $\nu_e$  can exchange charged boson  $W$  with an Earth electron. Thus neutrino-electron coherent forward scattering via  $W$  exchange opens up an extra source of interaction energy  $V_W$  suffered exclusively by electron neutrinos. Obviously, this additional energy being from weak interactions origin has to be proportional to  $G_F$ , the Fermi coupling constant. In addition, the interaction energy coming from  $\nu_e - e$  scattering grows with the number of targets,  $N_e$ , the number of electrons per unit volume (given by the density of the Earth). Putting everything together it is not difficult to see that

$$V_W = +\sqrt{2} G_F N_e , \quad (22)$$

clearly, this interaction energy affects also antineutrinos (in a opposite way though). It changes sign if we switch the  $\nu_e$  by  $\bar{\nu}_e$ .

The interactions mediated by neutral currents correspond to the case where a neutrino in matter interacts with a matter electron, proton, or neutron by exchanging a neutral  $Z$  boson. According to the Standard Model weak interactions are flavour blind. Every flavour of neutrino enjoys them, and the amplitude for this  $Z$  exchange is always the same. It also teaches us that, at zero momentum transfer, electrons and protons couple to the  $Z$  boson with equal strength. The interaction has though, opposite sign. Therefore, counting on the fact that the matter through which our neutrino moves is electrically neutral (it contains equal number of electrons and protons), the contribution of both, electrons and protons to coherent forward neutrino scattering through  $Z$  exchange will add up to zero. Consequently only interactions with neutrons will survive so that, the effect of the  $Z$  exchange contribution to the interaction potential energy  $V_Z$  reduces exclusively to that with neutrons and will be proportional to  $N_n$ , the number density of neutrons. It goes without saying that it will be equal to all flavours. This time, we find that

$$V_Z = -\frac{\sqrt{2}}{2} G_F N_n , \quad (23)$$

as was the case before, for  $V_W$ , this contribution will flip sign if we replace the neutrinos by anti-neutrinos.

But if, as we said, the Standard Model interactions do not change neutrino flavour, neutrino flavour transitions or neutrino oscillations point undoubtedly to neutrino mass and mixing even when neutrinos are propagating through matter. Unless non-Standard-Model flavour changing interactions play a role.

Neutrino propagation in matter is easy to understand when analysed through the time dependent Schrödinger equation in the lab frame

$$i \frac{\partial}{\partial t} |\nu(t)\rangle = \mathcal{H} |\nu(t)\rangle . \quad (24)$$

where,  $|\nu(t)\rangle$  is a (three component) neutrino vector state, in which each neutrino flavour corresponds to one component. In the same way, the Hamiltonian  $\mathcal{H}$  is a (three  $\times$  three) matrix in flavour space. To make our lives easy, lets analyse the case where only two neutrino flavours are relevant, say  $\nu_e$  and  $\nu_\mu$ . Then

$$|\nu(t)\rangle = \begin{pmatrix} f_e(t) \\ f_\mu(t) \end{pmatrix} , \quad (25)$$

with  $f_i(t)^2$  the amplitude of the neutrino to be a  $\nu_i$  at time  $t$ . This time the Hamiltonian,  $\mathcal{H}$ , is a  $2 \times 2$  matrix in neutrino flavour space, i.e.,  $\nu_e - \nu_\mu$  space.

It will prove to be clarifying to work out the two flavour case in vacuum first, and add matter effects afterwards. Using Eq. (2) to express  $|\nu_\alpha\rangle$  as a linear combination of mass eigenstates, we can see that the  $\nu_\alpha - \nu_\beta$  matrix element of the Hamiltonian in vacuum,  $\mathcal{H}_{\text{Vac}}$ , can be written as

$$\begin{aligned} \langle \nu_\alpha | \mathcal{H}_{\text{Vac}} | \nu_\beta \rangle &= \langle \sum_i U_{\alpha i}^* \nu_i | \mathcal{H}_{\text{Vac}} | \sum_j U_{\beta j}^* \nu_j \rangle \\ &= \sum_j U_{\alpha j} U_{\beta j}^* \sqrt{p^2 + m_j^2} . \end{aligned} \quad (26)$$

where we are supposing that the neutrinos belong to a beam where all its mass components (the mass eigenstates) share the same definite momentum  $p$ . As we have already mentioned, despite this supposition being technically wrong, it leads anyway to the right transition amplitude. In the second line of Eq. (26), we have used that the neutrinos  $\nu_j$  with momentum  $p$ , the mass eigenstates, are the asymptotic states of the hamiltonian,  $\mathcal{H}_{\text{Vac}}$  for which constitute an orthonormal basis, satisfy

$$\mathcal{H}_{\text{Vac}} | \nu_j \rangle = E_j | \nu_j \rangle \quad (27)$$

and have the standard dispersion relation,  $E_j = \sqrt{p^2 + m_j^2}$ .

As we have already mentioned, neutrino oscillations are the archetype quantum interference phenomenon, where only the *relative* phases of the interfering states play a role. Therefore, only the *relative* energies of these states, which set their relative phases, are relevant. As a consequence, if it proves to be convenient (and it will), we can feel free to happily remove from the Hamiltonian  $\mathcal{H}$  any contribution proportional to the identity matrix  $I$ . As we have said, this subtraction will leave unaffected the differences between the eigenvalues of  $\mathcal{H}$ , and therefore will leave unaffected the prediction of  $\mathcal{H}$  for flavour transitions.

It goes without saying that as in this case only two neutrinos are relevant, there are only two mass eigenstates,  $\nu_1$  and  $\nu_2$ , and only one mass splitting  $\Delta m^2 \equiv m_2^2 - m_1^2$ , and therefore there should be, as before a unitary  $U$  matrix given by Eq. (20) which rotates from one basis to the other. Inserting it into Eq. (26), and assuming that our neutrinos have low masses as compared to their momenta, i.e.,  $(p^2 + m_j^2)^{1/2} \cong p + m_j^2/2p$ , and removing from  $\mathcal{H}_{\text{Vac}}$  a term proportional to the the identity matrix (a removal we know is going to be harmless), we get

$$\mathcal{H}_{\text{Vac}} = \frac{\Delta m^2}{4E} \begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix} . \quad (28)$$

To write this expression, the highly relativistic approximation, which says that  $p \cong E$  is used. Where  $E$  is the average energy of the neutrino mass eigenstates in our neutrino beam of ultra high momentum  $p$ .

It is not difficult to corroborate that the Hamiltonian  $\mathcal{H}_{\text{Vac}}$  of Eq. (28) for the two neutrino scenario would give an identical oscillation probability, Eq. (21), as the one we have already obtained in a different way. An easy way to do it is to analyse the transition probability for the process  $\nu_e \rightarrow \nu_\mu$ . From Eq. (20) it is clear that in terms of the mixing angle, the electron neutrino state composition is

$$|\nu_e\rangle = |\nu_1\rangle \cos \theta + |\nu_2\rangle \sin \theta , \quad (29)$$

while that of the muon neutrino is given by

$$|\nu_\mu\rangle = -|\nu_1\rangle \sin \theta + |\nu_2\rangle \cos \theta . \quad (30)$$

In the same way, we can also write the eigenvalues of the vacuum hamiltonian  $\mathcal{H}_{\text{Vac}}$ , Eq.25, in terms of the mass squared differences as

$$\lambda_1 = -\frac{\Delta m^2}{4E} , \quad \lambda_2 = +\frac{\Delta m^2}{4E} . \quad (31)$$

The mass eigenbasis of this Hamiltonian,  $|\nu_1\rangle$  and  $|\nu_2\rangle$ , can also be written in terms of flavour eigenbasis  $|\nu_e\rangle$  and  $|\nu_\mu\rangle$  by means of Eqs. (29) and (30). Therefore, the Schrödinger equation of Eq. (24), with the identification of  $\mathcal{H}$  in this case with  $\mathcal{H}_{\text{Vac}}$  tells us that if at time  $t = 0$  we begin from a  $|\nu_e\rangle$ , then once some time  $t$  elapses this  $|\nu_e\rangle$  will progress into the state given by

$$|\nu(t)\rangle = |\nu_1\rangle e^{+i\frac{\Delta m^2}{4E}t} \cos\theta + |\nu_2\rangle e^{-i\frac{\Delta m^2}{4E}t} \sin\theta . \quad (32)$$

Thus, the probability  $P(\nu_e \rightarrow \nu_\mu)$  that this evolved neutrino be detected as a different flavour  $\nu_\mu$ , from Eqs. (30) and (32), is given by,

$$\begin{aligned} P(\nu_e \rightarrow \nu_\mu) &= |\langle \nu_\mu | \nu(t) \rangle|^2 \\ &= |\sin\theta \cos\theta (-e^{i\frac{\Delta m^2}{4E}t} + e^{-i\frac{\Delta m^2}{4E}t})|^2 \\ &= \sin^2 2\theta \sin^2 \left( \Delta m^2 \frac{L}{4E} \right) . \end{aligned} \quad (33)$$

Where we have substituted the time  $t$  travelled by our highly relativistic state by the distance  $L$  it has covered. The flavour transition or oscillation probability of Eq. (33), as expected, is exactly the same we have found before, Eq. (21).

We can now move on to analyse neutrino propagation in matter. In this case, the  $2 \times 2$  Hamiltonian representing the propagation in vacuum  $\mathcal{H}_{\text{Vac}}$  receives the two additional contributions we have discussed before, and becomes  $\mathcal{H}_M$ , which is given by

$$\mathcal{H}_M = \mathcal{H}_{\text{Vac}} + V_W \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + V_Z \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \quad (34)$$

In the new Hamiltonian, the first additional contribution corresponds to the interaction potential due to the charged bosons exchange, Eq. (22). As this interaction is suffered only by  $\nu_e$ , this contribution is different from zero only in the  $\mathcal{H}_M(1,1)$  element or the  $\nu_e - \nu_e$  element. The second additional contribution, the last term of Eq. (34) comes from the  $Z$  boson exchange, Eq. (23). Since this interaction is flavour blind, it affects every neutrino flavour in the same way, its contribution to  $\mathcal{H}_M$  is proportional to the identity matrix, and can be safely neglected. Thus

$$\mathcal{H}_M = \mathcal{H}_{\text{Vac}} + \frac{V_W}{2} + \frac{V_W}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} , \quad (35)$$

where (for reasons that are going to become clear later) we have divided the  $W$ -exchange contribution into two pieces, one proportional to the identity (that we will disregard in the next step) and, a piece that it is not proportional to the identity, that we will keep. Disregarding the first piece as promised, we have from Eqs. (28) and (35)

$$\mathcal{H}_M = \frac{\Delta m^2}{4E} \begin{pmatrix} -(\cos 2\theta - A) & \sin 2\theta \\ \sin 2\theta & (\cos 2\theta - A) \end{pmatrix} , \quad (36)$$

where we have defined

$$A \equiv \frac{V_W/2}{\Delta m^2/4E} = \frac{2\sqrt{2}G_F N_e E}{\Delta m^2} . \quad (37)$$

Clearly,  $A$  parameterizes the relative size of the matter effects as compared to the vacuum contribution given by the neutrino squared-mass splitting and signals the situations when they become important.

Now, if we introduce (a physically meaningful) short-hand notation

$$\Delta m_M^2 \equiv \Delta m^2 \sqrt{\sin^2 2\theta + (\cos 2\theta - A)^2} \quad (38)$$

and

$$\sin^2 2\theta^M \equiv \frac{\sin^2 2\theta}{\sin^2 2\theta + (\cos 2\theta - A)^2} , \quad (39)$$

then the Hamiltonian in a medium  $\mathcal{H}_M$  turns out to be

$$\mathcal{H}_M = \frac{\Delta m_M^2}{4E} \begin{pmatrix} -\cos 2\theta^M & \sin 2\theta^M \\ \sin 2\theta^M & \cos 2\theta^M \end{pmatrix} . \quad (40)$$

and can be diagonalised by inspection, i.e., as a result of our choice, the Hamiltonian in a medium,  $\mathcal{H}_M$ , becomes formally indistinguishable to the vacuum one,  $\mathcal{H}_{\text{vac}}$ , Eq. (28). The difference being that in this case what used to be the vacuum parameters  $\Delta m^2$  and  $\theta$  are presently given by the matter ones,  $\Delta m_M^2$  and  $\theta^M$ , respectively.

Obviously, the mass eigenstates and eigenvalues (which determine the mixing angle) of  $\mathcal{H}_M$  are not identical to the ones in vacuum. The mass squared difference of the matter eigenstates is not the same as the vacuum  $\Delta m^2$ , and the same happens with the mixing angle. The eigenstates in matter, *ie.* the files of the unitary matrix that rotates from the flavour basis to the mass basis, are different from the vacuum eigenvalues that form the vacuum mixing matrix, and therefore  $\theta_M$  is not  $\theta$ . But, the matter Hamiltonian  $\mathcal{H}_M$  does indeed contain all about the propagation of neutrinos in matter, in the same way  $\mathcal{H}_{\text{vac}}$  contains all about the propagation in vacuum.

According to Eq. (40),  $\mathcal{H}_M$  has the same functional dependence on the matter parameters  $\Delta m_M^2$  and  $\theta^M$  as the vacuum Hamiltonian  $\mathcal{H}_{\text{vac}}$ , Eq. (28), on the vacuum ones,  $\Delta m^2$  and  $\theta$ . Therefore,  $\Delta m_M^2$  can be identified with an effective mass squared difference in matter, and accordingly  $\theta^M$  can be unidentifed with an effective mixing angle in matter.

In a typical experimental set-up where the neutrino beam is generated by an accelerator and sent away to a detector that is, say, several hundred, or even thousand kilometres away, it traverses through earth matter, but only superficially, it does not get deep into the earth. Then, during this voyage the matter density encountered by such a beam can be taken to be approximately constant<sup>1</sup>. But if the density of the earth's matter is constant, the same happens with the electron density  $N_e$ , and the  $A$  parameter in which it is incorporated, which after all is determined by it. And it is also true about the Hamiltonian  $\mathcal{H}_M$ . They all become approximately constant, and therefore quite identical to the vacuum Hamiltonian  $\mathcal{H}_{\text{vac}}$ , except for the particular values of their parameters. By comparing Eqs. (40) and (28), we can immediately conclude that exactly in the same way  $\mathcal{H}_{\text{vac}}$  gives rise to vacuum oscillations with probability  $P(\nu_e \rightarrow \nu_\mu)$  of Eq. (33),  $\mathcal{H}_M$  must give rise to matter oscillations with probability

$$P_M(\nu_e \rightarrow \nu_\mu) = \sin^2 2\theta^M \sin^2 \left( \Delta m_M^2 \frac{L}{4E} \right) . \quad (41)$$

Namely, the transition and survival probabilities in matter are the same as those in vacuum, except that the vacuum parameters  $\Delta m^2$  and  $\theta$  are now replaced by their matter counterparts,  $\Delta m_M^2$  and  $\theta^M$ .

In theory, judging simply by its potential, matter effects can have very drastic repercussions in the oscillation probabilities. The exact impact (if any) can be estimated only after the details of the experimental set-up of the experiment in question are given. As a rule of thumb, to guess the importance of matter effects, we should keep in mind that for neutrinos propagating through the earth's mantle (not deeper than 200 km below the surface) and if the kinematic phase associated to the solar mass difference is still negligible,

$$A \cong \frac{E}{13 \text{ GeV}} \quad (42)$$

so that only for beam energies of several GeV matter effects do matter.

<sup>1</sup>This approximation is clearly not valid for neutrinos that cross the Earth

And how much do they matter? They matter a lot! From Eq. (39) for the matter mixing angle,  $\theta^M$ , we can appreciate that even when the vacuum mixing angle  $\theta$  is incredible small, say,  $\sin^2 2\theta = 10^{-4}$ , if we get to have  $A \cong \cos 2\theta$ , i.e., for energies of a few tens of GeV, then  $\sin^2 2\theta^M$  can be brutally enhanced as compared to its vacuum value and can even reach maximal mixing, i.e.  $\sin^2 2\theta^M = 1$ . This wild enhancement of a small mixing angle in vacuum up to a sizeable (even maximal) one in matter is the “resonant” enhancement, the largest possible version of the Mikheyev-Smirnov-Wolfenstein effect [6–9]. In the beginning of solar neutrino experiments, people entertained the idea that this brutal enhancement was actually taking place while neutrinos crossed the sun. Nonetheless, as we will see soon the mixing angle associated with solar neutrinos is quite sizeable ( $\sim 34^\circ$ ) already in vacuum [10]. Then, although matter effects on the sun are important and they do enhance the solar mixing angle, unfortunately they are not as drastic as we once dreamt. Nevertheless, for long-baselines they will play (they are already playing!) a key role in the determination of the ordering of the neutrino spectrum.

## 4 Evidence for neutrino oscillations

### 4.1 Atmospheric and Accelerator Neutrinos

Almost twenty years have elapsed since we were presented solid and convincing evidence of neutrino masses and mixings, and since then, the evidence has only grown. SuperKamiokande (SK) was the first experiment to present compelling evidence of  $\nu_\mu$  disappearance in their atmospheric neutrino fluxes, see [11]. In Fig. 2 the zenith angle (the angle subtended with the horizontal) dependence of the multi-GeV  $\nu_\mu$  sample is shown together with the disappearance as a function of  $L/E$  plot. These data fit amazingly well the naive two component neutrino hypothesis with

$$\Delta m_{\text{atm}}^2 = 2 - 3 \times 10^{-3} \text{eV}^2 \quad \text{and} \quad \sin^2 \theta_{\text{atm}} = 0.50 \pm 0.13 \quad (43)$$

Roughly speaking SK corresponds to an  $L/E$  for oscillations of 500 km/GeV and almost maximal mixing (the mass eigenstates are nearly even admixtures of muon and tau neutrinos). No signal of an involvement of the third flavour,  $\nu_e$  is found so the assumption is that atmospheric neutrino disappearance is basically  $\nu_\mu \rightarrow \nu_\tau$ . Notice however, that the first NOvA results seem to point toward a mixing angle which is not maximal (excluding maximal mixing at the 2 sigma level).

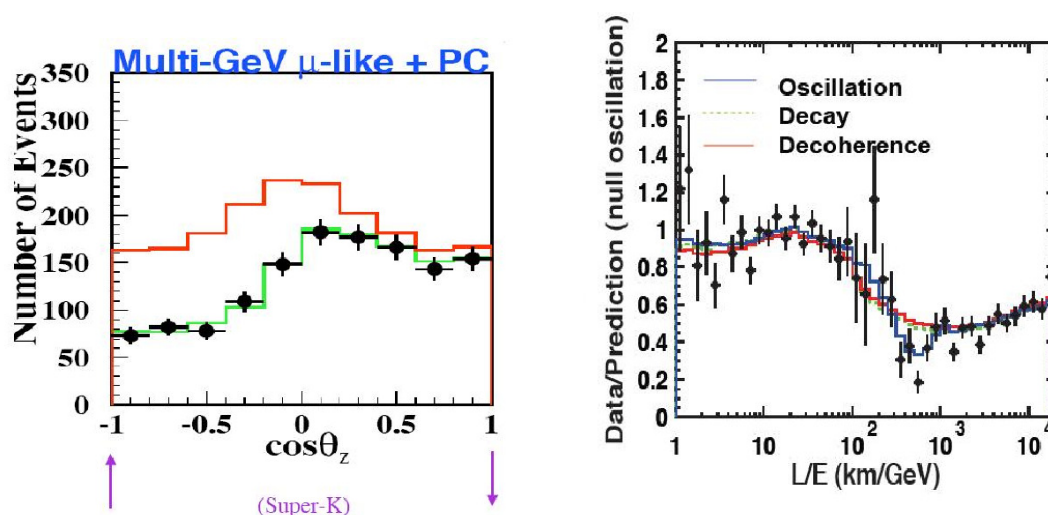
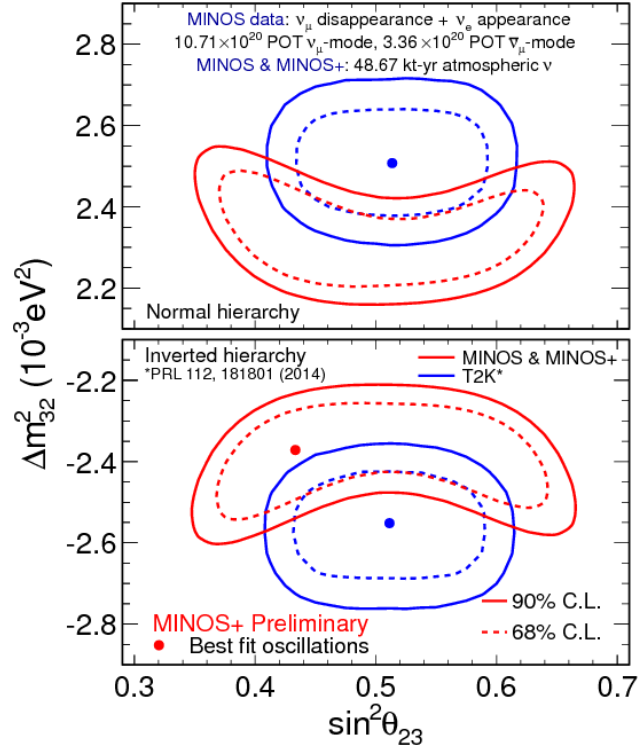


Fig. 2: Superkamiokande’s evidence for neutrino oscillations both in the zenith angle and  $L/E$  plots

After atmospheric neutrino oscillations were established, a new series of neutrino experiments were built, sending (man-made) beams of  $\nu_\mu$  neutrinos to detectors located at large distances: the K2K



(T2K) experiment [12,13], sends neutrinos from the KEK accelerator complex to the old SK mine, with a baseline of 120 (235) km while the MINOS (NOvA) experiment [14,15], sends its beam from Fermilab, near Chicago, to the Soudan mine (Ash river) in Minnesota, a baseline of 735 (810) km. All these experiments have seen evidence for  $\nu_\mu$  disappearance consistent with the one found by SK. Their results are summarised in Fig. 3.



**Fig. 3:** Allowed regions in the  $\Delta m_{\text{atm}}^2$  vs  $\sin^2 \theta_{\text{atm}}$  plane for MINOS data as well as for T2K data and two of the SK analyses. MINOS's best fit point is at  $\sin^2 \theta_{\text{atm}} = .51$  and  $\Delta m_{\text{atm}}^2 = 2.37 \times 10^{-3} \text{eV}^2$ . Notice that new NOvA data seem to exclude maximal mixing at the 2 sigma level

## 4.2 Reactor and Solar Neutrinos

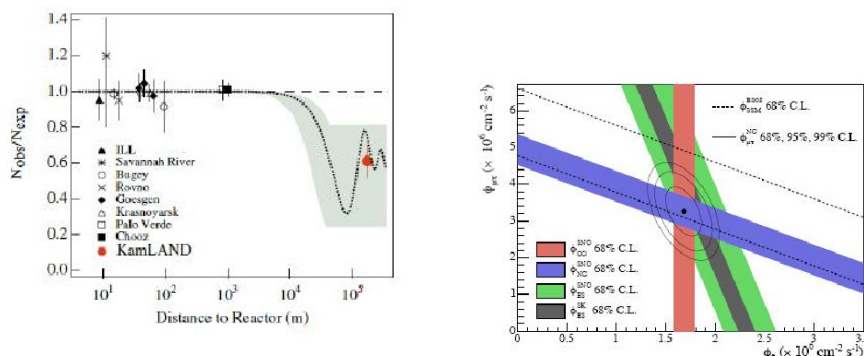
The KamLAND reactor experiment, an antineutrino disappearance experiment, receiving neutrinos from sixteen different reactors, at distances ranging from hundred to thousand kilometres, with an average baseline of 180 km and neutrinos of a few eV, [16, 17], has seen evidence of neutrino oscillations. Such evidence was collected not only at a different  $L/E$  than the atmospheric and accelerator experiments but also consists on oscillations involving electron neutrinos,  $\nu_e$ , the ones which were not involved before. These oscillations have also been seen for neutrinos coming from the sun (the sun produces only electron neutrinos). However, in order to compare the two experiments we should assume that neutrinos (solar) and antineutrinos (reactor) behave in the same way, *ie.* assume CPT conservation. The best fit values in

the two neutrino scenario for the KamLAND experiment are

$$\Delta m_{\odot}^2 = 8.0 \pm 0.4 \times 10^{-5} \text{eV}^2 \quad \text{and} \quad \sin^2 \theta_{\odot} = 0.31 \pm 0.03 \quad (44)$$

In this case, the  $L/E$  involved is 15 km/MeV which is more than an order of magnitude larger than the atmospheric scale and the mixing angle, although large, is clearly not maximal.

Fig. 4 shows the disappearance probability for the  $\bar{\nu}_e$  for KamLAND as well as several older reactor experiments with shorter baselines<sup>2</sup>. The second panel depicts the flavour content of the  $^8\text{B}$  Boron solar neutrino flux (with GeV energies) measured by SNO, [18], and SK, [19]. The reactor outcome can be explained in terms of two flavour oscillations in vacuum, given that the fit to the disappearance probability, is appropriately averaged over  $E$  and  $L$ .



**Fig. 4:** Disappearance of the  $\bar{\nu}_e$  observed by reactor experiments as a function of distance from the reactor. The flavour content of the  $^8\text{B}$  Boron solar neutrinos for the various reactions for SNO and SK. CC:  $\nu_e + d \rightarrow e^- + p + p$ , NC:  $\nu_x + d \rightarrow \nu_x + p + n$  and ES:  $\nu_\alpha + e^- \rightarrow \nu_\alpha + e^-$

The analysis of neutrinos originating from the sun is marginally more complex than the one we did before because it should incorporate the matter effects that the neutrinos endure since they are born (at the centre of the sun) until they abandon it, which are imperative at least for the  $^8\text{B}$  Boron neutrinos. The pp and  $^7\text{Be}$  neutrinos are less energetic and therefore are not significantly altered by the presence of matter and leave the sun as though it were ethereal.  $^8\text{B}$  Boron neutrinos on the other hand, leave the sun unequivocally influenced by the presence of matter and this is evidenced by the fact that they leave the sun as  $\nu_2$ , the second mass eigenstate and therefore do not experience oscillations. This distinction among neutrinos coming from different reaction chains is, as mentioned, due mainly to their disparities at birth. While pp ( $^7\text{Be}$ ) neutrinos are created with an average energy of 0.2 MeV (0.9 MeV),  $^8\text{B}$  are born with 10 MeV and as we have seen the impact of matter effects grows with the energy of the neutrino.

However, we ought to emphasise that we do not really see solar neutrino oscillations. To trace the oscillation pattern, to be able to test its distinctive shape, we need a kinematic phase of order one otherwise the oscillations either do not develop or get averaged to 1/2. In the case of neutrinos coming from the sun the kinematic phase is

$$\Delta_{\odot} = \frac{\Delta m_{\odot}^2 L}{4E} = 10^{7 \pm 1}. \quad (45)$$

Consequently, solar neutrinos behave as "effectively incoherent" mass eigenstates once they leave the sun, and remain so once they reach the earth. Consequently the  $\nu_e$  disappearance or survival probability

<sup>2</sup>Shorter baseline reactor neutrino experiments, which has seen no evidence of flux depletion suffer the so-called reactor neutrino anomaly, which may point toward the existence of light sterile states

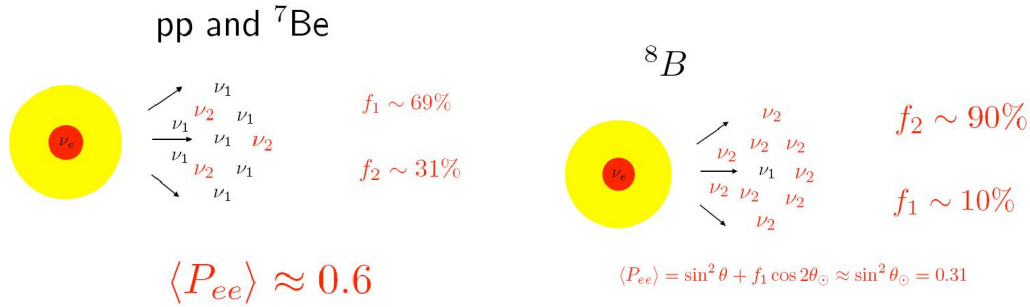
is given by

$$\langle P_{ee} \rangle = f_1 \cos^2 \theta_\odot + f_2 \sin^2 \theta_\odot \quad (46)$$

where  $f_1$  is the  $\nu_1$  content or fraction of  $\nu_\mu$  and  $f_2$  is the  $\nu_2$  content of  $\nu_\mu$  and therefore both fractions satisfy

$$f_1 + f_2 = 1. \quad (47)$$

Nevertheless, as we have already mentioned, solar neutrinos originating from the pp and  ${}^7\text{Be}$  chains are not affected by the solar matter and oscillate as in vacuum and thus, in their case  $f_1 \approx \cos^2 \theta_\odot = 0.69$  and  $f_2 \approx \sin^2 \theta_\odot = 0.31$ . In the  ${}^8\text{B}$  a neutrino case, however, the impact of solar matter is sizeable and the corresponding fractions are substantially altered, see Fig. 5.



**Fig. 5:** The sun produces  $\nu_e$  in the core but once they exit the sun thinking about them in the mass eigenstate basis is useful. The fraction of  $\nu_1$  and  $\nu_2$  is energy dependent above 1 MeV and has a dramatic effect on the  ${}^8\text{B}$  solar neutrinos, as first observed by Davis.

In a two neutrino scenario, the day-time CC/NC measured by SNO, which is roughly identical to the day-time average  $\nu_e$  survival probability,  $\langle P_{ee} \rangle$ , reads

$$\left. \frac{CC}{NC} \right|_{\text{day}} = \langle P_{ee} \rangle = f_1 \cos^2 \theta_\odot + f_2 \sin^2 \theta_\odot, \quad (48)$$

where  $f_1$  and  $f_2 = 1 - f_1$  are the  $\nu_1$  and  $\nu_2$  contents of the muon neutrino, respectively, averaged over the  ${}^8\text{B}$  neutrino energy spectrum appropriately weighted with the charged current current cross section. Therefore, the  $\nu_1$  fraction (or how much  $f_2$  differs from 100%) is given by

$$f_1 = \frac{\left( \left. \frac{CC}{NC} \right|_{\text{day}} - \sin^2 \theta_\odot \right)}{\cos 2\theta_\odot} = \frac{(0.347 - 0.311)}{0.378} \approx 10\% \quad (49)$$

where the central values of the last SNO analysis, [18], were used. As there are strong correlations between the uncertainties of the CC/NC ratio and  $\sin^2 \theta_\odot$  it is not obvious how to estimate the uncertainty on  $f_1$  from their analysis. Note, that if the fraction of  $\nu_2$  were 100%, then  $\left. \frac{CC}{NC} \right|_{\text{day}} = \sin^2 \theta_\odot$ .

Utilising the analytic analysis of the Mikheyev-Smirnov-Wolfenstein (MSW) effect, gave in [20], one can obtain the mass eigenstate fractions in a medium, which are given by

$$f_2 = 1 - f_1 = \langle \sin^2 \theta_\odot^M + P_x \cos 2\theta_\odot^M \rangle_{s\text{B}}, \quad (50)$$

with  $\theta_\odot^M$  being the mixing angle as given at the  $\nu_e$  production point and  $P_x$  is the probability of the neutrino to hop from one mass eigenstate to the second one during the Mikheyev-Smirnov resonance

crossing. The average  $\langle \dots \rangle_{\text{SBN}}$  is over the electron density of the  ${}^8\text{B}$   $\nu_e$  production region in the centre of the Sun as given by the Solar Standard Model and the energy spectrum of  ${}^8\text{B}$  neutrinos has been appropriately weighted with SNO's charged current cross section. All in all, the  ${}^8\text{B}$  energy weighted average content of  $\nu_2$ 's measured by SNO is

$$f_2 = 91 \pm 2\% \text{ at the } 95\% \text{ C.L.} \quad (51)$$

Therefore, it is obvious that the  ${}^8\text{B}$  solar neutrinos are the purest mass eigenstate neutrino beam known so far and SK super famous picture of the sun taken (from underground) with neutrinos is made with approximately 90% of  $\nu_2$ , *ie.* almost a pure beam of mass eigenstates.

On March 8, 2012 a newly built reactor neutrino experiment, the Daya Bay experiment, located in China, announced the measurement of the third mixing angle [21], the only one which was still missing and found it to be

$$\sin^2(2\theta_{12}) = 0.092 \pm 0.017 \quad (52)$$

Following this announcement, several experiments confirmed the finding and during the last years the last mixing angle to be measured became the best (most precisely) measured one. The fact that this angle, although smaller than the other two, is still sizeable opens the door to a new generation of neutrino experiments aiming to answer the open questions in the field.

## 5 $\nu$ Standard Model

Now that we have comprehended the physics behind neutrinos oscillations and have leaned the experimental evidence about the parameters driving these oscillations, we can move ahead and construct the Neutrino Standard Model:

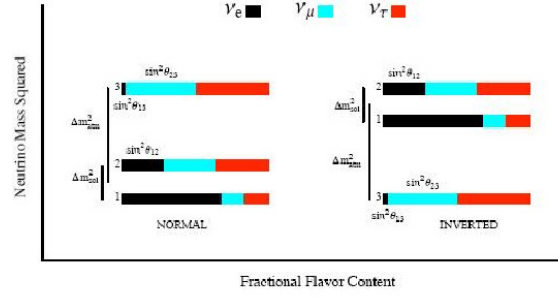
- it comprises three light ( $m_i < 1 \text{ eV}$ ) neutrinos, *ie.* it involves just two mass differences  $\Delta m_{\text{atm}}^2 \approx 2.5 \times 10^{-3} \text{ eV}^2$  and  $\Delta m_{\text{solar}}^2 \approx 8.0 \times 10^{-5} \text{ eV}^2$ .
- so far we have not seen any solid experimental indication (or need) for additional neutrinos<sup>3</sup>. As we have measured long time ago the invisible width of the  $Z$  boson and found it to be 3, within errors, if additional neutrinos are going to be incorporated into the model, they cannot couple to the  $Z$  boson, *ie.* they cannot enjoy weak interactions, so we call them sterile. However, as sterile neutrinos have not been seen (although they may have been hinted), and are not needed to explain any solid experimental evidence, our Neutrino Standard Model will contain just the three active flavours:  $e$ ,  $\mu$  and  $\tau$ .
- the unitary mixing matrix which rotates from the flavour to the mass basis, called the PMNS matrix, comprises three mixing angles (the so called solar mixing angle:  $\theta_{12}$ , the atmospheric mixing angle  $\theta_{23}$ , and the last to be measured, the reactor mixing angle  $\theta_{13}$ ), one Dirac phase ( $\delta$ ) and potentially two Majorana phases ( $\alpha$ ,  $\beta$ ) and is given by

$$|\nu_\alpha\rangle = U_{\alpha i} |\nu_i\rangle$$

$$U_{\alpha i} = \begin{pmatrix} 1 & & & \\ & c_{23} & s_{23} & \\ & -s_{23} & c_{23} & \\ & & & 1 \end{pmatrix} \begin{pmatrix} c_{13} & & s_{13}e^{-i\delta} & \\ & 1 & & \\ -s_{13}e^{i\delta} & & c_{13} & \\ & & & 1 \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & & \\ -s_{12} & c_{12} & & \\ & & & 1 \end{pmatrix} \begin{pmatrix} 1 & & & \\ & e^{i\alpha} & & \\ & & & e^{i\beta} \end{pmatrix}$$

where  $s_{ij} = \sin \theta_{ij}$  and  $c_{ij} = \cos \theta_{ij}$ . Courtesy of the hierarchy in mass differences (and to a less extent to the smallness of the reactor mixing angle) we are permitted to recognise the (23) label in the three neutrino scenario as the atmospheric  $\Delta m_{\text{atm}}^2$  we obtained in the two neutrino scenario, in

<sup>3</sup>Although it must be noted that there are several not significant hints pointing in this direction



**Fig. 6:** Flavour content of the three neutrino mass eigenstates (not including the dependence on the cosine of the CP violating phase  $\delta$ ). If CPT is conserved, the flavour content must be the same for neutrinos and anti-neutrinos. Notice that oscillation experiments cannot tell us how far above zero the entire spectrum lies.

a similar fashion the (12) label can be assimilated to the solar  $\Delta m_{\odot}^2$ . The (13) sector drives the  $\nu_e$  flavour oscillations at the atmospheric scale, and the depletion in reactor neutrino fluxes see [23]. According to the experiments done so far, the three sigma ranges for the neutrino mixing angles are

$$0.267 < \sin^2 \theta_{12} < 0.344 \quad ; \quad 0.342 < \sin^2 \theta_{23} < 0.667 \quad ; \quad 0.0156 < \sin^2 \theta_{13} < 0.0299$$

while the corresponding ones for the mass splittings are

$$2.24 \times 10^{-3} \text{eV}^2 < |\Delta m_{32}^2| < 2.70 \times 10^{-3} \text{eV}^2$$

and

$$7. \times 10^{-5} \text{eV}^2 < \Delta m_{21}^2 < 8.09 \times 10^{-5} \text{eV}^2.$$

These mixing angles and mass splittings are summarised in Fig. 6.

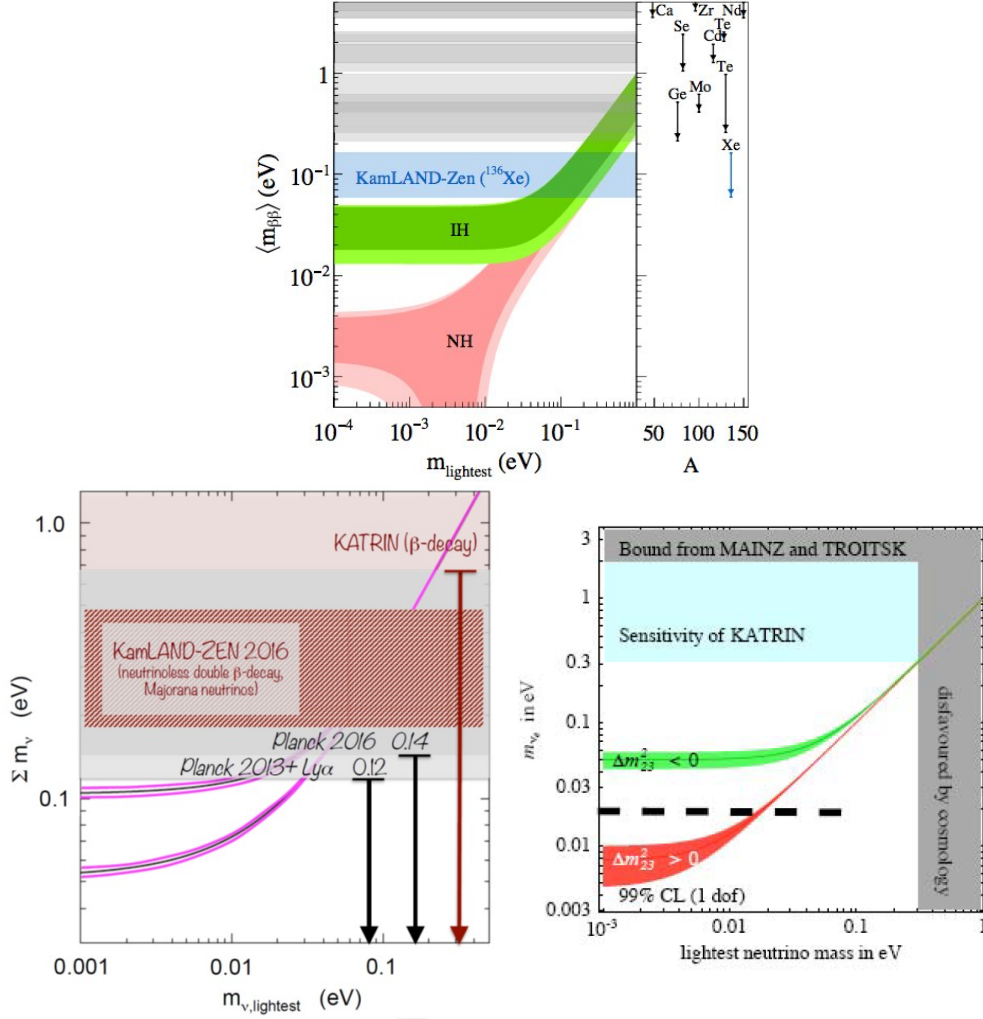
- As oscillation experiments only explore the two mass differences, two ordering are possible, as shown in Fig. 6. They are called normal and inverted hierarchy and roughly identify whether the mass eigenstate with the smaller electron neutrino content is the lightest or the heaviest.
- The absolute mass scale of the neutrinos, or the mass of the lightest neutrino is not know yet, but cosmological bounds already say that the heaviest one must be lighter than about .5 eV.
- As transition or survival probabilities depend on the combination  $U_{\alpha i}^* U_{\beta i}$  no trace of the Majorana phases could appear on oscillation phenomena, however they will have observable effects in those processes where the Majorana character of the neutrino is essential for the process to happen, like neutrino-less double beta decay.

## 6 Neutrino mass and character

### 6.1 Absolute Neutrino Mass

The absolute mass scale of the neutrino, *ie.* the mass of the lightest/heaviest neutrino, cannot be obtained from oscillation experiments, however this does not mean we have no access to it. Direct experiments like tritium beta decay, or neutrinoless double beta decay and indirect ones, like cosmological observations, have potential to feed us the information on the absolute scale of neutrino mass, we so desperately need. The Katrin tritium beta decay experiment, [24], has sensitivity down to 200 meV for the "mass" of  $\nu_e$  defined as

$$m_{\nu_e} = |U_{e1}|^2 m_1 + |U_{e2}|^2 m_2 + |U_{e3}|^2 m_3. \quad (53)$$



**Fig. 7:** The effective mass measured in double  $\beta$  decay, in cosmology and in Tritium  $\beta$  decay versus the mass of the lightest neutrino. Below the dashed lines, only the normal hierarchy is allowed. Notice that while double  $\beta$  decay experiments bound the neutrino mass only in the Majorana case, Planck bounds apply for either case

Neutrino-less double beta decay experiments, see [25] for a review, do not measure the absolute mass of the neutrino directly but a particular combination of neutrino masses and mixings,

$$m_{\beta\beta} = \left| \sum m_i U_{ei}^2 \right| = \left| m_a c_{13}^2 c_{12}^2 + m_2 c_{13}^2 s_{12}^2 e^{2i\alpha} + m_3 s_{13}^2 e^{2i\beta} \right|, \quad (54)$$

where it is understood that neutrinos are taken to be Majorana particles, *ie.* truly neutral particles (having all their quantum numbers to be zero). The new generation of experiments seeks to reach below 10 meV for  $m_{\beta\beta}$  in double beta decay.

Cosmological probes (CMB and Large Scale Structure experiments) measure the sum of the neutrino masses

$$m_{\text{cosmo}} = \sum_i m_i. \quad (55)$$

and may have a say on the mass ordering (direct or inverted spectrum) as well as test other neutrino properties like neutrino asymmetries [26]. If  $\sum m_i \approx 10$  eV, the energy balance of the universe saturates the bound coming from its critical density. The current limit, [27], is a few % of this number,  $\sim .5$

eV. These bounds are model dependent but they do all give numbers of the same order of magnitude. However, given the systematic uncertainties characteristic of cosmology, a solid limit of less than 100 meV seems way too aggressive.

Fig. 7 shows the allowed parameter space for the neutrino masses (as a function of the absolute scale) for both the normal and inverted hierarchy.

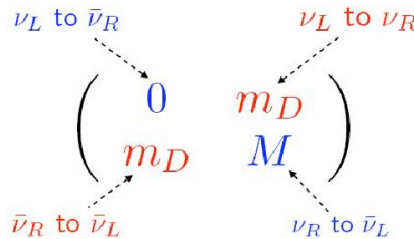
### 6.2 Majorana vs Dirac

A fermion mass is nothing but a coupling between a left handed state and a right handed one. Thus, if we examine a massive fermion at rest, then one can regard this state as a linear combination of two massless particles, one right handed and one left handed. If the particle we are examining is electrically charged, like an electron or a muon, both particles, the left handed as well as the right handed must have the same charge (we want the mass term to be electrically neutral). This is a Dirac mass term. However, for a neutral particle, like a sterile neutrino, a new possibility opens up, the left handed particle can be coupled to the right handed anti-particle, (a term which would have a net charge, if the fields are not absolutely and totally neutral) this is a Majorana mass term.

Thus a truly and absolutely neutral particle (who will inevitably be its own antiparticle) does have two ways of getting a mass term, a la Dirac or a la Majorana, and if there are no reasons to forbid one of them, will have them both, as shown in Fig. 6.2.



In the case of a neutrino, the left chiral field couples to  $SU(2) \times U(1)$  implying that a Majorana mass term is forbidden by gauge symmetry. However, the right chiral field carries no quantum numbers, is totally and absolutely neutral. Then, the Majorana mass term is unprotected by any symmetry and it is expected to be very large, of the order of the largest scale in the theory. On the other hand, Dirac mass terms are expected to be of the order of the electroweak scale times a Yukawa coupling, giving a mass of the order of magnitude of the charged lepton or quark masses. Putting all the pieces together, the mass matrix for the neutrinos results as in Fig. 8.



**Fig. 8:** The neutrino mass matrix with the various right to left couplings,  $M_D$  is the Dirac mass terms while 0 and  $M$  are Majorana masses for the charged and uncharged (under  $SU(2) \times U(1)$ ) chiral components

To get the mass eigenstates we need to diagonalise the neutrino mass matrix. By doing so, one is left with two Majorana neutrinos, one super-heavy Majorana neutrino with mass  $\simeq M$  and one super-light Majorana neutrino with mass  $m_D^2/M$ , *ie.* one mass goes up while the other sinks, this is what

we call the seesaw mechanism, [28–30]<sup>4</sup>. The light neutrino(s) is(are) the one(s) observed in current experiments (its mass differences) while the heavy neutrino(s) are not accessible to current experiments and could be responsible for explaining the baryon asymmetry of the universe through the generation of a lepton asymmetry at very high energy scales since its decays can in principle be CP violating (they depend on the two Majorana phases on the PNMS matrix which are invisible for oscillations). The super heavy Majorana neutrinos being their masses so large can play a role at very high energies and can be related to inflation [31].

If neutrinos are Majorana particles lepton number is no longer a good quantum number and a plethora of new processes forbidden by lepton number conservation can take place, it is not only neutrino-less double beta decay. For example, a muon neutrino can produce a positively charged muon. However, this process and any processes of this kind, would be suppressed by  $(m_\nu/E)^2$  which is tiny,  $10^{-20}$ , and therefore, although they are technically allowed, are experimentally unobservable. To most stringent limit nowadays comes from KamLAND-zen [32], and constraints the half-life of neutrino-less double beta decay to be  $T_{1/2}^{0\nu} > 1.07 \times 10^{26}$  years at 90% C.L. Forthcoming experiments such as GERDA-PhaseII, Majorana, SuperNEMO, CUORE, and nEXO will improve this sensitivity by one order of magnitude.

Recently low energy seesaw models [33] have experienced a revival and are actively being explored [34]. In such models the heavy states, of only few tens of TeV can be searched for at the LHC. The heavy right handed states in these models will be produced at LHC either through Yukawa couplings or through gauge coupling to right handed gauge bosons. Some models contain also additional scalar that can be looked for.

## 7 Conclusions

The experimental observations of neutrino oscillations, meaning that neutrinos have mass and mix, answered questions that had endured since the establishment of the Standard Model. As those veils have disappeared, new questions open up and challenge our understanding

- what is the true nature of the neutrinos ? are they Majorana particles or Dirac ones ? are neutrinos totally neutral ?
- is there any new scale associated to neutrinos masses ? can it be accessible at colliders ?
- is the spectrum normal or inverted ? is the lightest neutrino the one with the least electron content on it, or is it the heaviest one ?
- is CP violated (is  $\sin \delta \neq 0$ ) ? if so, is this phase related at any rate with the baryon asymmetry of the Universe ? what about the other two phases ?
- which is the absolute mass scale of the neutrinos ?
- are there new interactions ? are neutrinos related to the open questions in cosmology, like dark matter and/or dark energy ? do (presumably heavy) neutrinos play a role in inflation ?
- can neutrinos violate CPT [35]? what about Lorentz invariance ?
- if we ever measure a different spectrum for neutrinos and antineutrinos (after matter effects are properly taken into account), how can we distinguish whether it is due to a true (genuine) CTP violation or to a non-standard neutrino interaction ?
- are these intriguing signals in short baseline reactor neutrino experiments (the missing fluxes) a real effect ? Do they imply the existence of sterile neutrinos ?

We would like to answer these questions. For doing it, we are doing right now, and we plan to do new experiments. These experiments will, for sure bring some answers and clearly open new, pressing questions. Only one thing is clear. Our journey into the neutrino world is just beginning.

<sup>4</sup>Depending on the envisioned high energy theory, the simplest see saw mechanism can be categorised into three different classes or types (as they are called) depending on their scalar content.



## Acknowledgements

I would like to thank the students and the organisers of the European School on HEP for giving me the opportunity to present these lectures in such a wonderful atmosphere (and for pampering me beyond my expectations, which were not low). I enjoyed each day of the school enormously. Support from the MEC and FEDER (EC) Grants SEV-2014-0398 and FPA2014-54459 and the Generalitat Valenciana under grant PROMETEOII/2013/017 is acknowledged. The author's work has received funding from the European Union Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant Elusives ITN agreement No 674896 and InvisiblesPlus RISE, agreement No 690575.

## References

- [1] A copy of the letter in both, German and English can be found in <http://microboone-docdb.fnal.gov/cgi-bin/RetrieveFile?docid=953;filename=pauli>
- [2] Z. Maki, M. Nakagawa and S. Sakata, *Prog. Theor. Phys.* **28**, 870 (1962).
- [3] H. J. Lipkin, *Phys. Lett. B* **579**, 355 (2004) [arXiv:hep-ph/0304187].
- [4] L. Stodolsky, *Phys. Rev. D* **58**, 036006 (1998) [arXiv:hep-ph/9802387].
- [5] An analysis of CPT violation in the neutrino sector can be found in: G. Barenboim and J. D. Lykken, *Phys. Lett. B* **554** (2003) 73 [arXiv:hep-ph/0210411] ; G. Barenboim, J. F. Beacom, L. Borissoff and B. Kayser, *Phys. Lett. B* **537** (2002) 227 [arXiv:hep-ph/0203261].
- [6] L. Wolfenstein, *Phys. Rev. D* **17**, 2369 (1978).
- [7] S. P. Mikheev and A. Y. Smirnov, "Resonance enhancement of oscillations in matter and solar neutrino Sov. J. Nucl. Phys. **42**, 913 (1985) [*Yad. Fiz.* **42**, 1441 (1985)].
- [8] S. P. Mikheev and A. Y. Smirnov, "Neutrino oscillations in a variable-density medium and  $\nu$ -bursts due Sov. Phys. JETP **64**, 4 (1986) [*Zh. Eksp. Teor. Fiz.* **91**, 7 (1986)] [arXiv:0706.0454 [hep-ph]].
- [9] S. P. Mikheev and A. Y. Smirnov, "Resonant amplification of neutrino oscillations in matter and solar Nuovo Cim. C **9**, 17 (1986).
- [10] N. Tolich [SNO Collaboration], *J. Phys. Conf. Ser.* **375** (2012) 042049.
- [11] Y. Ashie *et al.* [Super-Kamiokande Collaboration], "A measurement of atmospheric neutrino oscillation parameters by *Phys. Rev. D* **71**, 112005 (2005) [arXiv:hep-ex/0501064]; Y. Takeuchi [Super-Kamiokande Collaboration], arXiv:1112.3425 [hep-ex].
- [12] M. H. Ahn *et al.* [K2K Collaboration], *Phys. Rev. D* **74** (2006) 072003 [hep-ex/0606032]; C. Mariani [K2K Collaboration], *AIP Conf. Proc.* **981** (2008) 247. doi:10.1063/1.2898949
- [13] M. Scott [T2K Collaboration], arXiv:1606.01217 [hep-ex].
- [14] A. B. Sousa [MINOS and MINOS+ Collaborations], *AIP Conf. Proc.* **1666** (2015) 110004 doi:10.1063/1.4915576 [arXiv:1502.07715 [hep-ex]]; L. H. Whitehead [MINOS Collaboration], *Nucl. Phys. B* **908** (2016) 130 doi:10.1016/j.nuclphysb.2016.03.004 [arXiv:1601.05233 [hep-ex]].
- [15] P. Adamson *et al.* [NOvA Collaboration], *Phys. Rev. D* **93** (2016) no.5, 051104 doi:10.1103/PhysRevD.93.051104 [arXiv:1601.05037 [hep-ex]].
- [16] I. Shimizu [KamLAND Collaboration], *Nucl. Phys. Proc. Suppl.* **188** (2009) 84. doi:10.1016/j.nuclphysbps.2009.02.020
- [17] M. P. Decowski [KamLAND Collaboration], *Nucl. Phys. B* **908** (2016) 52. doi:10.1016/j.nuclphysb.2016.04.014
- [18] B. Aharmim *et al.* [SNO Collaboration], *Phys. Rev. C* **88** (2013) 025501 doi:10.1103/PhysRevC.88.025501 [arXiv:1109.0763 [nucl-ex]].
- [19] A. Hime, *Nucl. Phys. Proc. Suppl.* **221** (2011) 110. doi:10.1016/j.nuclphysbps.2011.03.104
- [20] S. J. Parke and T. P. Walker, *Phys. Rev. Lett.* **57**, 2322 (1986) [Erratum-*ibid.* **57**, 3124 (1986)].

- [21] F. P. An *et al.* [DAYA-BAY Collaboration], Phys. Rev. Lett. **108** (2012) 171803 [arXiv:1203.1669 [hep-ex]].
- [22] F. P. An *et al.* [Daya Bay Collaboration], arXiv:1610.04802 [hep-ex]; S. B. Kim [RENO Collaboration], Nucl. Phys. B **908** (2016) 94; Y. Abe *et al.* [Double Chooz Collaboration], JHEP **1410** (2014) 086 Erratum: [JHEP **1502** (2015) 074] [arXiv:1406.7763 [hep-ex]].
- [23] F. P. An *et al.* [Daya Bay Collaboration], Phys. Rev. Lett. **116** (2016) no.6, 061801 doi:10.1103/PhysRevLett.116.061801 [arXiv:1508.04233 [hep-ex]].
- [24] S. Mertens [KATRIN Collaboration], Phys. Procedia **61** (2015) 267.
- [25] S. R. Elliott and P. Vogel, Ann. Rev. Nucl. Part. Sci. **52**, 115 (2002) [arXiv:hep-ph/0202264].
- [26] G. Barenboim, W. H. Kinney and W. I. Park, arXiv:1609.01584 [hep-ph]; G. Barenboim, W. H. Kinney and W. I. Park, arXiv:1609.03200 [astro-ph.CO].
- [27] M. Lattanzi [Planck Collaboration], J. Phys. Conf. Ser. **718** (2016) no.3, 032008. doi:10.1088/1742-6596/718/3/032008
- [28] M. Gell-Mann, P. Ramond and R. Slansky, in *Supergravity*, edited by P.van Nieuwenhuizen and D. Freedman, (North-Holland,1979), p.315.
- [29] R. N. Mohapatra and G. Senjanovic, Phys. Rev. Lett. **44**, 912 (1980).
- [30] M. Fukugita and T. Yanagida, Phys. Lett. B **174**, 45 (1986).
- [31] G. Barenboim, JHEP **0903** (2009) 102 [arXiv:0811.2998 [hep-ph]]; G. Barenboim, Phys. Rev. D **82** (2010) 093014 [arXiv:1009.2504 [hep-ph]].
- [32] Y. Gando [KamLAND-Zen Collaboration], Nucl. Part. Phys. Proc. **273-275** (2016) 1842. doi:10.1016/j.nuclphysbps.2015.09.297
- [33] F. Borzumati and Y. Nomura, Phys. Rev. D **64** (2001) 053005 doi:10.1103/PhysRevD.64.053005 [hep-ph/0007018].
- [34] C. G. Cely, A. Ibarra, E. Molinaro and S. T. Petcov, Phys. Lett. B **718** (2013) 957 doi:10.1016/j.physletb.2012.11.026 [arXiv:1208.3654 [hep-ph]].
- [35] G. Barenboim, L. Borisso, J. D. Lykken and A. Y. Smirnov, JHEP **0210** (2002) 001 [hep-ph/0108199]. G. Barenboim and J. D. Lykken, Phys. Rev. D **80** (2009) 113008 [arXiv:0908.2993 [hep-ph]].

## Beyond the Standard Model

*B.C. Allanach*

Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences,  
University of Cambridge, Cambridge, United Kingdom

### Abstract

We cover some current topics in Beyond the Standard Model phenomenology, with an emphasis on collider (particularly Large Hadron Collider) phenomenology. We begin with a review of the Standard Model and some unresolved mysteries that it leaves. Then, we shall heuristically introduce supersymmetry, grand unified theories and extra dimensions as paradigms for expanding the Standard Model. The collider phenomenology of such models is too rich and complex to review, but we give some key examples of how the new states associated with the models might be inferred in Large Hadron Collider events<sup>1</sup>. Before concluding, we finish with a brief description of a quantum field theory approximation that can be used in some cases to reduce model dependence: effective field theory.

### Keywords

CERN report; contribution; supersymmetry, extra dimensions, large hadron collider; effective field theories; gauge unification

## 1 Introduction

We must remember that the Standard Model of particle physics is a remarkably successful physical theory. It has been tested in literally thousands of different and diverse ways. Some of its predictions (for example the anomalous electron magnetic moment) have been verified to one part in  $10^{10}$ , whereas some of them (particularly the ones involving low energies and the strong interactions) have only been tested at the 10% level. However, there is to date no unambiguous direct collider measurement which rules it out. The more precise predictions are sensitive to higher loops of Standard Model particles (and in principle could be affected by loops involving beyond the Standard Model particles). Going beyond the Standard Model successfully then should not upset any of these successful predictions, and so any extension is likely to only be a small perturbation, at least at the energy scales currently being probed. Let us now turn to the fundamentals that The Standard Model is built upon.

### 1.1 A basic theory: quantum field theory

Microscopically we have *quantum mechanics* and *special relativity* as two fundamental theories. A consistent framework incorporating these two theories is *quantum field theory (QFT)*. In this theory the fundamental entities are quantum fields. Their excitations correspond to the physically observable elementary particles which are the basic constituents of matter as well as the mediators of all the known interactions. Therefore, fields have a particle-like character. Particles can be classified in two general classes: bosons (spin  $s = n \in \mathbb{Z}$ ) and fermions ( $s = n + \frac{1}{2} \forall n \in \mathbb{Z}$ ). Bosons and fermions have very different physical behaviour. The main difference is that fermions can be shown to satisfy the Pauli “exclusion principle”, which states that two identical fermions cannot occupy the same quantum state, and therefore explaining the vast diversity of atoms.

---

<sup>1</sup>A large portion of these notes is based on Prof. Fernando Quevedo’s excellent Cambridge Part III “Supersymmetry and extra dimensions” course [1], with his permission.

All apparently elementary matter particles are fermions, for example the leptons (including electrons and neutrinos) and quarks (that make protons, neutrons and all other hadrons). Bosons on the other hand include the photon (particle of light and mediator of electromagnetic interaction), and the mediators of all the other interactions. They are not constrained by the Pauli principle. As we shall see, *supersymmetry* is a symmetry that unifies bosons and fermions despite all their differences.

## 1.2 Basic principle: symmetry

If QFT is the basic framework to study elementary processes, one tool to learn about these processes is the concept of *symmetry*.

A symmetry is a transformation that can be made to a physical system leaving the physical observables unchanged. Throughout the history of science symmetry has played a very important role in better understanding nature.

## 1.3 Classes of symmetries

For elementary particles, we can define two general classes of symmetries:

- *Space-time symmetries*: These symmetries correspond to transformations on a field theory acting explicitly on the space-time coordinates,

$$x^\mu \mapsto x'^\mu \quad \forall \{\mu, \nu\} = \{0, 1, 2, 3\}. \quad (1)$$

Some examples are rotations, translations and, more generally, *Lorentz- and Poincaré transformations* defining special relativity as well as *general coordinate transformations* that define *general relativity*.

- *Internal symmetries*: These are symmetries that correspond to transformations of the different fields in a field theory,

$$\Phi^a(x) \mapsto M^a_b \Phi^b(x). \quad (2)$$

Roman indices  $a, b$  label the corresponding fields<sup>2</sup>. If  $M^a_b$  is constant then the symmetry is a *global symmetry*; in case of space-time dependent  $M^a_b(x)$  the symmetry is called a *local symmetry* or a *gauge symmetry*.

## 1.4 Importance of symmetries

Symmetry is important for various reasons:

- *Labelling and classifying particles*: Symmetries label and classify particles according to the different conserved quantum numbers identified by the space-time and internal symmetries (mass, spin, charge, colour, etc.). In this regard symmetries actually “define” an elementary particle according to the behaviour of the corresponding field with respect to the different symmetries.
- Symmetries determine the *interactions* among particles, by means of the *gauge principle*, for instance. It is important that *most QFTs of vector bosons are sick: they are non-renormalisable in a way that makes them unpredictable*. The counter example to this is *gauge theory*, where vector bosons are *necessarily in the adjoint representation* of the gauge group. As an illustration, consider the Lagrangian

$$\mathcal{L} = \partial_\mu \phi \partial^\mu \phi^* - V(\phi, \phi^*) \quad (3)$$

which is invariant under rotations in the complex plane

$$\phi \mapsto \exp(i\alpha) \phi, \quad (4)$$

---

<sup>2</sup>Unless otherwise noted, we follow the convention that repeated indices are summed over.

as long as  $\alpha$  is a constant (this corresponds to a global symmetry). If  $\alpha = \alpha(x)$ , the kinetic term is no longer invariant:

$$\partial_\mu \phi \mapsto \exp(i\alpha) (\partial_\mu \phi + i(\partial_\mu \alpha)\phi). \quad (5)$$

However, the covariant derivative  $D_\mu$ , defined as

$$D_\mu \phi = \partial_\mu \phi + iA_\mu \phi, \quad (6)$$

transforms like  $\phi$  itself, if the gauge - potential  $A_\mu$  transforms to  $A_\mu - \partial_\mu \alpha$ :

$$\begin{aligned} D_\mu \phi &\mapsto \exp(i\alpha) (\partial_\mu \phi + i(\partial_\mu \alpha)\phi + i(A_\mu - \partial_\mu \alpha)\phi) \\ &= \exp(i\alpha) D_\mu \phi, \end{aligned}$$

so we rewrite the Lagrangian to ensure gauge invariance:

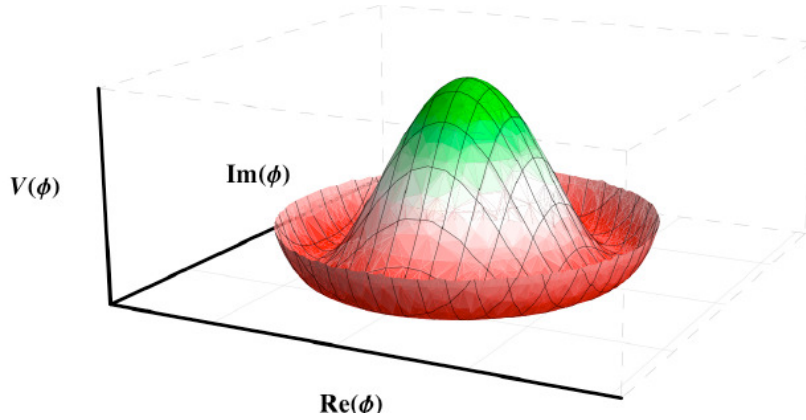
$$\mathcal{L} = D_\mu \phi (D^\mu \phi)^* - V(\phi, \phi^*). \quad (7)$$

The scalar field  $\phi$  couples to the gauge field  $A_\mu$  via  $A_\mu \phi A^\mu \phi$ , similarly, the Dirac Lagrangian

$$\mathcal{L} = \bar{\Psi} \gamma^\mu D_\mu \Psi \quad (8)$$

contains an interaction term  $\bar{\Psi} A_\mu \Psi$ . This interaction provides the three point vertex that describes interactions of electrons and photons, illustrating how photons mediate the electromagnetic interactions.

- Symmetries can hide or be *spontaneously broken*: Consider the potential  $V(\phi, \phi^*)$  in the scalar field Lagrangian above.



**Fig. 1:** The Mexican hat potential for  $V = (a - b|\phi|^2)^2$  with  $a, b \geq 0$ . From Ref. [1].

If  $V(\phi, \phi^*) = V(|\phi|^2)$ , then it is symmetric for  $\phi \mapsto \exp(i\alpha)\phi$ . If the potential is of the type

$$V = a|\phi|^2 + b|\phi|^4 \forall a, b \geq 0, \quad (9)$$

then the minimum is at  $\langle \phi \rangle = 0$  (here  $\langle \phi \rangle \equiv \langle 0|\phi|0 \rangle$  denotes the *vacuum expectation value (VEV)* of the field  $\phi$ ). The vacuum state is then also symmetric under the symmetry since the origin is invariant. However if the potential is of the form

$$V = (a - b|\phi|^2)^2 \forall a, b \geq 0, \quad (10)$$

the symmetry of  $V$  is lost in the ground state  $\langle \phi \rangle \neq 0$ . The existence of hidden symmetries is important for at least two reasons:

- (i) This is a natural way to introduce an energy scale in the system, determined by the non vanishing VEV. In particular, in the Standard Model, the electroweak scale  $M_{ew} \sim 10^2$  GeV defines the basic scale of mass for the particles of the standard model, the electroweak gauge bosons and the matter fields, through their Yukawa couplings, obtain their mass from the VEV.
- (ii) The existence of hidden symmetries implies that the fundamental symmetries of nature may be larger than is apparent. This is because the only manifest symmetries we can observe are the symmetries of the vacuum we live in and not those of the full underlying theory. This opens-up an essentially unlimited resource to consider physical theories with an indefinite number of symmetries even though they are not explicitly realized in nature. The standard model is one typical example and supersymmetry and theories of extra dimensions are further examples.

### 1.4.1 The Standard Model

The Standard Model is well defined and currently well confirmed by experiments. It is based on the two classes of symmetry:

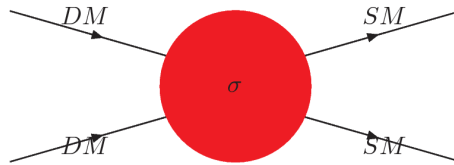
- *space-time symmetry*: Poincaré symmetry in 4 dimensions.
- *internal symmetry*: gauged  $G_{SM} = SU(3)_c \times SU(2)_L \times U(1)_Y$  symmetry, where  $SU(3)_c$  defines the strong interactions.  $SU(2)_L \times U(1)_Y$  is spontaneously broken by the *Higgs* mechanism to  $U(1)_{em}$ . The gauge fields are spin-1 bosons, for example the photon  $A^\mu$ , or gluons  $G^{a=1,\dots,8}$ . Matter fields (quarks and leptons) have spin  $1/2\hbar$  and come in three ‘families’ (successively heavier copies). The Higgs boson (a particle has been discovered at the LHC whose properties are consistent with the Standard Model Higgs boson) is the spin zero particle that spontaneously breaks the  $SU(2)_L \times U(1)_Y$ . The  $W^\pm$  and  $Z^0$  bosons get a mass via the Higgs mechanism and therefore the weak interactions are short range. This is also the source of masses for all quarks and leptons. The sub-index  $L$  in  $SU(2)_L$  refers to the fact that the Standard Model does not preserve parity and differentiates between left-handed and right-handed particles. In the Standard Model only left-handed particles transform non-trivially under  $SU(2)_L$ . The gauge particles have all spin  $s = 1\hbar$  and mediate each of the three forces: photons ( $\gamma$ ) for  $U(1)$  electromagnetism, gluons for  $SU(3)_C$  of strong interactions, and the massive  $W^\pm$  and  $Z^0$  bosons for the weak interactions.

## 1.5 Problems of the Standard Model

The Standard Model is one of the cornerstones of all science and one of the great triumphs of the past century. It has been carefully experimentally verified in many ways, especially during the past 20 years. However, there are still some unresolved issues or mysteries:

- The hierarchy problem. The Higgs mass is  $m_h \approx 125$  GeV, whereas the gravitational scale is  $M_{Planck} \sim \sqrt{G} \sim 10^{19}$  GeV. The ‘hierarchy problem’ is: why is  $m_h/M_{Planck} \sim 10^{-17}$  so much smaller than 1? In a fundamental theory, one might expect them to be the same order. In QFT, one sees that quantum corrections (loops) to  $m_h$  are expected to be of order of the heaviest scale in the theory divided by  $4\pi$ . The question of why the hierarchy is stable with respect to the quantum corrections is called the *technical hierarchy problem*, and is arguably the main motivation for weak-scale supersymmetry.
- The cosmological constant ( $\Lambda$ ) problem: probably the biggest unsolved problem in fundamental physics.  $\Lambda$  is the energy density of free space time. The cosmological constant problem is: Why is  $(\Lambda/M_{Planck})^4 \sim 10^{-120} \ll 1$ ?

- The Standard Model has around 20 parameters, which must be measured then set ‘by hand’. Many consider that a more satisfying fundamental theory would relate all of these parameters to less (or ideally one) fundamental parameter.
- What particle constitutes the inferred cold dark matter in the universe? It is not contained in the Standard Model. Planck and large scale structure data favour a cosmological constant-cold dark matter model, where approximately 22% of the universe’s energy budget lies in dark matter, only 4% in ordinary matter, and some 74% in mysterious dark energy<sup>3</sup>. Neutrinos constitute a hot component of dark matter (since they are relativistic when they decouple from the thermal plasma i.e. they smooth density perturbations in the early universe on smaller scales), so they are not good candidates.



**Fig. 2:** For time  $t \rightarrow$  (i.e. time increasing toward the right), this describes *annihilation*: once the particle physics model is set, a calculation tells us how much is thermally produced in the early universe. This also is a diagram for dark matter indirect detection, for example by dark matter collecting in the core of the sun and annihilating into neutrinos which could be detected by the IceCube experiment. For  $t \leftarrow$ , the diagram depicts *collider production* at (e.g.) the LHC, whereas for  $t \uparrow$ , it’s *direct detection*, where dark matter colliding with heavy nuclei may produce measurable nuclear recoils.

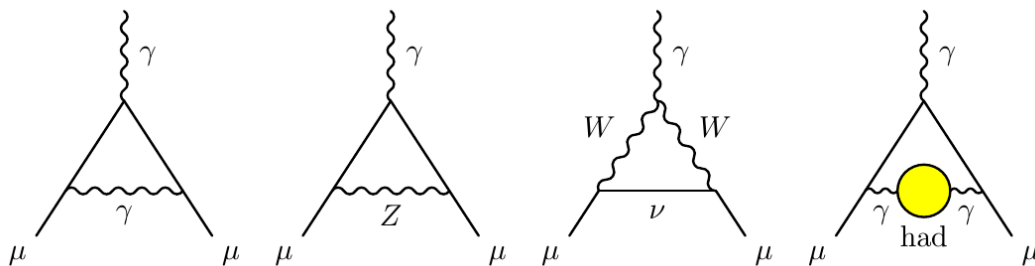
- The anomalous magnetic moment of the muon: This is a particular interaction between the photon and the muon: the Dirac equation predicts a muon magnetic moment

$$\vec{M} = g_\mu \frac{e}{2m_\mu} \vec{S}, \tag{11}$$

and at tree level,  $g_\mu = 2$ . However, it can be measured very precisely by storing muons in a ring with magnetic fields, then measuring the *precession frequency* of their spins. The ‘anomalous’ part comes from loops involving various particles. Defining  $a_\mu \equiv \frac{g_\mu - 2}{2}$  [2],

$$\begin{aligned} a_\mu^{\text{exp}} &= 11659209.1(5.4)(3.3) \times 10^{-10}, & a_\mu^{\text{SM}} &= 11659180.3(4.2)(2.6) \times 10^{-10}, \\ \Rightarrow \Delta a_\mu &= a_\mu^{\text{exp}} - a_\mu^{\text{SM}} = 28.8(6.3)(4.9) \times 10^{-10}, \end{aligned} \tag{12}$$

<sup>3</sup>A tiny negative energy density of space-time,  $\Lambda \sim \mathcal{O}(10^{-3} \text{ eV})^4$ .



**Fig. 3:** Some SM contributions to the anomalous magnetic moment of the muon. From Ref. [2].



where the first number in brackets labels the statistical error and the second the systematic error. The measurement of  $(g - 2)_\mu$  thus differs with the SM prediction at around the  $\sim 3.6\sigma$  level (and has done for some 20 years). There should be a new more accurate measurement from the Muon  $g - 2$  experiment at the Fermilab collider in 2017. If one adds new particles to the SM, it is possible that they could travel in loops in diagrams similar to those in Fig. 3, and introduce a non-standard contribution to explain the discrepancy between the SM prediction and the SM measurement.

We wish to find extensions that could solve some or all of the problems mentioned above in order to generalize the Standard Model. Experiments are a traditional way of making progress in science. We need experiments to explore energies above the currently attainable scales and discover new particles and underlying principles that generalize the Standard Model. This approach is of course being followed at the LHC. The LHC will explore physics at the TeV scale, an interesting and important régime for new physics beyond the Standard Model. Notice that directly exploring energies closer to the Planck scale  $M_{Planck} \approx 10^{19}$  GeV is out of the reach for many years to come.

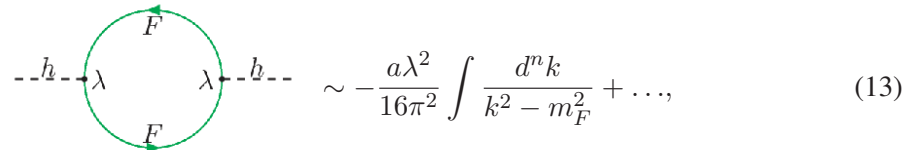
### 1.5.1 The technical hierarchy problem

The Planck mass  $M_{pl} \approx 10^{19}$  GeV is an energy scale associated with gravity and the electroweak scale  $M_{ew} \approx 10^2$  GeV is an energy scale associated with the electroweak symmetry breaking scale of the Standard Model. The hierarchy problem involves these two scales being so different in magnitude. Actually the problem can be formulated in two parts:

- (i) Why is  $M_{ew} \ll M_{xpl}$  at tree level? This question is known as ‘the hierarchy problem’. There are many solutions, once the SM is extended.
- (ii) Once we have solved (i), we ask why is the hierarchy stable under quantum corrections? This is the ‘technical hierarchy problem’ and does not have many full/effective solutions, aside from supersymmetry (SUSY).

Let us now think some more about the technical hierarchy problem. In the Standard Model we know that:

- Vector bosons are massless due to gauge invariance, that means, a direct mass term for the gauge particles  $M^2 A_\mu A^\mu$  is not allowed by gauge invariance ( $A_\mu \rightarrow A_\mu + \partial_\mu \alpha$  for a  $U(1)$  field, for example).
- Chiral fermion masses  $m \bar{\psi}_L \psi_R$  are also forbidden for all quarks and leptons by gauge invariance (because, for example,  $\psi_L$  and  $\psi_R$  have different hypercharges). Recall that these particles receive a mass only through the Yukawa couplings to the Higgs (e.g.  $H \bar{\psi}_L \psi_R$  giving a Dirac mass to  $\psi$  after  $H$  gets a non-zero value<sup>4</sup>).
- The Higgs boson is the only fundamental scalar particle in the Standard Model. There is no symmetry banning its mass term  $m_h^2 H^\dagger H$  in the Standard Model Lagrangian. If the heaviest state in the theory has a mass squared of  $\Lambda^2$ , loops give corrections of order  $\Lambda^2/(16\pi^2)$  to the scalar mass squared. The corrections come from both bosons and fermions running in loops, for example:



$$\begin{array}{c} \text{---}\lambda\text{---} \\ \text{---}\lambda\text{---} \end{array} \sim -\frac{a\lambda^2}{16\pi^2} \int \frac{d^n k}{k^2 - m_F^2} + \dots, \quad (13)$$

where  $a$  is some dimensionless  $\mathcal{O}(1)$  constant. The quantum correction to the Higgs mass from this diagram are:

$$m_h^{phys^2} = (125 \text{ GeV}/c^2)^2 = m_h^{tree^2} + \mathcal{O}(m_F^2/(16\pi^2)). \quad (14)$$

<sup>4</sup>With  $R$ -parity conservation (see below), the minimal supersymmetric standard model does not give neutrinos mass. Thus one must augment the model in some way: one can do this by adding right-handed neutrinos to the model.



Experimentally, the Higgs mass is measured to be  $m_h \approx 125$  GeV. The Standard Model is considered to be unnatural since the loop corrections are typically much larger: the largest are expected to be<sup>5</sup>  $\sim \mathcal{O}(10^{17})$  GeV. Therefore even if we start with a tree-level Higgs mass of order the electroweak scale, loop corrections would bring it up to almost the highest scale in the theory:  $\Lambda/(4\pi)$ , since we expect  $m_F \sim \mathcal{O}(\Lambda)$ . This would ruin the hierarchy between large and small scales. It is possible to adjust or “fine tune” the loop corrections such as to keep the Higgs light, but this would require cancellations between the apparently unrelated tree-level and loop contributions to some 15 significant figures. This fine tuning is considered unnatural and an explanation of why the Higgs mass (and the whole electroweak scale) can be naturally maintained to be hierarchically smaller than the Planck scale or any other large cutoff scale  $\Lambda$  is required.

### 1.5.2 Modifications of the Standard Model

In order to go beyond the Standard Model we can follow several avenues, for example:

- Add new particles and/or interactions (e.g. a dark matter particle).
  - More symmetries. For example,
- (i) Internal symmetries, for example *grand unified theories (GUTs)* in which the symmetries of the Standard Model are themselves the result of the breaking of a yet larger symmetry group:

$$G_{\text{GUT}} \xrightarrow{M \approx 10^{16} \text{ GeV}} G_{\text{SM}} \xrightarrow{M \approx 10^2 \text{ GeV}} SU(3)_c \times U(1)_Y, \quad (15)$$

Let’s take one of the simplest examples,  $G_{\text{GUT}} = SU(5)$ :

$$\underline{\mathbf{5}} = \begin{pmatrix} d \\ d \\ d \\ e^+ \\ \bar{\nu}_e \end{pmatrix}_R, \quad \underline{\mathbf{10}} = \begin{pmatrix} 0 & \bar{u} & -\bar{u} & -u & -d \\ & 0 & \bar{u} & -u & d \\ & & 0 & -u & d \\ & & & 0 & e^+ \\ & & & & 0 \end{pmatrix}_L. \quad (16)$$

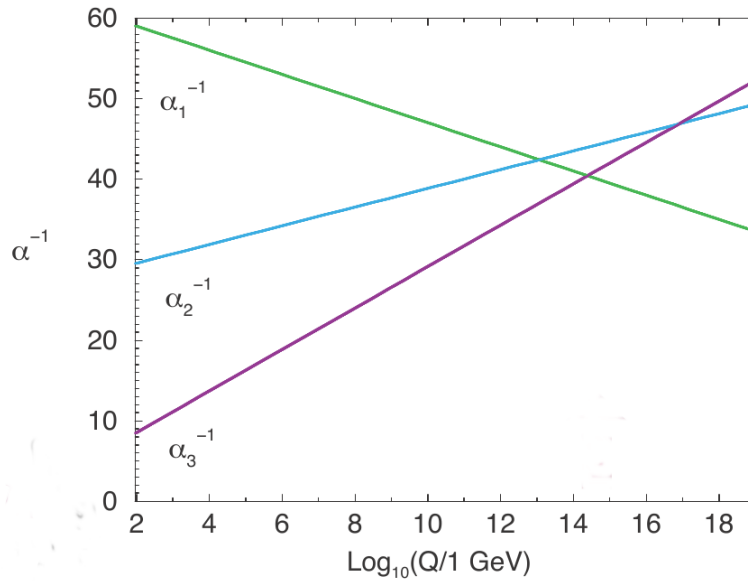
(The  $\underline{\mathbf{10}}$  is an anti-symmetric matrix; we have omitted the lower left-hand half of it because the entries are simply related to those above the diagonal). Thus, we see how quarks and leptons become unified within multiplets of  $G_{\text{GUT}}$ .

The GUT proposal is very elegant because it unifies, in one single symmetry, the three gauge interactions of the Standard Model. It leaves unanswered most of the open questions above, except for the fact that it reduces the number of independent parameters due to the fact that there is only one gauge coupling at large energies. This is expected to “run” at low energies and give rise to the three different couplings of the Standard Model (one corresponding to each group factor). Unfortunately, with our present precision understanding of the gauge couplings and spectrum of the Standard Model, the running of the three gauge couplings does **not** unify at a single coupling at higher energies but they cross each other at different energies: see Fig. 4. Because leptons and quarks are unified within GUT multiplets, they predict e.g.  $m_e(M_{\text{GUT}}) = m_d(M_{\text{GUT}})$ , which also doesn’t work, and in practice further model building is required.

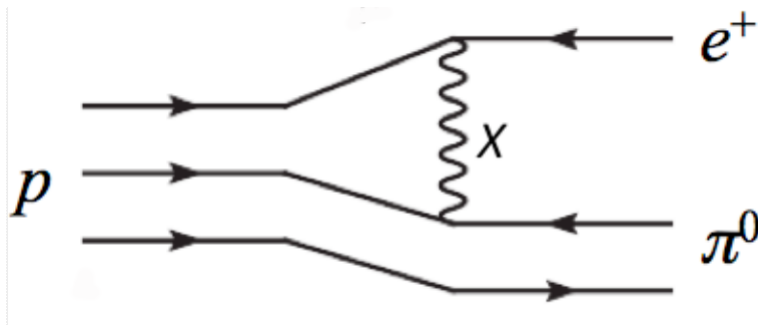
GUTs have heavy  $X$  and  $Y$  gauge boson particles of order the gauge unification scale, which arise from a GUT Higgs mechanism (in a completely analogous way to the way in which the  $W^\pm$  and  $Z^0$  bosons acquire their mass). They predict *proton decay*, which isn’t observed at super-Kamiokande. The current constraint from super-Kamiokande is that the proton lifetime  $\tau_{p \rightarrow e^+ \pi^0} > 10^{34}$  years. However, estimating  $M_{\text{GUT}} \sim 10^{15}$  GeV from Fig. 4, we predict, for ordinary GUTs, a proton lifetime of

$$\tau \approx \frac{M_{\text{GUT}}^4}{\alpha^2 m_p^5} = 4.5 \times 10^{29 \pm 1.7} \text{ years}, \quad (17)$$

<sup>5</sup>This does rely on quantum gravity yielding an effective quantum field theory that acts in the usual way.



**Fig. 4:** Gauge unification doesn't work in the Standard Model: the three gauge couplings  $\alpha_1, \alpha_2, \alpha_3$  should all unify at a single renormalisation scale  $Q$ . One needs to add some additional particles of mass below  $10^{14}$  GeV in order to make this work. Experiments (LEP and LHC experiments, for example) fix the gauge couplings at the left-hand side of the figure, and renormalisation within QFT is used to evolve them to the right. From Ref. [2].



**Fig. 5:** Example  $p \rightarrow e^+ \pi^0$  process from GUTs. From Ref. [2].

which easily is in contravention of the Super Kamiokande bound.

- (ii) *Supersymmetry*. For a phenomenological review of supersymmetry, see Ref. [3]. Supersymmetry is an external, or space-time, symmetry. Supersymmetry solves the technical hierarchy problem due to cancellations between the contributions of bosons and fermions to the electroweak scale, defined by the Higgs mass. Combined with the GUT idea, it also solves the unification of the three gauge couplings at one single point at larger energies. Supersymmetry also provides the most studied example for dark matter candidates. Moreover, it provides well defined QFTs in which the régime of strong coupling can be better studied than in non-supersymmetric models.
- (iii) *Extra spatial dimensions*. More general space-time symmetries open up many more interesting avenues for investigation. These can be of two types. First we can add more dimensions to space-time, extending the Poincaré symmetries of the Standard Model and the general coordinate transformations of general relativity. This is the well known *Kaluza Klein theory* in which our observation of a 4 dimensional universe is only due to the fact that we have limitations about “seeing” other dimensions of space-time that may be hidden to our observations. In recent years

this has been extended to the *brane world scenario* in which our 4 dimensional universe is only a brane or surface inside a higher dimensional universe. These ideas lead to a different perspective on the hierarchy problem and also may help unify internal and space-time symmetries.

- Beyond QFT: A QFT with Supersymmetry and extra dimensions does not address the problem of quantising gravity. For this purpose, the current best hope is string theory which goes beyond the basic framework of QFT. It so happens that for its consistency, string theory requires supersymmetry and extra dimensions.

## 1.6 Supersymmetry algebra

### 1.6.1 History of supersymmetry

- In the 1960's, the study of strong interactions lead to the discovery of many hadrons. These were successfully organized into multiplets of  $SU(3)_f$ , the  $f$  referring to flavour. This procedure was known as the *eight fold way* of Gell-Mann and Neeman. Questions arose about bigger multiplets including particles of different spins.
- In a famous *No-go theorem* (Coleman, Mandula 1967) said that the most general symmetry of the  $S$  - matrix (which still has non-trivial scattering) is Poincaré  $\times$  internal. The implication is that there is no symmetry that mixes up the internal and external symmetries in a non-trivial way, or that mixes particles of different spin, and still has scattering.
- Golfand and Licktmann (1971) extended the Poincaré algebra to include spinor generators  $Q_\alpha$ , where  $\alpha = 1, 2$ .
- Ramond, Neveu-Schwarz, Gervais, Sakita (1971) derived supersymmetry in 2 dimensions (from string theory).
- Wess and Zumino (1974) wrote down supersymmetric field theories in 4 dimensions. They opened the way for many other contributions to the field. This is often seen as the actual starting point for the systematic study of supersymmetry.
- Haag, Lopuszanski, Sohnius (1975): generalized the Coleman Mandula theorem to show that the only non-trivial quantum field theories have a symmetry group of super Poincaré group in a direct product with internal symmetries.

### 1.6.2 Graded algebra

The Poincaré algebra consists of commutation relations between 4-momentum operators  $P^\mu$  (generating translations in space and time) and  $M^{\mu\nu}$ , generating Lorentz boosts and rotations. Particles of the Standard Model are all irreducible representations of the Poincaré group.

To implement supersymmetry, we *extend* the Poincaré algebra non-trivially. The *Coleman Mandula theorem* stated that in 3+1 dimensions, one cannot do this in a non-trivial way and still have non-zero scattering amplitudes. In other words, there is no non-trivial mix of Poincaré and internal symmetries with non-zero scattering except for the direct product

$$\text{Poincaré} \times \text{internal}.$$

However (as usual with no-go theorems) there was a loop-hole because of an implicit axiom: the proof only considered “*bosonic* generators”.

We wish to turn bosons into fermions, thus we need to introduce a fermionic generator  $Q$ . Heuristically:

$$Q|\text{boson}\rangle \propto |\text{fermion}\rangle, \quad Q|\text{fermion}\rangle \propto |\text{boson}\rangle.$$



superpartners couple to other fields with the same strength as their SM partners) remain valid even after SUSY breaking. In particular, Eqs. 19 and 20 become

$$m_{\tilde{f}_{L,R}}^2 = m_F^2 + \Delta m^2, \quad (21)$$

whilst the scalars  $\tilde{f}_{L,R}$  and the fermion  $F$  still couple to the Higgs field  $h$  with the same strength coupling  $\lambda$ :

$$+ \text{---} \tilde{h} \text{---} \lambda \text{---} \text{---} \tilde{h} \text{---} \lambda \text{---} \sim \mathcal{O} \left( \frac{m_h^2 \log(M_Z/m_F) + \Delta m^2}{16\pi^2} \right). \quad (22)$$

Thus, as long as the *splitting* between the particles in a super multiplet is small, and as long as certain SUSY relations are preserved (such as the coupling of the Higgs field to the scalar and fermionic components of a super multiplet being equal), one still obtains only reasonable corrections to the Higgs mass squared, even if the fields  $F$  and  $\tilde{f}_{L,R}$  are very heavy. The fact that we require  $\Delta m^2/(16\pi^2)$  to be not much larger than  $m_h^2 = (125 \text{ GeV})^2 \Rightarrow \Delta m^2 < \mathcal{O}(1 \text{ TeV}^2)$ . This is then the main argument for why supersymmetric partners of SM particles should not be much heavier than the TeV scale, because otherwise its correction to the Higgs mass would be too large. Given that the LHC currently operates at a centre of mass energy of 13 TeV, this implies that there ought to be enough energy to pair produce such particles.

## 2.1 Particles

First of all, we have vector superfields containing the Standard Model gauge bosons. We write their representations under  $(SU(3)_C, SU(2)_L, U(1)_Y)$  as (pre-Higgs mechanism):

- gluons/gluinos  $G = (8, 1, 0)$
- $W$  bosons/winos  $W = (1, 3, 0)$
- $B$  bosons/gauginos  $B = (1, 1, 0),$

which contains the gauge boson of  $U(1)_Y$ .

Secondly, there are chiral superfields containing Standard Model matter and Higgs fields. Since chiral superfields only contain left-handed fermions, we place charge conjugated, i.e. *anti* right handed fermionic fields (which are actually left-handed), denoted by  $^c$  ( $\{i, j, k\} \in \{1, 2, 3\}$  are family indices):

- (s)quarks: lepton number  $L = 0$ , whereas baryon number  $B = 1/3$  for a (s)quark,  $B = -1/3$  for an anti-quark.

$$\underbrace{Q_i = \left(3, 2, \frac{1}{6}\right)}_{\text{left-handed}}, \underbrace{u_i^c = \left(\bar{3}, 1, -\frac{2}{3}\right), d_i^c = \left(\bar{3}, 1, \frac{1}{3}\right)}_{\text{anti (right-handed)}}$$

- (s)leptons  $L = 1$  for a lepton,  $L = -1$  for an anti-lepton.  $B = 0$ .

$$\underbrace{L_i = \left(1, 2, -\frac{1}{2}\right)}_{\text{left-handed}}, \underbrace{e_i^c = (1, 1, +1)}_{\text{anti (right-handed)}}$$

– Higgs bosons/higgsinos:  $B = L = 0$ .

$$H_2 = (1, 2, \frac{1}{2}), H_1 = (1, 2, -\frac{1}{2})$$

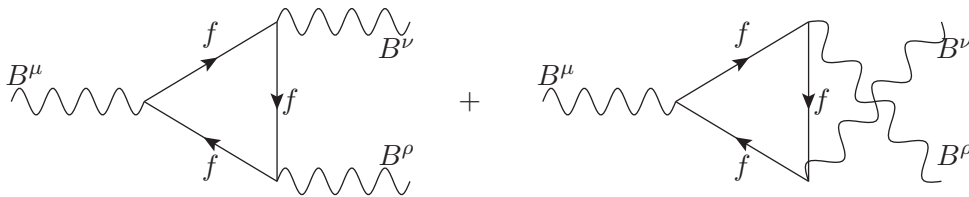
the second of which is a new Higgs doublet not present in the Standard Model. Thus, the MSSM is a *two Higgs doublet model*. The extra Higgs doublet is needed in order to avoid a gauge anomaly, and to give masses to down-type quarks and leptons.

Note that after the breaking of electroweak symmetry (see the Standard Model course), the electric charge generator is  $Q = T_3^{SU(2)_L} + Y/2$ . Baryon and lepton number correspond to multiplicative discrete perturbative symmetries in the SM, and are thus conserved, perturbatively.

Chiral fermions may generate an *anomaly* in the theory, as shown by Fig. 6. This is where a symmetry that is present in the tree-level Lagrangian is broken by quantum corrections. Here, the symmetry is  $U(1)_Y$ : all chiral fermions in the theory travel in the loop, and yield a logarithmic divergence proportional to

$$A \equiv \sum_{LH f_i} Y_i^3 - \sum_{RH f_i} Y_i^3 \tag{23}$$

multiplied by some kinematic factor which is the same for each fermion. If  $A$  is non-zero, one must renormalize the diagram away by adding a  $B_\mu B_\nu B_\rho$  counter term in the Lagrangian. But this breaks  $U(1)_Y$ , meaning that  $U(1)_Y$  would not be a consistent symmetry at the quantum level. Fortunately,



**Fig. 6:** Anomalous Feynman diagrams proportional to  $\text{Tr}\{Y^3\}$ . The sum of them must vanish for  $U(1)_Y$  to be a valid symmetry at the quantum level. Hyper-charged chiral fermions  $f$  travel in the loop contributing to a three-hypercharge gauge boson  $B$  vertex. From Ref. [1].

$A = 0$  for each fermion family in the Standard Model. Contributions are from (the factors of 3 are from the different colours of the quarks, whereas the factors of 2 come from the different  $SU(2)_L$  degrees of freedom):

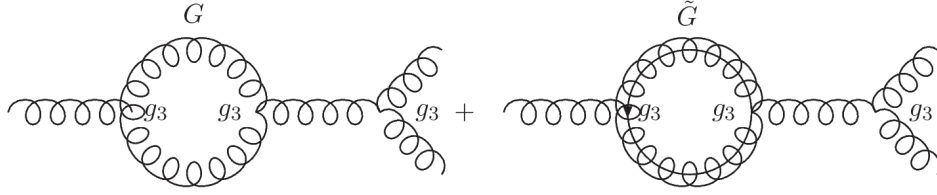
$$\underbrace{3 \times 2 \times (\frac{1}{6})^3}_{Q_L} + \underbrace{3 \times (-\frac{2}{3})^3}_{u_R^c} + \underbrace{3 \times (\frac{1}{3})^3}_{d_R^c} + \underbrace{2 \times (-\frac{1}{2})^3}_{L_L} + \underbrace{1^3}_{e_R^c} = 0.$$

In SUSY, we add the Higgsino doublet  $\tilde{H}_1$ , which yields a non-zero contribution to  $A$ . This must be cancelled by another Higgsino doublet with opposite  $Y$ :  $\tilde{H}_2$ .

There is another special super multiplet sometimes considered to be part of the MSSM with  $B = L = 0$ . This is the *gravity super multiplet*, with the spin  $2\hbar$  graviton and a spin  $3/2\hbar$  gravitino. Usually, after SUSY breaking (see later), the only component of the gravitino that couples with non-negligible strength is its spin  $1/2\hbar$  component.

$$G = (1, 1, 1)$$





**Fig. 7:** Example Feynman diagrams leading to renormalisation of the strong coupling constant  $g_3$ . The left-hand diagram renormalizes the QCD gauge coupling in the Standard Model, whereas in the MSSM, we have additional contributions from supersymmetric particles such as the one on the right-hand side with gluinos in the loop. There are other contributing diagrams, some involving loops of quarks and squarks, for instance.

## 2.2 Interactions

- Gauge couplings are renormalized, which ends up giving them *renormalisation scale dependence*, which matches onto dependence upon the energy scale at which one is probing them:

$$\mu \frac{dg_a(\mu)}{d\mu} = \beta_a g_a^3(\mu), \Rightarrow g_a^{-2}(\mu) = g_a^{-2}(\mu_0) - 2\beta_a \ln \frac{\mu}{\mu_0} \quad (24)$$

where  $\beta_a$  is a constant determined by which particles travel in the loop in the theory. For ordinary QCD it is  $\beta_3 = -7/(16\pi^2)$  whereas for the MSSM, it is  $\beta_3 = -3/(16\pi^2)$  because of additional contributions from squarks and gluinos to the loops, as in Fig. 7.

Eq. 24 is used to extrapolate gauge couplings measured at some energy scale  $\mu_0$  (often taken to be  $M_Z$ , from LEP constraints) to some other scale  $\mu$ . With the SUSY contributions in the MSSM, the gauge couplings almost meet at a renormalisation scale  $E \approx 2 \times 10^{16}$  GeV (see Fig. 8), whereas with just the Standard Model contributions, they do not meet each other at all: see Fig. 4. The meeting of the gauge couplings is a necessary condition for a Grand Unified Theory, which only has one gauge coupling (above  $M_{GUT} \approx 2 \times 10^{16}$  GeV).  $\alpha_1(M_Z)$  and  $\alpha_2(M_Z)$  are both known with high accuracy from the LEP experiments, so we can use them to predict  $M_{GUT} \sim 10^{16}$  GeV and  $\alpha_s(M_Z) = 0.129 \pm 0.002$ . The experimental determination<sup>7</sup> of  $\alpha_s(M_Z) = 0.119 \pm 0.002$ , so the naive prediction is some  $5\sigma$  out. However, this small difference is easily explained by GUT threshold corrections (for example because the  $X$  or  $Y$  bosons are a factor of a few lighter than  $M_{GUT}$  and change the running near the GUT scale) in explicit GUT models.

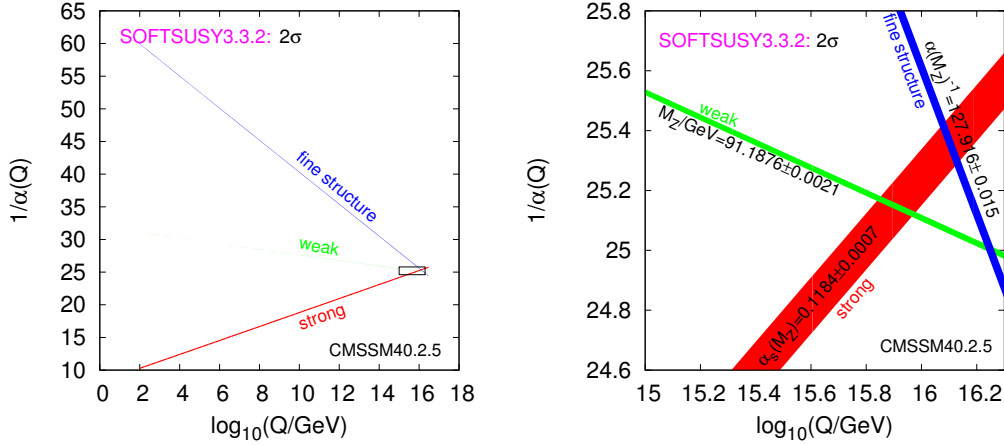
Gauge couplings are renormalized, which ends up giving them *renormalisation scale dependence*, which matches onto dependence upon the energy scale at which one is probing them (one achieves a worse approximation in a truncated perturbation series by picking the renormalisation scale to be vastly different to the energy scales probed in some process): integrating both sides,

$$\mu \frac{dg_a(\mu)}{d\mu} = \beta_a g_a^3(\mu), \Rightarrow g_a^{-2}(\mu) = g_a^{-2}(\mu_0) - 2\beta_a \ln \frac{\mu}{\mu_0} \quad (25)$$

where  $\beta_a$  is a constant determined by which particles travel in the loop in the theory. For ordinary QCD it is  $\beta_3 = -7/(16\pi^2)$  whereas for the MSSM, it is  $\beta_3 = -3/(16\pi^2)$  because of additional contributions from squarks and gluinos to the loops.

- A ‘superpotential’ is like a Lagrangian energy density for SUSY theories: it encodes some of the interactions between the chiral superfields in a way that preserves SUSY. A superpotential term  $W = \lambda\Phi^3$  for a chiral superfield  $\Phi = (\varphi, \psi)$  encodes both a Yukawa interaction  $\mathcal{L} = -\lambda\varphi\psi\psi$  and a scalar interaction  $\mathcal{L} = -|\lambda|^2|\varphi|^4$ , for example.

<sup>7</sup>We quote SM gauge couplings in the  $\overline{MS}$  scheme.



**Fig. 8:** Gauge unification in the MSSM: the thickness of the lines corresponds to the  $2\sigma$  error bars. The right-hand panel shows a zoom of the unification region near  $Q \sim 10^{16}$  GeV.

We write down a superpotential containing all terms which are renormalisable and consistent with our symmetries. If one does this, one obtains two classes of terms,  $W = W_{R_p} + W_{R_{PV}}$ . The terms in  $W_{R_p}$  all conserve baryon number  $B$  and lepton number  $L$ , whereas those in  $W_{R_{PV}}$  break either  $B$  or  $L$ :

$$W_{R_p} = (Y_U)_{ij} Q_i H_2 u_j^c + (Y_D)_{ij} Q_i H_1 d_j^c + Y_E L_i H_1 e_j^c + \mu H_1 H_2 \quad (26)$$

$$W_{R_{PV}} = \lambda_{ijk} L_i L_j e_k^c + \lambda'_{ijk} L_i Q_j d_k^c + \lambda''_{ijk} u_i^c d_j^c d_k^c + \kappa_i L_i H_2, \quad (27)$$

where we have suppressed gauge indices. Since superfields commute in  $W$ ,

$$H_1^a H_1^b \epsilon_{ab} = \frac{1}{2}(H_1^a H_1^b + H_1^b H_1^a) \epsilon_{ab} = \frac{1}{2} H_1^a H_1^b (\epsilon_{ab} + \epsilon_{ba}) = 0 \quad (28)$$

The first three terms in  $W_{R_p}$  correspond to standard Yukawa couplings and give masses to up quarks, down quarks and leptons, as we shall see. Writing  $x = 1, 2, 3$  as a fundamental  $SU(3)$  index,  $a, b = 1, 2$  as fundamental  $SU(2)$  indices, the first term in  $W_{R_p}$  becomes

$$(Y_U)_{ij} Q_i^{xa} H_2^b u_{jx}^c \epsilon_{ab} = (Y_U)_{ij} [u_L^x H_2^0 u_{jx}^c - d_L^x H_2^+ u_{jx}^c]. \quad (29)$$

Once the neutral Higgs component develops a vacuum expectation value,  $H_2^0 \equiv (v_2 + h_2^0)/\sqrt{2}$ , the first term becomes  $(Y_U)_{ij} v_2/\sqrt{2} u_L^x u_{jx}^c + \dots$ , yielding a Dirac mass matrix  $m_u \equiv (Y_U)_{ij} v_2/\sqrt{2}$  for the up quarks. The down quark and lepton masses proceed in an analogous manner. The fourth term is a mass term for the two Higgs(ino) fields.

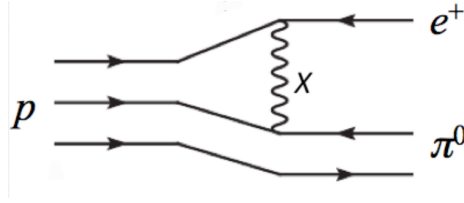
If all of the terms in  $W_{R_{PV}}$  are present, the interaction shown in Fig. 9 would allow proton decay  $p \rightarrow e^+ + \pi^0$  within seconds because

$$\Gamma(p \rightarrow e^+ \pi^0) \approx \frac{\lambda_{11k}^{\prime 2} \lambda_{11k}^{\prime\prime 2}}{16\pi^2 \tilde{m}_{d_k}^4} M_p^5, \quad (30)$$

whereas experiments say that it should be  $> 10^{34}$  years. Alternatively, we could make the RPV couplings very small to make the proton long-lived, by imposing the implied bound on  $\Gamma(p \rightarrow e^+ \pi^0)$ :

$$\lambda'_{11k} \cdot \lambda''_{11k} < 10^{-27} \left( \frac{\tilde{m}_{d_k}}{100 \text{ GeV}} \right)^2. \quad (31)$$





**Fig. 9:** Proton decay  $p \rightarrow e^+\pi^0$  due to baryon- and lepton number violating interactions. Both  $B$  and  $L$  violating terms must be present for the proton to decay. The matrix element is proportional to  $\lambda''_{1j1}{}^* \times \lambda'_{11j}{}^*$ .

In order to forbid proton decay an extra symmetry should be imposed. One symmetry that works is a discrete multiplicative symmetry  $R$  parity defined as

$$R \equiv (-1)^{3(B-L)+2S} = \begin{cases} +1 & : \text{Standard Model particles,} \\ -1 & : \text{superpartners} \end{cases} \quad (32)$$

It forbids all of the terms in  $W_{RPV}$ , but there exist other examples which only ban some subset.

$R$  parity would have important physical implications:

- The lightest superpartner (LSP) is stable, because it is  $R$ -parity odd.
- Cosmological constraints then say that a stable LSP must be electrically and colour-neutral (higgsino, photino, zino). It is then a good candidate for cold weakly interacting dark matter.
- In colliders, the initial state is  $R_p = +1$ , implying that superparticles are produced in pairs. When a superparticle decays, it must decay to another (lighter) superparticle plus some standard model particles.
- One ends up with LSPs at the end of the decays. These do not interact with the detector, and hence appear as unbalanced or ‘missing’ momentum.

Note that the terms in  $W_{RPV}$  can lead to Majorana fermion structure<sup>8</sup>. For instance,  $W = \lambda''_{112} u_1^c d_1^c d_2^c$ : we take the  $F$ -terms as usual in order to find the Lagrangian in terms of components:

$$\mathcal{L} = \frac{1}{2} \left( \lambda''_{112} \tilde{u}_1^* d_{1R}^\dagger C d_{2R}^* - (\lambda''_{112})^* \tilde{u}_1 d_{1R}^T C^* d_{2R} \right)$$

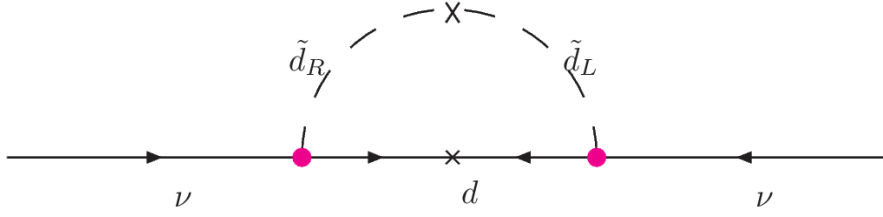
plus supersymmetric copies, where  $C$  is the charge conjugation matrix and  $T$  denotes transpose.

RPV has several potential motivations and characteristics:

- It has many additional search possibilities<sup>9</sup>
- Dark matter changes character: one loses the usual neutralino dark matter candidate. However, the SUSY breaking sector always contains other fields that may be used instead, for example the gravitino or hidden sector fields. Either of these two candidates is so weakly coupled that direct or indirect dark matter detection becomes extremely unlikely, although inference of its production at colliders is still possible.
- Neutrino masses and mixings are generated by the  $L$  violating couplings in diagrams like those in Fig. 10, and the mechanism of their generation is potentially testable at the LHC (unlike, for example, the seesaw mechanism of producing neutrino masses).

<sup>8</sup>This is a familiar structure for people extending the Standard Model to include neutrino masses.

<sup>9</sup>This leads us to a conjecture: *any experimental excess can be explained by RPV SUSY*. We have not found any counter-examples to this yet. This in turn leads to Butterworth’s corollary: *RPV is the last refuge of the ambulance chasing scoundrel*.



**Fig. 10:** RPV generation of neutrino masses and mixings. Here, the dots show the  $L$  violating RPV couplings.

### 2.3 Supersymmetry breaking in the MSSM

An operator called the *supertrace* treats bosonic and fermionic parts of a super multiplet differently. It is defined as

$$\text{STr}\{M^2\} \equiv \sum_j (-1)^{2j+1} (2j+1) m_j^2 = 0, \quad (33)$$

where  $j$  represents the ‘spin’ of the particles in some super multiplet. This is generic for tree level directly broken SUSY. Thus, we cannot break supersymmetry directly in the MSSM, since it preserves  $\text{STr}\{M^2\} = 0$ . Applying this to the photon, say:  $-3m_\gamma^2 + 2m_{\tilde{\gamma}}^2 = 0$ , which would predict a massless photino that hasn’t been observed. Applying it to up quarks:  $2m_u^2 - m_{\tilde{u}_L}^2 - m_{\tilde{u}_R}^2 = 0$ , thus one up squark must be *lighter* than the up quark, again this hasn’t been observed. We introduce a *hidden* sector, which breaks SUSY and has its own fields (which do not directly interact with MSSM fields) and interactions, and an additional *messenger sector* to communicate the SUSY breaking to the observable sector fields:

$$\left( \begin{array}{c} \text{observable} \\ \text{sector, MSSM} \end{array} \right) \longleftrightarrow \left( \begin{array}{c} \text{messenger -} \\ \text{sector} \end{array} \right) \longleftrightarrow \left( \begin{array}{c} \text{hidden} \\ \text{sector} \end{array} \right).$$

This gets around the supertrace rule. There is typically an overall gauge group

$$(SU(3) \times SU(2) \times U(1)) \times G_{\text{SUSY}} \equiv G_{SM} \times G_{\text{SUSY}},$$

where the MSSM fields are singlets of  $G_{\text{SUSY}}$  and the hidden sector fields are singlets of  $G_{SM}$ .

We have already seen several examples of SUSY breaking theories. One popular SUSY-breaking sector in the MSSM context is that of *gaugino condensation*: here, some asymptotically free gauge coupling  $g$  becomes large at some energy scale  $\Lambda$ .  $g$  will renormalize like Eq. 24 with some beta function coefficient. Solving the equation, with  $g^{-2}(\Lambda) \rightarrow 0$ , we obtain  $\Lambda = M \exp[g^{-2}(M)/\beta]$ .  $M$  could be some large scale such as the string scale,  $\sim 5 \times 10^{17}$  GeV. It is easy to arrange for  $\Lambda \ll M$  because of the exponential suppression. When the gauge coupling becomes large, and the theory becomes non-perturbative, one can obtain  $\langle \tilde{g}\tilde{g} \rangle \sim \mathcal{O}(\Lambda^3)$ , breaking SUSY dynamically<sup>10</sup>.

The SUSY breaking fields have couplings with the messenger sector, which in turn have couplings with the MSSM fields, and carry the SUSY breaking over to them. There are several possibilities for the messenger sector fields, which may determine the explicit form of SUSY breaking terms in the MSSM, including (note here that  $M_{\text{SUSY}}$  is the SUSY breaking in the hidden sector, whereas  $\Delta m$  is the SUSY breaking that ends up in the MSSM fields):

- gravity mediated SUSY

If the mediating field couples with gravitational strength to the standard model, the couplings are suppressed by the inverse Planck mass  $M_{\text{pl}}$ , the natural scale of gravity. The SUSY breaking mass splitting between MSSM particles and superparticles,  $\Delta m$ , becomes

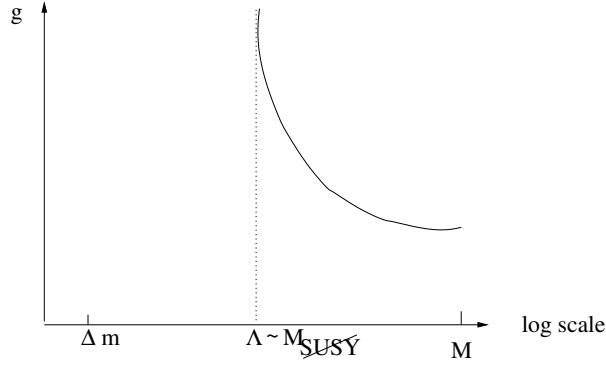
$$\Delta m = \frac{M_{\text{SUSY}}^2}{M_{\text{pl}}}. \quad (34)$$

<sup>10</sup>Here,  $\tilde{g}$  is the gaugino of the hidden sector gauge group, and  $\beta$  is the hidden gauge group beta function coefficient.

We want  $\Delta m \approx 1\text{TeV}$  and we know that  $M_{\text{pl}} \approx 10^{19}\text{GeV}$ , so

$$M_{\text{SUSY}} = \sqrt{\Delta m \cdot M_{\text{pl}}} \approx 10^{11}\text{GeV}. \quad (35)$$

The gravitino gets a mass  $m_{\frac{3}{2}}$  of  $\Delta m$  order TeV from the ‘super Higgs mechanism’.



**Fig. 11:** Gaugino condensation and supergravity mediated SUSY breaking. From Ref. [1].

– gauge mediated **SUSY**

Messenger fields are charged under both  $G_{\text{SM}}$  and  $G_{\text{SUSY}}$ . Gauge loops transmit SUSY breaking to the MSSM fields. Thus,  $\Delta m \sim M_{\text{SUSY}}/(16\pi^2)$  is required to be of order TeV. In this case, the gravitino mass  $m_{\frac{3}{2}} \sim \frac{M_{\text{SUSY}}^2}{M_{\text{pl}}} \sim \text{eV}$  and the gravitino is the LSP.

– anomaly mediated **SUSY**

In this case, the auxiliary fields of supergravity get a vacuum expectation value. The effects are always present, but suppressed by loop factors. They may be dominant if the tree-level contribution is suppressed for some reason.

Each of these scenarios has phenomenological advantages and disadvantages and solving their problems is an active field of research. In all scenarios, the Lagrangian for the observable sector has contributions

$$\mathcal{L} = \mathcal{L}_{\text{SUSY}} + \mathcal{L}_{\text{SUSY}}. \quad (36)$$

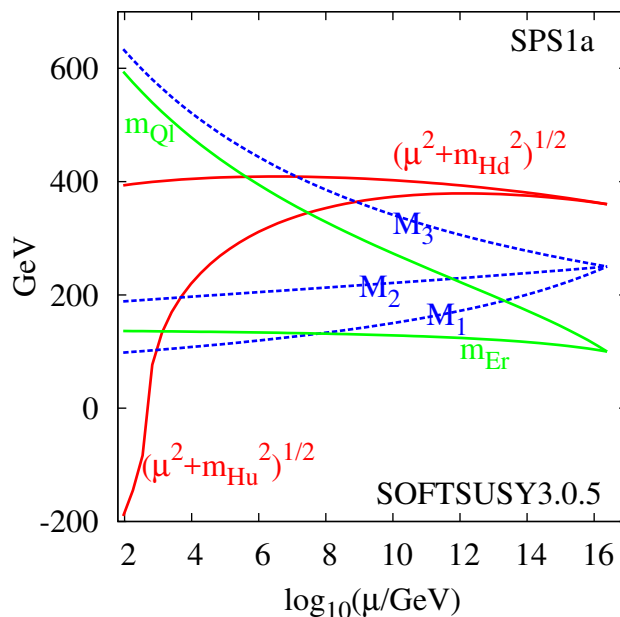
In the second term, we write down all renormalisable symmetry invariant terms which do not reintroduce the hierarchy problem. They are of the form (where  $i$  and  $j$  label different fields):

$$\mathcal{L}_{\text{SUSY}} = \underbrace{m_{ij}^2 \varphi_i^* \varphi_j + m'_{ij}{}^2 (\varphi_i \varphi_j + h.c.)}_{\text{scalar masses}} + \left( \underbrace{\frac{1}{2} M_\lambda \lambda \lambda}_{\text{gaugino masses}} + \underbrace{A_{ijk} \varphi_i \varphi_j \varphi_k}_{\text{trilinear couplings}} + h.c. \right). \quad (37)$$

$M_\lambda, m'_{ij}, m_{ij}^2, A_{ijk}$  are called *soft SUSY breaking terms*: they do not reintroduce quadratic divergences into the theory. Particular forms of SUSY breaking mediation can give relations between the different soft SUSY breaking terms. They determine the amount by which supersymmetry is expected to be broken in the observable sector, and the masses of the superparticles for which the LHC is searching.

Explicitly, we parameterize all of the terms that softly break SUSY in the  $R_p$  preserving MSSM, suppressing gauge indices:

$$\begin{aligned} \mathcal{L}_{R_p}^{\text{SUSY}} = & (A_U)_{ij} \tilde{Q}_{Li} H_2 \tilde{u}_{Rj}^* + (A_D)_{ij} \tilde{Q}_{Li} H_1 \tilde{d}_{Rj}^* + (A_E)_{ij} \tilde{L}_{Li} H_1 \tilde{e}_{Rj}^* + \\ & \tilde{Q}_{Li}^* (m_{\tilde{Q}}^2)_{ij} \tilde{Q}_{Lj} + \tilde{L}_{Li}^* (m_{\tilde{L}}^2)_{ij} \tilde{L}_{Lj} + \tilde{u}_{Ri} (m_{\tilde{U}}^2)_{ij} \tilde{u}_{Rj}^* + \tilde{d}_{Ri} (m_{\tilde{D}}^2)_{ij} \tilde{d}_{Rj}^* + \tilde{e}_{Ri} (m_{\tilde{E}}^2)_{ij} \tilde{e}_{Rj}^* + \end{aligned}$$



**Fig. 12:** An example of renormalisation in the MSSM using the program SOFTSUSY [4] to calculate the renormalisation. A particular high energy theory is assumed, which has GUT symmetry and implies that the gauginos are all mass degenerate at the GUT scale. The scalars (e.g the right-handed electron  $Er$  and the left-handed squarks  $Ql$ ) are also mass-degenerate at the GUT scale. Below the GUT scale though, the masses split and renormalize separately. When we are scattering at energies  $\sim O(100)$  GeV, it is a good approximation to use the masses evaluated at that renormalisation scale  $\mu \approx E$ . We see that one of the Higgs mass squared parameters,  $\mu^2 + M_{Hu}^2$ , becomes negative at the electroweak scale, triggering electroweak symmetry breaking.

$$(m_3^2 H_1 H_2 + h.c.) + m_1^2 |H_1|^2 + m_2^2 |H_2|^2 + \frac{1}{2} M_3 \tilde{g} \tilde{g} + \frac{1}{2} M_2 \tilde{W} \tilde{W} + \frac{1}{2} M_1 \tilde{B} \tilde{B}.$$

Sometimes,  $m_3^2$  is written as  $\mu B$ . Often, specific high scale models provide relations between these many parameters. For instance, the Constrained MSSM (which may come from some particular string theory or other field theory) specifies the constraints

$$\begin{aligned} M_1 &= M_2 = M_3 =: M_{1/2} \\ m_Q^2 &= m_L^2 = m_{\tilde{U}}^2 = m_D^2 = m_E^2 \equiv m_0^2 I_3 \\ m_1^2 &= m_2^2 = m_0^2 \\ A_U &= A_0 Y_U, \quad A_D = A_0 Y_D, \quad A_E = A_0 Y_E \end{aligned}$$

where  $I_3$  is the 3 by 3 identity matrix. Thus in the ‘CMSSM’, we reduce the large number of free SUSY breaking parameters down to<sup>11</sup> 3:  $M_{1/2}$ ,  $m_0$  and  $A_0$ . These relations hold at the GUT scale, and receive large quantum corrections, as Fig. 12 shows.

## 2.4 States after electroweak symmetry breaking

With two complex Higgs doublets, we count 8 real degrees of freedom. 3 of these are ‘eaten’ by the longitudinal components of the  $W^\pm$  and  $Z^0$  bosons, leaving a total of five physical Higgs fields: two

<sup>11</sup>One should really include  $\tan \beta = v_2/v_1$  as well, the ratio of the two Higgs vacuum expectation values.

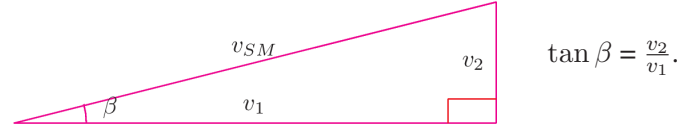
$CP$ -even (in mass order)  $h^0, H^0$ , one  $CP$ -odd  $A^0$  and two charged Higgs'  $H^\pm$ . The other SUSY particles that have identical quantum numbers under  $QED \times QCD$  mix after electroweak symmetry breaking: for example the bino, wino, and two neutral Higgsinos mix. Their mass eigenstates are called *neutralinos*, conventionally written in order of their masses  $\chi_{1,2,3,4}^0$ .  $\chi_1^0$  typically has a special status in that is a good candidate for dark matter if it is the *lightest supersymmetric particle* and  $R_p$  is conserved. The scalar partner of the left-handed top (called the 'left-handed stop') mixes with the right-handed stop to form two mass eigenstates:  $\tilde{t}_{1,2}$ . This analogously occurs for the sbottoms and staus as well. The charged Higgsinos mix with the winos to form mass eigenstates called 'charginos':  $\chi_{1,2}^\pm$ .

## 2.5 The Neutral Higgs Potential

Both Higgs' of the MSSM acquire vacuum expectation values:

$$\begin{pmatrix} H_1^0 \\ H_1^- \end{pmatrix} \rightarrow \begin{pmatrix} v_1 \\ 0 \end{pmatrix} \quad \begin{pmatrix} H_2^+ \\ H_2^0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ v_2 \end{pmatrix} \quad (38)$$

and to get the value of  $M_W$  to match with experimental data, we require  $v_{SM} = 246$  GeV. In a two-Higgs doublet model, this leads to the following construction:



$\tan \beta$  is a parameter which changes the phenomenology of the model because the third family Yukawa couplings depend upon it, and they are comparatively large dimensionless couplings. The Yukawa terms from the MSSM superpotential are:

$$\mathcal{L} = h_t \bar{t}_L H_2^0 t_R + h_b \bar{b}_L H_1^0 b_R + h_\tau \bar{\tau}_L H_1^0 \tau_R + \text{H.c.} + \dots \quad (39)$$

$$\Rightarrow \frac{m_t}{\sin \beta} = \frac{h_t v_{SM}}{\sqrt{2}}, \quad \frac{m_{b,\tau}}{\cos \beta} = \frac{h_{b,\tau} v_{SM}}{\sqrt{2}}, \quad (40)$$

after electroweak symmetry breaking and the neutral components of Higgs' are replaced by their vacuum expectation values:  $H_i^0 = (v_i^0 + H_i^0)/\sqrt{2}$ .

Picking out only the terms involving the neutral Higgs fields  $H_1^0$  and  $H_2^0$ , we have the neutral Higgs potential

$$V = (|\mu|^2 + m_{H_2}^2)|H_2^0|^2 + (|\mu|^2 + m_{H_1}^2)|H_1^0|^2 - \mu B(H_2^0 H_1^0 + \text{H.c.}) + \frac{1}{8}(g^2 + g'^2)(|H_2^0|^2 - |H_1^0|^2)^2. \quad (41)$$

The vacuum minimizes this potential with respect to both of the neutral components:

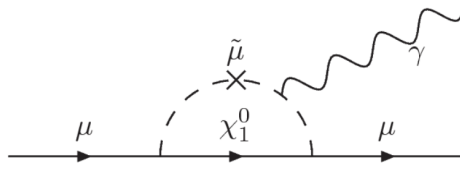
$$\frac{\partial V}{\partial H_2^0} = \frac{\partial V}{\partial H_1^0} = 0 \Rightarrow \mu B = \frac{\sin 2\beta}{2}(\bar{m}_{H_1}^2 + \bar{m}_{H_2}^2 + 2\mu^2), \mu^2 = \frac{\bar{m}_{H_1}^2 - \bar{m}_{H_2}^2 \tan^2 \beta}{\tan^2 \beta - 1} - \frac{M_Z^2}{2}. \quad (42)$$

These two conditions should be used to eliminate two of the MSSM's free parameters: often,  $|\mu|$  and  $B$  (although note that the sign of  $\mu$  is physical and not determined by Eq. 42).

## 2.6 Pros and Cons of the MSSM

We start with a list of unattractive features of the MSSM:

- There are  $\sim 100$  extra free parameters in the SUSY breaking sector, making for a complicated parameter space.



**Fig. 13:** Example one-loop diagram of particles contributing to  $(g - 2)_\mu$ .

- Nearly all of this parameter space is ruled out by flavour physics constraints: SUSY particles could heavily mix in general, then this mixing could appear in loops and make the quarks mix in a flavour changing neutral current, upon which there are very strong experimental bounds. It could be that this clue is merely telling us that there is more structure to the MSSM parameter space, though (like in the CMSSM).
- The  $\mu$  problem.  $\mu$  in  $W_{R_p}$  must be  $< \mathcal{O}(1)$  TeV, since it contributes at tree-level to  $m_h$ . Why should this be, when in principle we could put it to be  $\sim \mathcal{O}(M_{Pl})$ , because it does not break any SM symmetries? (Note though that once it is set to be small at tree-level, SUSY protects it from large quantum corrections).
- As lower limits on sparticle masses increase, the extent to which SUSY solves the hierarchy problem decreases.

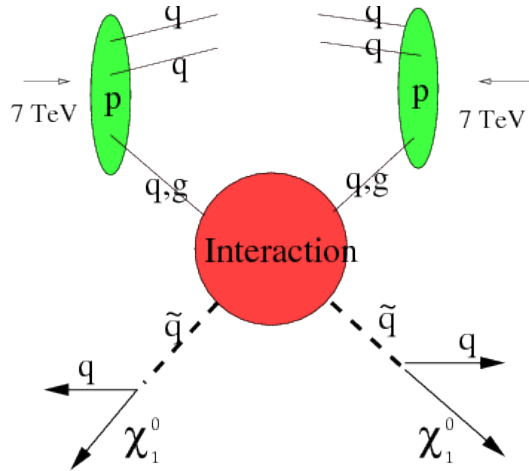
These SUSY problems can be solved with further model building.

We close with an ordered list of weak-scale SUSY's successes:

- SUSY solves the technical hierarchy problem.
- Gauge unification works.
- The MSSM contains a viable dark matter candidate, if  $R_p$  is conserved.
- Electroweak symmetry breaks radiatively.
- A one-loop diagram involving sneutrinos and charginos (and one involving smuons and neutralinos: see Fig. 13) contribute to the anomalous magnetic moment of the muon, and may solve the discrepancy between SM predictions and experimental measurements in Eq. 12.

## 2.7 LHC Production of SUSY Particles

One turns the energy of the LHC beams into mass via  $E = mc^2$ , hoping to produce pairs (if  $R_p$  is conserved) of SUSY particles that were too heavy to have been previously produced in lower energy machines. We show a schematic in Fig. 14: occasionally, high energy constituents of the proton (called ‘partons’: quarks or gluons) will collide, as in the figure. The idea is that these are most likely to make strongly interacting particles, all other things being equal (in the figure, we have the example of squark production). The rest of the broken protons typically will be boosted along the beam-line. The sparticles undergo subsequent decay (in the example in the figure, into a quark - which will form a jet of hadrons and a dark matter particle: the lightest neutralino). Since we have assumed  $R_p$  to be conserved, the  $\chi_1^0$  is stable but since it is weakly interacting, it passes through the rest of the detector without any interactions, stealing momentum from the collision. The decays of the initial pair of sparticles may be much more complex, going through cascade decays where at each stage there is a lighter sparticle and a SM particle produced.  $R_p$  conserving SUSY provides an example of how any dark matter candidate that is light enough and that (perhaps indirectly) couples to protons can be produced in LHC collisions. Jets and missing transverse momentum  $\vec{p}_T^{\text{miss}}$  (sometimes this is known under the misnomer ‘missing energy’) form a classic SUSY search, but also jets plus varying numbers of leptons (from sparticle



**Fig. 14:** Picture of the production of sparticles at a 14 TeV LHC.

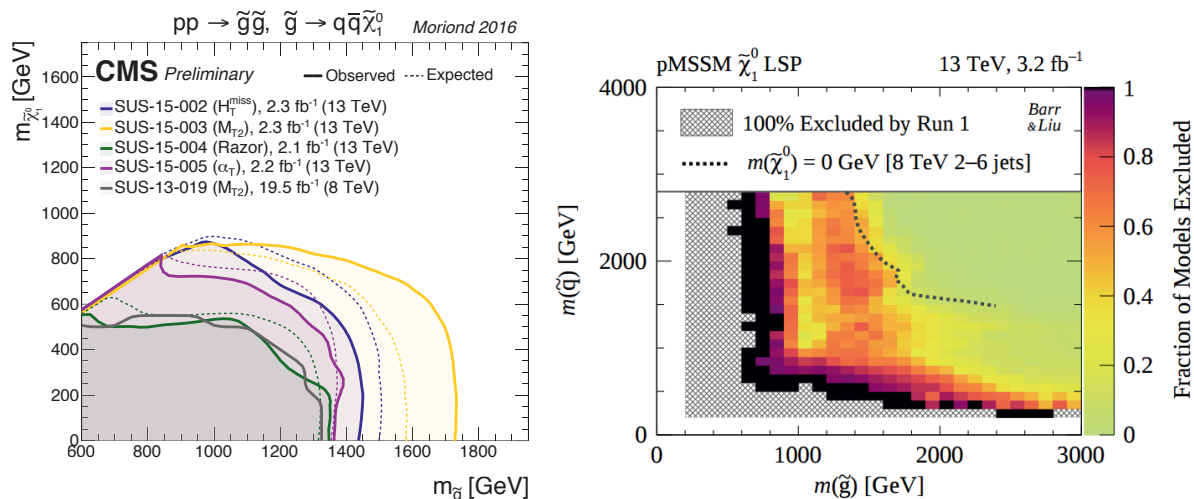
cascade decays) plus missing transverse momentum form another well-studied class. There is a SUSY monojet signature [5], although sparticles would likely be found in one of the other production channels first because the monojet signature is due to a strong times an electroweak matrix element. In the case of gauge mediated SUSY breaking models, the lightest neutralino may decay into a gravitino plus a photon, or a  $Z^0$ , and so for instance di-photon plus missing transverse momentum searches form another class. Since one obtains additional jets from showering off the initial state at the LHC, searches are often inclusive, meaning that one only selects a *minimum* number of hard jets.

Often, searches are interpreted in terms of ‘simplified models’: for instance, one studies gluino pair production, then assumes that each decays into 2 jets and missing transverse momentum: see Fig. 15. However, current bounds based on simplified models [7] often give much stronger bounds than in a more general MSSM set-up [8]. This is because simplified models tend to only assume a single decay mode of one sparticle (or a few decay modes of particular sparticles), whereas in full models there can be literally thousands of active decay chains, diluting the signal between many different search channels such that no one shows an excess. There are also cases of somewhat ‘compressed spectra’: when sparticles in decay chains are similar in mass, energy-momentum conservation means that they tend to produce fairly soft SM particles, which often fail analysis cuts. Because they are not dependent on the many MSSM parameters, simplified searches are very convenient for searches, being less model dependent. However, exclusion limits from simplified models are not easy to interpret in more realistic models, and tend to be far too restrictive unless one interprets them with care. In Fig. 15, we see this in action: for massless neutralinos, gluinos up to 1750 GeV are ruled out in the simplified model, whereas in a (more realistic) phenomenological MSSM approximation, we see that gluinos of 800 GeV are still allowed for some points.

### 3 Extra Dimensions

For a review of extra dimensions and their phenomenology, see Ref. [10]. As mentioned above, extra dimensions correspond to an expansion of the Poincaré symmetry: there are additional generators associated with translation invariance in each extra spatial dimension. Superstring theory also requires them in addition to supersymmetry for internal consistency, but any theory incorporating them must explain why we only observe 3+1 (i.e. three space-like and one time-like). There are a couple of possibilities to ‘hide’ the extra dimensions from our perception:





**Fig. 15:** Examples of interpretation of search limits: the left-hand panel shows simplified model exclusions from  $2.3\text{fb}^{-1}$  of 13 TeV LHC analyses for di-gluino production and is from Ref. [6]. On the right-hand side, we see a more complete description in terms of the MSSM interpreting  $3.2\text{fb}^{-1}$  of integrated luminosity and is from Ref. [9]. The simplified model exclusion on the left hand panel is that gluinos up to 1750 GeV are excluded (for zero neutralino mass) whereas on the right-hand panel, we see that models exist where gluinos of 800 GeV are allowed: these pass a list of negative searches for SUSY at 13 TeV in ATLAS. The ‘fraction of models excluded’ starts from a list of models in MSSM parameter space that had good dark matter properties, and otherwise passed the constraints from Run I. If there are *any* points with the masses listed on the axis, the fraction of models excluded is less than 1.

- *We are stuck on a brane:* meaning that the bulk of space-time has more than 3+1 dimensions, but SM fields are stuck on a 3+1 dimensional hypersurface: a ‘brane’. Gravity travels wherever space-time is, so that it *must* feel the effect of the additional dimensions. That’s because gravity is described by a quantum fluctuation of the metric, and the bulk metric is defined in the bulk space-time.
- *The extra dimensions are curled up on themselves:* each point in our 3+1 dimensional space time has a circle, or some other compact manifold, where one can travel – albeit periodically – in the extra dimensions, which are in an orthogonal direction to all of the other dimensions. If such manifolds are not too large (less than a millimeter, certainly), then current experimental bounds upon gravitational forces acting at relatively small distances may still not rule the model out.

We illustrate the two cases in Fig. 16. In the figure, we have taken the example of string theory to illustrate the brane case, but it is essentially valid in the field theory limit as well: SM fields may be confined to a hypersurface of the bulk space-time, whereas gravity travels everywhere.

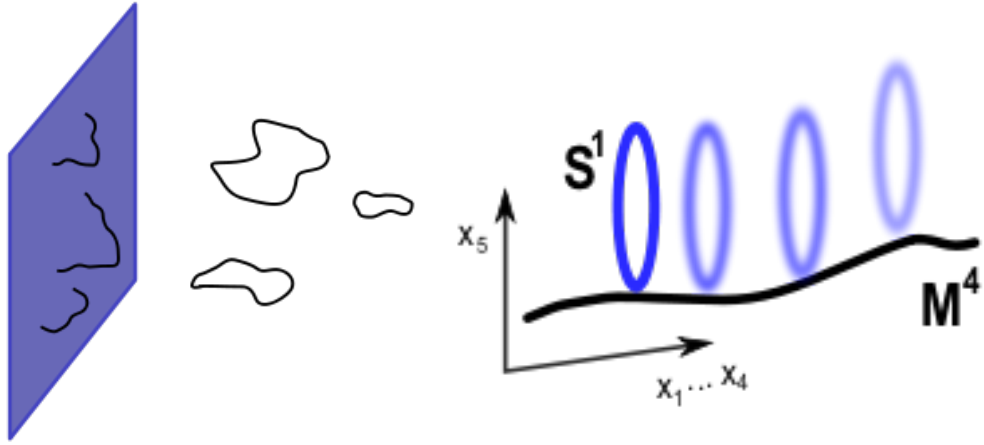
### 3.1 Compactification and a Scalar Field in 5 Dimensions

Taking compactified extra dimensions as an example, consider a massless five dimensional (5D) scalar field (i.e. a scalar field living in a 5-dimensional bulk space-time)  $\varphi(x^M)$ ,  $M = 0, 1, \dots, 4$  with action

$$\mathcal{S}_{5D} = \int d^5x \partial^M \varphi \partial_M \varphi. \quad (43)$$

We single the extra dimension out by calling it  $x^4 = y$ .  $y$  defines a circle of radius  $r$  with  $y \equiv y + 2\pi r$ . Our space time is now  $\mathbb{M}_4 \times S^1$ . Periodicity in the  $y$  direction implies that we may perform a discrete





**Fig. 16:** Picture of different extra-dimensional set-ups: the brane (on the left), where in string theory SM states appear as open strings whose ends end upon the brane but gravitons appear as close string states in the bulk, or compactification (on the right), in this example we have taken the example of a circle  $S^1$  times ordinary 4-dimensional Minkowski space  $M^4$ .

Fourier expansion

$$\varphi(x^\mu, y) = \sum_{n=-\infty}^{\infty} \varphi_n(x^\mu) \exp\left(\frac{iny}{r}\right). \quad (44)$$

Notice that the Fourier coefficients are functions of the standard 4D coordinates and therefore are (an infinite number of) 4D scalar fields. The equations of motion for the Fourier modes are the (in general massive) Klein-Gordon wave equations

$$\begin{aligned} \partial^M \partial_M \varphi = 0 &\Rightarrow \sum_{n=-\infty}^{\infty} \left( \partial^\mu \partial_\mu - \frac{n^2}{r^2} \right) \varphi_n(x^\mu) \exp\left(\frac{iny}{r}\right) = 0 \\ &\Rightarrow \boxed{\partial^\mu \partial_\mu \varphi_n(x^\mu) - \frac{n^2}{r^2} \varphi_n(x^\mu) = 0.} \end{aligned} \quad (45)$$

These are then an infinite number of Klein Gordon equations for massive 4D fields. This means that each Fourier mode  $\varphi_n$  is a 4D particle with mass  $m_n^2 = \frac{n^2}{r^2}$ . Only the zero mode ( $n = 0$ ) is massless. One can visualize the states as an infinite tower of massive states (with increasing mass proportional to  $n$ ). This is called a *Kaluza Klein tower* and the massive states ( $n \neq 0$ ) are called *Kaluza Klein-states* or *momentum states*, since they come from the momentum in the extra dimension:

In order to obtain the effective action in 4D for all these particles, let us plug the mode expansion of  $\varphi$  Eq. 44 into the original 5D action Eq. 43:

$$\begin{aligned} \mathcal{S}_{5D} &= \int d^4x \int dy \sum_{n=-\infty}^{\infty} \left( \partial^\mu \varphi_n(x^\mu) \partial_\mu \varphi_n(x^\mu)^* - \frac{n^2}{r^2} |\varphi_n|^2 \right) \\ &= 2\pi r \int d^4x \left( \partial^\mu \varphi_0(x^\mu) \partial_\mu \varphi_0(x^\mu)^* + \dots \right) = 2\pi r \mathcal{S}_{4D} + \dots \end{aligned}$$

This means that the 5D action reduces to one 4D action for a massless scalar field plus an infinite sum of massive scalar actions in 4D. If we are only interested in energies smaller than the  $\frac{1}{r}$  scale, we may concentrate only on the action of the massless mode.



**Fig. 17:** The Kaluza Klein tower of massive states due to an extra  $S^1$  dimension. Masses  $m_n = |n|/r$  grow linearly with the fifth dimension's wave number  $n \in \mathbb{Z}$ .

### 3.2 Compactification of a Vector Field in 5 Dimensions

Vector fields are decomposed in a completely analogous way:  $\{A_M\} = \{A_\mu, A_4 = \phi\}$ . Consider the action

$$\mathcal{S}_{5D} = \int d^5x \frac{1}{g_{5D}^2} F_{MN} F^{MN} \quad (46)$$

with a field strength

$$F_{MN} = \partial_M A_N - \partial_N A_M \quad (47)$$

implying

$$\partial^M \partial_M A_N - \partial^M \partial_N A_M = 0. \quad (48)$$

If we now choose a gauge, e.g. the transverse gauge:

$$\partial^M A_M = 0, A_0 = 0 \Rightarrow \partial^M \partial_M A_N = 0, \quad (49)$$

then this obviously becomes equivalent to the scalar field case (for each component  $A_M$ ) indicating an infinite tower of massive states for each massless state in 5D. In order to find the 4D effective action we once again plug this into the 5D action:

$$\begin{aligned} \mathcal{S}_{5D} &\mapsto \mathcal{S}_{4D} \\ &= \int d^4x \left( \frac{2\pi r}{g_{5D}^2} F_{(0)}^{\mu\nu} F_{(0)\mu\nu} + \frac{2\pi r}{g_{5D}^2} \partial_\mu \rho_0 \partial^\mu \rho_0 + \dots \right). \end{aligned}$$

Therefore we end up with a 4D theory of a massless gauge particle  $F_{(0)}^{\mu\nu}$ , a massless scalar  $\rho_0$  from the massless Kaluza-Klein state of  $\phi$  and infinite towers of massive vector and scalar fields. Notice that the gauge couplings of 4- and 5 dimensional actions (coefficients of  $F_{MN}F^{MN}$  and  $F_{\mu\nu}F^{\mu\nu}$ ) are related by

$$\frac{1}{g_4^2} = \frac{2\pi r}{g_5^2}. \quad (50)$$

In  $D$  space time dimensions, this generalizes to

$$\frac{1}{g_4^2} = \frac{V_{D-4}}{g_D^2} \quad (51)$$

where  $V_n$  is the volume of the  $n$  dimensional compact space (e.g. an  $n$  sphere of radius  $r$ ).

### 3.2.1 The electric (and gravitational) potential

We apply Gauss' law for the electric field  $\vec{E}$  and the potential  $\Phi$  of a point charge  $Q$ :

$$\oint_{S^2} \vec{E} \cdot d\vec{S} = Q \Rightarrow \|\vec{E}\| \propto \frac{1}{R^2}, \Phi \propto \frac{1}{R} : 4D$$

$$\oint_{S^3} \vec{E} \cdot d\vec{S} = Q \Rightarrow \|\vec{E}\| \propto \frac{1}{R^3}, \Phi \propto \frac{1}{R^2} : 5D$$

Thus, the apparent behaviour of the force depends upon whether we are sensitive to the extra dimension or not: if we test the force at distances smaller than its size (i.e. at energies high enough to probe such small distance scales), it falls off as  $1/R^3$ : the field lines have an extra dimension to travel in. If we test the force at larger distances than the size of the extra dimension, we obtain the usual  $1/R^2$  law.

In  $D$  space time dimensions

$$\|\vec{E}\| \propto \frac{1}{R^{D-2}}, \Phi \propto \frac{1}{R^{D-3}}. \quad (52)$$

If one dimension is compactified (radius  $r$ ) like in  $\mathbb{M}_4 \times S^1$ , then we have two limits

$$\|\vec{E}\| \propto \begin{cases} \frac{1}{R^3} & : R < r \\ \frac{1}{R^2} & : R \gg r \end{cases}. \quad (53)$$

Analogous arguments hold for gravitational fields and their potentials, but we shall not detail them here, preferring instead to sketch the resulting field content.

### 3.2.2 Sketch of Compactified Gravitation

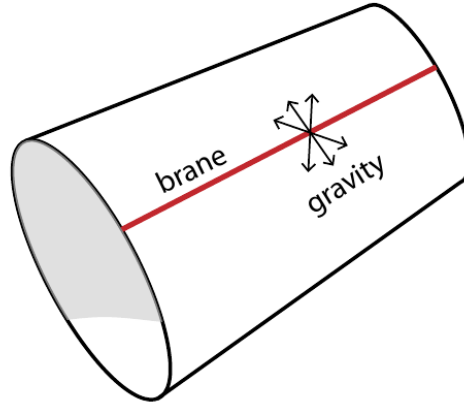
The spin  $2\hbar$  graviton  $G_{MN}$  becomes the 4D graviton  $g_{\mu\nu}$ , some gravivectors  $G_{\mu n}$  and some graviscalars  $G_{mn}$  (where  $m, n = 4, \dots, D-1$ ), along with their infinite Kaluza-Klein towers. The Planck mass squared  $M_{Pl}^2 = M_D^{D-2} V_{D-4} \sim M_D^{D-2} r^{D-4}$  is a derived quantity. Fixing  $D$ , we can fix  $M_D$  and  $r$  to get the correct result for  $M_{Pl} \sim 10^{19}$  GeV. So far, we require  $M_D > 1$  TeV and  $r < 10^{-16}$  cm from Standard Model measurements since no significant confirmed signature of extra dimensions has been seen at the time of writing.

## 3.3 Brane Worlds

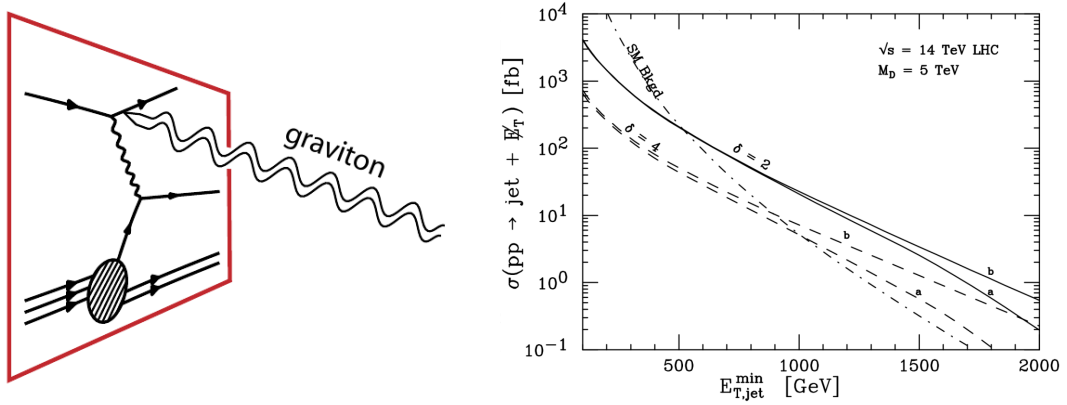
In the brane world scenario, we are trapped on a 3+1 surface in a  $D+1$  dimensional bulk space-time (see Fig. 18). There are two cases here: *large extra dimensions* and warped space-times. Since gravity itself is so weak, the constraints on brane world scenarios are quite weak: the extra dimension is constrained to be of a size  $r < 0.1$  mm or so, potentially much larger than the  $10^{-16}$  cm of the Standard Model, hence the name *large extra dimensions*.

### 3.3.1 Large extra dimensions

There is the possibility to try to solve the hierarchy problem with the large extra dimensions scenario if we put  $M_D \sim 1$  TeV. The idea is that this *is* the fundamental scale: there is no high scale associated with  $M_{Pl}$  fundamentally - it is an illusion caused by the presence of the extra dimensions. In 5D for example,  $M_{Pl}^2 = M_D^{D-2} V_{D-4} \Rightarrow r \sim 10^8$  km, clearly ruled out by observations. Already in 6D though,  $r = 0.1$  mm - consistent with experiments that measure the gravitational force on small distance scales. This rephrases the hierarchy problem to the question "why are the extra dimensions so large compared with  $10^{-16}$  cm?"



**Fig. 18:** Force field lines feel the effect of the extra dimensions. Here we show a 3+1 dimensional brane, where gravity spreads into the extra dimension and feels its effect.



**Fig. 19:** Example production of  $\vec{p}_T^{\text{miss}}$  at a 14 TeV LHC through production of a Kaluza-Klein tower of graviton states that propagate off into the bulk. The left-hand panel shows a heuristic picture: the red hypersurface representing the brane, and the graviton tower being emitted into the bulk. In the right-hand panel, it is shown how the cross-section varies with the transverse momentum of the jet ( $E_{T,\text{jet}}^{\text{min}}$ ) for the SM background, and the case of  $d$  extra dimensions. The lines a (b) are constructed by integrating the cross-section over  $\hat{s} < M_D^2$  (all  $\hat{s}$ ), respectively. Both the diagram and the plot are from Ref. [11].

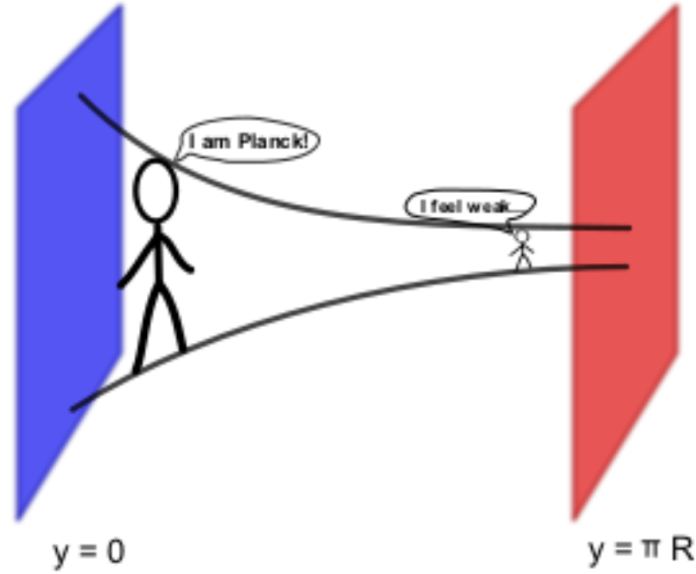
Graviton phenomenology: each Kaluza-Klein mode couples weakly  $\propto 1/M_{Pl}$ , but there are so many modes that after summing over them, you end up with  $1/M_D$  suppression only! One can approximate them by a *continuum* of modes with a cut-off. The graviton tower propagates into the bulk and takes away missing momentum leading to a  $pp \rightarrow j + \vec{p}_T^{\text{miss}}$  signature (for example) by the process shown in Fig. 19.

### 3.3.2 Warped (or ‘Randall-Sundrum’ space-times

Warped space-times are where the metric exponentially warps along the extra dimension  $y$ :

$$ds^2 = e^{-|ky|} \eta_{\mu\nu} dx^\mu dx^\nu + dy^2. \quad (54)$$

The metric changes from  $y = 0$  to  $y = \pi r$  via  $\eta_{\mu\nu} \mapsto e^{-k\pi r} \eta_{\mu\nu}$ . Here, we set  $M_D = M_{Pl}$ , but this gets



**Fig. 20:** Picture of the Randall-Sundrum I set-up. On the left-hand side at  $y = 0$  we have the Planck brane, which is warped down to the weak brane at the right hand side ( $y = \pi R$ ). The idea is that the Higgs boson (and some other fields) are localized on the weak brane.

warped down to the weak brane:

$$\Lambda_\pi \sim M_{Pl} e^{-k\pi r} \sim \mathcal{O}(\text{TeV}), \quad (55)$$

if  $r \sim 10/k$ . Here,  $k$  is of order  $M_{Pl}$  and so we have a *small extra dimension*, but the warping explains the smallness of the weak scale. Note that we still have to stabilize the separation between the branes, which can involve extra tuning unless extra structure is added to the model.

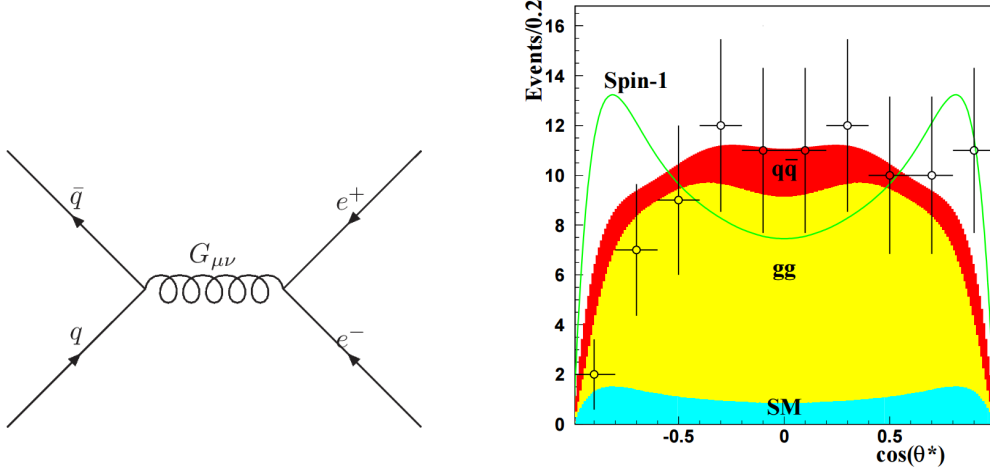
The interaction Lagrangian is

$$\mathcal{L}_I = -G^{\mu\nu} T_{\mu\nu} / \Lambda_\pi, \quad (56)$$

where  $T_{\mu\nu}$  is the *stress energy tensor*, containing products of the other Standard Model fields.  $\Lambda_\pi \sim \mathcal{O}(\text{TeV})$ , so the interaction leads to electroweak-strength cross sections, not gravitationally suppressed ones. Thus, the LHC can produce the resonance: one will tend to produce the lightest one most often, as it is less suppressed by parton distribution functions. The ratios of masses of higher modes are given by zeros of Bessel functions, so they are not as regular as they are in large extra dimensions.

Randall-Sundrum phenomenology: one looks for the TeV scale first resonances, which are weakly coupled to Standard Model states. If only gravity travels in the extra dimensions, then the resonance is the ‘Randall-Sundrum graviton’: it has universal coupling to all particles via Eq. 56 and so it can decay into  $q\bar{q}$ ,  $WW$ ,  $ZZ$ ,  $\gamma\gamma$ ,  $gg$ ,  $l^+l^-$  or  $h^0h^0$  with branching ratios that are of a similar order of magnitude to each other. Flavour considerations imply that this isn’t the end of the story: one requires additional flavour structure, otherwise the model violates flavour bounds from experiment. One common way of adding flavour structure is to allow the other particles into the bulk, but have different profiles of fermions in the bulk, leading to different overlaps with the weak brane, where the Higgs field is localized (the overlap would be proportional to the particle in question’s Yukawa coupling). In this case, one could look for the first Kaluza Klein modes of gauge bosons and fermions, too.

Kaluza Klein modes that have masses that are heavier than the centre of mass energy of the beams may also be looked for via their virtual effects. Searching for particles that mediate interactions that are



**Fig. 21:** Production of an RS graviton and subsequent decay into  $e^+e^-$ : both figures are taken from Ref. [12]. On the right-hand side, the angular distribution of the electron with respect to the beam line in the rest-frame ( $\theta^*$ ) of  $G_{\mu\nu}$  is shown for a 1.5 TeV graviton. The different shaded colours show the contribution from  $q\bar{q}$  collisions, from  $gg$  and from other SM processes. This is contrasted against a straw-man spin  $1\hbar$  distribution in the green line. The typical expected size of statistical uncertainties resulting from  $100 \text{ fb}^{-1}$  of integrated luminosity at a 14 TeV LHC is shown on the points. This would be enough to discriminate against the spin  $1\hbar$  hypothesis, which is much more forward than the spin  $2\hbar$  hypothesis.

occurring at collisions with less energy than their mass has been historically very important (particularly in terms of the weak interactions which were indirectly observed before the discovery of the  $W^\pm$  and  $Z^0$  bosons). Such a kinematic situation can be approximated by *effective field theories*, which in turn reduces model dependence. We now sketch effective field theories, along with caveats pertinent to their use.

#### 4 Effective Field Theories

At low momenta  $p^\mu$ , we can model the effects of particles with a much heavier mass  $M^2 \gg p^2$  and a small width  $\Gamma \ll M$  with *effective field theory*. This squeezes a propagator down to a point:

$$\frac{1}{p^2 - M^2 + im\Gamma} \approx -\frac{1}{M^2}, \quad (57)$$

in a fairly model independent way. Thus, for example a  $W$  boson coupling like

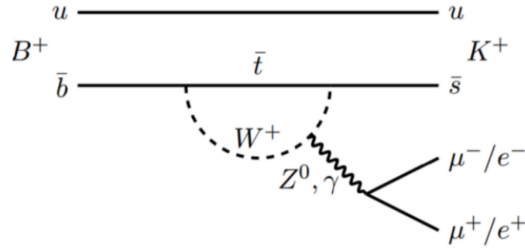
$$\mathcal{L} = -\frac{g}{2\sqrt{2}}\bar{e}\gamma^\rho(1 - \gamma_5)W_\rho\nu_e - \frac{g}{2\sqrt{2}}\bar{\nu}_\mu\gamma^\rho(1 - \gamma_5)W_\rho\mu \quad (58)$$

becomes

$$\mathcal{L} \approx -\frac{G_F}{\sqrt{2}}(\bar{e}\gamma^\rho(1 - \gamma_5)\nu_e)(\bar{\nu}_\mu\gamma^\rho(1 - \gamma_5)W_\rho\mu), \quad (59)$$

where  $G_F = \sqrt{2}g^2/(8M_W^2)$ . One has to be careful at the LHC with the range of validity of the effective field theory, however, because the LHC has a large centre of mass energy. If some of the collisions have  $p^2 \geq M^2$ , then for those collisions the effective field theory is a bad approximation: there, one becomes sensitive to the full structure of the propagator. Effective field theory methods can be useful for parameterising searches for new physics at low momentum: these four-fermion operators are often called *contact operators*, e.g. for some fermionic dark matter particle  $\chi$ ,

$$\mathcal{L} = \frac{\lambda^2}{M^2}(\bar{q}\gamma^\mu q)(\chi\gamma_\mu\chi) \quad (60)$$



**Fig. 22:** Example Feynman diagram contributing to  $B$  meson decays that form the variable  $R_K$ .

for some coupling strength  $\lambda$  [15]. However, for dark matter production at the LHC (e.g. in the monojet channel), the energies are often higher than the messenger mass and so a more precise (simplified?) model is needed [16]. Such a move to more specified models increases model dependence, but may be necessary if one requires a large régime of validity for one’s description of high energy collisions.

## 5 Conclusion

Unfortunately, at the time of writing, 13 TeV collisions at the LHC have yet to yield unambiguous and confirmed discoveries of new physics. In some channels, around  $13 \text{ fb}^{-1}$  of integrated luminosity has been analysed. However, there is plenty of room for new physics to be hiding: in more data or in other analyses. I personally and perhaps naively expect some signal to show up in the first  $100 \text{ fb}^{-1}$  of Run II data. Certainly it seems unlikely that if there are no significant excesses in that amount of data, there is unlikely to be<sup>12</sup> a discovery at Run II.

On the other hand, there are several interesting excesses in  $B$  physics measurements as compared to SM measurements, which we have not explicitly discussed in these Beyond the Standard Model lectures. Probably the theoretically cleanest of these is that of

$$R_K \equiv \frac{BR(B^\pm \rightarrow K^\pm \mu^+ \mu^-)}{BR(B^\pm \rightarrow K^\pm e^+ e^-)} = 0.745_{-0.074}^{+0.090} \pm 0.036 \quad (61)$$

from the LHCb experiment [13]. Large theoretical uncertainties associated with mesonic physics cancel well in such a ratio, particularly when one is probing final states involving leptons. In the SM,  $R_K = 1.00$  is a firm prediction from diagrams like Fig. 22, and so the measurement in Eq. 61 indicates non-SM lepton flavour non-universality at the  $3.6\sigma$  level. In fact, a fit to this and other data indicates that a new physics effective field theory operator on top of the SM

$$\mathcal{L} = C_9(\bar{s}_L \gamma^\mu b_L)(\bar{\mu} \gamma_\mu \mu) + \dots \quad (62)$$

is preferred to be non-zero at the  $4.3\sigma$  level [14].

We close with a quote from William Blake from *The Marriage of Heaven and Hell*:

“The road of excess leads to the palace of wisdom”.

## Acknowledgements

This work has been partially supported by STFC grant ST/L000385/1. We thank the TH Department at CERN for hospitality offered while these notes were written up. We also thank the organizers of the school, and the students whose enthusiasm in physics and Karaoke provided plentiful inspiration.

<sup>12</sup>One caveat to this statement is if a decision is made to increase the beam energy, for example from 13 TeV to 14 TeV. Increasing the beam energy tends to quickly increase the search sensitivity.

## References

- [1] F. Quevedo, S. Krippendorff and O. Schlotterer, arXiv:1011.1491 [hep-th].
- [2] K.A.Olive *et al* (Particle Data Group), Chin. Phys. **C38** (2014) 090001.
- [3] S. P. Martin, Adv. Ser. Direct. High Energy Phys. **21** (2010) 1 [Adv. Ser. Direct. High Energy Phys. **18** (1998) 1] doi:10.1142/9789812839657\_0001, 10.1142/9789814307505\_0001 [hep-ph/9709356].
- [4] B. C. Allanach, Comput. Phys. Commun. **143** (2002) 305 doi:10.1016/S0010-4655(01)00460-X [hep-ph/0104145].
- [5] B. C. Allanach, S. Grab and H. E. Haber, JHEP **1101** (2011) 138 Erratum: [JHEP **1107** (2011) 087] Erratum: [JHEP **1109** (2011) 027] doi:10.1007/JHEP07(2011)087, 10.1007/JHEP09(2011)027, 10.1007/JHEP01(2011)138 [arXiv:1010.4261 [hep-ph]].
- [6] C. Young (on behalf of ATLAS and CMS Collaborations), talk at Electroweak Session of Moriond 2016 conference, La Thuile, Italy.
- [7] D. Alves *et al.* [LHC New Physics Working Group Collaboration], J. Phys. G **39** (2012) 105005 doi:10.1088/0954-3899/39/10/105005 [arXiv:1105.2838 [hep-ph]].
- [8] G. Aad *et al.* [ATLAS Collaboration], JHEP **1510** (2015) 134 doi:10.1007/JHEP10(2015)134 [arXiv:1508.06608 [hep-ex]].
- [9] A. Barr and J. Liu, arXiv:1605.09502 [hep-ph].
- [10] S. Raychaydhuri and K. Sridhar, Cambridge Monographs on Mathematical Physics, Cambridge University Press (2016), ISBN 9780521768566.
- [11] G. F. Giudice, R. Rattazzi and J. D. Wells, Nucl. Phys. B **544** (1999) 3 doi:10.1016/S0550-3213(99)00044-9 [hep-ph/9811291].
- [12] B. C. Allanach, K. Odagiri, M. A. Parker and B. R. Webber, JHEP **0009** (2000) 019 doi:10.1088/1126-6708/2000/09/019 [hep-ph/0006114].
- [13] R. Aaij *et al.* [LHCb Collaboration], Phys. Rev. Lett. **113** (2014) 151601 doi:10.1103/PhysRevLett.113.151601 [arXiv:1406.6482 [hep-ex]].
- [14] W. Altmannshofer and D. M. Straub, Eur. Phys. J. C **75** (2015) no.8, 382 doi:10.1140/epjc/s10052-015-3602-7 [arXiv:1411.3161 [hep-ph]].
- [15] J. Goodman, M. Ibe, A. Rajaraman, W. Shepherd, T. M. P. Tait and H. B. Yu, Phys. Rev. D **82** (2010) 116010 doi:10.1103/PhysRevD.82.116010 [arXiv:1008.1783 [hep-ph]].
- [16] J. Abdallah *et al.*, Phys. Dark Univ. **9-10** (2015) 8 doi:10.1016/j.dark.2015.08.001 [arXiv:1506.03116 [hep-ph]].



## Physics at the LHC Run-2 and Beyond

A. Hoecker

CERN, Geneva, Switzerland

### Abstract

After an astounding Run-1 with 8 TeV proton–proton collisions featuring among others the discovery of the Higgs boson, Run-2 of the Large Hadron Collider (LHC) has started in 2015 colliding protons with unprecedented 13 TeV centre-of-mass energy. The higher energy and large expected integrated luminosity significantly increases the discovery potential for new physics, and allows for more detailed Higgs boson studies as well as improved Standard Model measurements. The lecture discusses methods, recent results and future prospects in proton–proton physics at the LHC.<sup>1</sup>

### Keywords

LHC; ATLAS; CMS; LHCb; Higgs; Standard Model; BSM physics.

## 1 Introduction

The Large Hadron Collider (LHC) at CERN probes nature at the smallest distances ever explored on Earth to study and improve our current knowledge of space and time, matter and force as it is encoded in the Standard Model (SM) of particle physics. The SM is *the* legacy of 20<sup>th</sup> century particle physics: it unifies quantum mechanics, special relativity and field theory; it unifies electromagnetic and weak interactions; it describes (about) all laboratory data. Does the SM deliver a complete answer to the complexity of the world generated by the simultaneous existence of very small as well as very large, seemingly fundamental numbers? We have reasons to believe that this is not the case.

The SM is made of spin one-half matter particles consisting of three generations of massive quarks and leptons, and force carriers in the form of partially massive spin one gauge bosons. An additional doublet of complex scalar fields, the Brout-Englert-Higgs (BEH) field  $\phi$ , is dictated by the requirement of local gauge symmetry [1]. Its condensation after spontaneous symmetry breaking at low temperature is responsible for the masses of the SM gauge bosons and (chiral Dirac-) fermions, leaving the electromagnetic force with infinite range, but making the weak force short-ranged (about  $10^{-15}$  cm). The new field is not only a constant background field, but it has its own massive quantum, the scalar Higgs boson. Being a boson, we might want to call it a fifth force. However, unlike the other forces, the new force is not a gauge force. Its non-universal coupling to masses of fermions and gauge bosons reminds us of classical gravitation, but the BEH force is much stronger than gravity and short-ranged.

The potential of the scalar double field  $\phi$  consists in its simplest form at low temperature of three terms: a quadratic term with negative coefficient  $\mu$ , a quartic term with positive coefficient  $\lambda$  realising the “Mexican hat” shape, and a Yukawa term describing the helicity-changing couplings between the BEH field and the fermions. The discovery of the Higgs boson and measurement of its mass fixes the coefficients to  $\lambda = m_H^2/2v^2 \approx 0.13$  and  $|\mu| = \sqrt{\lambda} \cdot v = m_H/\sqrt{2} \approx 89$  GeV, where  $v = |\mu|/\sqrt{\lambda} = (\sqrt{2} \cdot G_F)^{-1/2} \approx 246$  GeV is the vacuum expectation value of  $\phi$ .

---

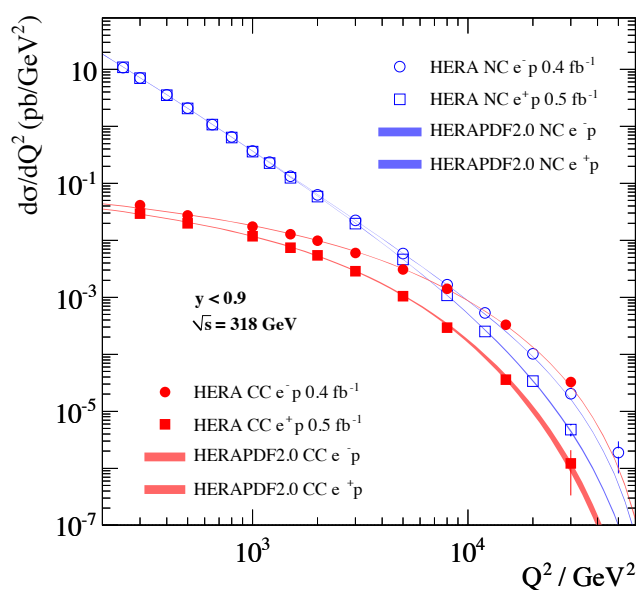
<sup>1</sup>As I am a member of the ATLAS experiment, for practical reasons, this lecture writeup leans somewhat towards ATLAS results. In the majority of the cases, the plots shown can be interchanged against those from CMS (and vice versa) without altering the message.

We may wonder how the potential evolves with the decreasing temperature of the expanding universe. Above the critical temperature  $T_{EW}$  of approximately 100 GeV, during the earliest  $10^{-11}$  second after the big bang where the universe covered a causal domain of a few cm, the evolution of the potential is such that its minimum is  $\langle 0|\phi|0\rangle_{T>T_{EW}} = 0$ . Gauge symmetry is respected, ie, all matter particles are massless and weak interaction is long-ranged. A spontaneous phase transition at  $T_{EW}$  (within the SM expected to be continuous, that is of second order) displaces the ground state of the BEH field to  $\langle 0|\phi|0\rangle_{T<T_{EW}} = v$ , breaking gauge symmetry. This spontaneous symmetry breaking corresponds to choosing a direction in the  $SU(2)_L \times U(1)_Y$  group space. The condensed field fills all space-time, but without orientation as it has no spin. (One could imagine it as a Lorentz-invariant ether [2]). The massive gauge bosons and fermions interact with the condensate which effectively reduces their velocity. The acquired mass is proportional to the strength of that interaction. The action of the BEH field thus creates a “vacuum viscosity”.

There are many questions about the structure of the SM in particular related to the matter sector: a large mass hierarchy is observed; CP violation has been observed in the quark sector, consistent with a single CP-violating phase in the quark mixing matrix. In effect, three quark generations allow for exactly one such phase. The neutrino sector still bears many unknowns, among which the origin and values of neutrino masses, the neutrino nature, CP violation and whether or not there are sterile neutrinos that are singlets under the SM interactions (but possibly not under the new force).

Electroweak unification reduces the number of SM parameters from 20 to 19 (including the strong-CP parameter and neglecting the massive neutrino sector, which is irrelevant for LHC physics unless there are new right-handed neutrinos in reach of the LHC). Figure 1 demonstrates beautifully electroweak unification at work at the HERA collider in electron–proton and positron–proton scattering [3]. At low momentum transfer, neutral current processes with photon exchange producing an electron/positron in the final state dominate over charged current processes mediated via  $W$  bosons. Above 100 GeV, however, neutral and charged current processes are of similar size: electromagnetic and weak interactions are unified.<sup>2</sup> Electroweak unification relates the electromagnetic and weak coupling strengths to each other (the latter coupling given at lowest order by the ratio-squared of weak gauge boson masses). This relation has been tested experimentally to high precision [4].

The scales of particle physics reach from zero mass (for the photon and gluons) to as much as (and beyond) the Planck scale set by the strength of gravity. Within these two extremes lies the range of sensitivity of the LHC covering three orders of magnitude between roughly 10 GeV and 10 TeV, probing



**Fig. 1:** Differential cross-section versus momentum transfer-squared measured at the HERA collider for neutral (blue) and charged (red) current deep inelastic scattering processes [3]. The data points are integrated over the Bjorken- $x$  variable.

<sup>2</sup>Looking in more detail into Fig. 1, the neutral current cross sections for  $e^-p$  and  $e^+p$  are almost identical at small  $Q^2$  but start to diverge as  $Q^2$  grows. This is due to  $\gamma$ - $Z$  interference, which has the opposite effect on the  $e^-p$  and  $e^+p$  cross sections. The charged current cross sections also differ between  $e^-p$  and  $e^+p$  scattering, with two effects contributing: the helicity structure of the  $W^\pm$  exchange and the fact that charged current  $e^-p$  ( $e^+p$ ) scattering probes the  $u$ -valence ( $d$ -valence) quarks [3].

length scales down to an attometre. That range comprises the study of high-energetic radiation such as jets, charm and bottom flavour physics, top quarks,  $Z$ ,  $W$  bosons and the Higgs boson, and any new physics that may reside therein. The physics at scales above that of the LHC is highly speculative. There could be right-handed neutrinos of mass above  $10^{10}$  GeV, as predicted by the (type 1) seesaw mechanism, the Peccei-Quinn axion scale above  $10^{10}$  GeV to suppress strong CP violation, grand unification of the electroweak and strong forces at roughly  $10^{15}$  GeV, quantum gravity at roughly  $10^{18}$  GeV and the hypercharge Landau pole well above the Planck scale.

## 2 Particle physics at the dawn of the LHC

The Higgs boson — last of the particles? The SM predicts all properties, except for its mass. But before coming to the Higgs boson let us briefly recall the status of particle physics at the dawn of the LHC.

- LEP and SLC had ended their experimental programmes, with among their main results the proof of three light active neutrino flavours, and direct Higgs boson searches that excluded  $m_H < 114$  GeV. Moreover, SM tests to unprecedented precision were performed with no direct or indirect hint for beyond the SM (BSM) physics. Among these, asymptotic freedom of strong interactions was tested to the percent level through measurements of the strong coupling strength  $\alpha_s(\mu)$  at scales  $\mu = m_\tau$  and  $\mu = m_Z$ , respectively, and comparison with the accurately predicted evolution from the QCD renormalisation group. In both cases the extraction occurred by comparing experimental results for the inclusive  $\tau$  or  $Z$  hadronic widths (among other  $Z$  pole observables) with NNNLO (3NLO) perturbative QCD predictions.
- Precision measurements at the  $Z$  pole and of the top-quark and  $W$ -boson masses allowed to exclude an SM Higgs boson heavier than about 160 GeV at 95% confidence level. There are also theoretical arguments in favour of a not too heavy Higgs boson, which is required to moderate longitudinal weak-boson scattering at large momentum transfer. The evolution of the quartic coupling in the scalar potential with the energy scale  $\Lambda$  representing the SM cut-off scale where new physics occurs leads to constraints on  $m_H$  in terms of upper perturbativity and lower (meta)stability bounds. Indeed, the SM Higgs boson must steer a narrow course between two disastrous situations if the SM is to survive up to the Planck scale  $\Lambda = M_{\text{Planck}}$ .
- The Tevatron collider at Fermilab, USA still continued Run-2. That collider led to the discovery of the top quark and the measurement of its mass to better than 1%. The  $W$  boson mass was measured more precisely than at LEP and by today Tevatron dominates the world average. The mixing frequency of neutral  $B_s$  mesons was measured for the first time, and found in agreement with the SM prediction. The Higgs boson was beyond Tevatron’s sensitivity except for masses around 160 GeV, which could be excluded. No hint for BSM physics was seen.
- The  $B$  factory experiments BABAR at SLAC, USA and Belle at KEK, Japan were about to end with a precise confirmation of the Kobayashi-Maskawa paradigm of a phase in the three-generation CKM quark matrix being the sole responsible of the observed CP violation in the quark sector. Ambiguous initial hints about a possible difference in the unitarity triangle angle  $\beta$  extracted from tree and loop (“penguin”) processes disappeared with increasing statistics. The  $B$ -factory experiments measured many rare processes and observed for the first time CP violation in the charm sector.
- There was (and still is) no hint for charged-lepton flavour violation in spite of ever increasing experimental sensitivity. Any non-zero measurement would indicate new physics as the SM predictions via the massive neutrino sector are immeasurably small. Also, no sign of a CP-violating electric dipole moment (EDM) was seen in atoms or neutrons. The absence of a neutron EDM strongly constrains QCD induced CP violation that would be expected in the SM. Only the anomalous

magnetic moment of the muon exhibits a long-standing  $>3\sigma$  discrepancy between data and the SM prediction.

- The neutrino sector has seen a revolution after the discovery of neutrino oscillation and the measurement of all three angles of the neutrino mixing matrix. These measurements establish that neutrinos have mass, but their nature (Dirac versus Majorana), mass hierarchy (normal versus inverted), as well as CP violating mixing phase remain unknown and are the subject of intense experimental activity.
- Finally, there has been no signal other than gravitational effects for dark matter, no signs of axions or of proton decay.<sup>3</sup>

### 3 Experimental setup

Producing the Higgs boson and searching for new physics at the TeV scale requires a huge machine. Particle accelerators exploit three principles: (i) they look deep into matter requiring high energy to resolve small de Broglie wave lengths (particle accelerators are powerful microscopes), (ii) Einstein’s relation between energy and mass allows to produce potentially new heavy particles at high energy, and (iii) accelerators probe the conditions of the early universe through Boltzmann’s relation between energy and temperature.

Figure 2 gives a schematic view of CERN’s accelerator complex. It consists of a succession of machines that accelerate particles to increasingly higher energies and condition the particle beams. Each machine boosts the energy of a beam before injecting it into the next machine in the sequence. Protons are accelerated to 50 MeV in the Linac 2, to 1.4 GeV in the PS Booster, to 26 GeV in the PS (Proton Synchrotron) where also spatial proton bunches with 25 ns (7.5 m) distance and bunch trains are formed, 450 GeV in the SPS (Super Proton Synchrotron) before being injected in opposite directions into the LHC. Booster, PS and SPS have their own experimental halls where fractions of the beams are used for fixed target experiments at lower energies. The LHC [5] is a superconducting proton/ion accelerator and collider installed in a 26.7 km circumference, 70–140 m underground tunnel with 4 m cross-section diameter. Up to 2 800 bunches containing each more than 100 billion protons are accelerated within roughly 20 minutes from 450 GeV up to the design energy of 7 TeV per beam. So far, proton–proton collisions with centre-of-mass energies of 0.9, 2.8, 5, 7, 8, and 13 TeV were delivered by the LHC. In addition, data with 5 and 8 TeV proton–lead and 2.8 and 5 TeV lead–lead collisions were taken.

The most challenging component of the LHC are the 1 232 superconducting dipole magnets realised in a novel “2-in-1” design that guide the protons along their circular trajectory around the ring. The dipoles have a length of 14.3 m each, and are cooled to 1.9° K by means of a closed circuit of 120 tonnes liquid-helium. The LHC also features almost 400 focusing quadrupole magnets and 3 700 multipole corrector magnets. The maximum dipole field strength of 8.3 T, achieved with a current of almost 12 kA, limits the energy to which the protons can be accelerated. The proton’s energy is given by  $E_p[\text{TeV}] = \sqrt{4\pi\alpha \cdot B[\text{T}] \cdot r[\text{km}]}$  so that, with the radius  $r = 4.3$  km, one finds  $E_p \sim 7$  TeV taking into account that only roughly two-third of the ring are equipped with dipoles. Following the scale-energy relation  $\mu \approx 200 \text{ GeV am} / E[\text{TeV}]$ , does the LHC thus probe length scales of  $\mu \sim 10^{-20}$  m at 14 TeV centre-of-mass energy? It is not quite that small as the protons are composite particles whose energy is distributed among

<sup>3</sup>It is not possible to reach energies in the laboratory that would allow to directly study the physics at the expected grand unification scale. Even Enrico Fermi’s “Globatron” (that was to be built in 1994) would with current LHC magnet technology “only” reach insufficient 20 PeV proton–proton centre-of-mass energy. Proton decay is among the greatest mysteries in elementary particle physics. It is required for baryogenesis and predicted by grand unified theories (GUT). Its discovery could therefore provide a probe of GUT-scale physics. The best current limit on the partial proton lifetime from Super-Kamiokande, combining all its data from 1996 until now, is  $\tau(p \rightarrow e^+ \pi^0) > 1.7 \cdot 10^{34}$  years.

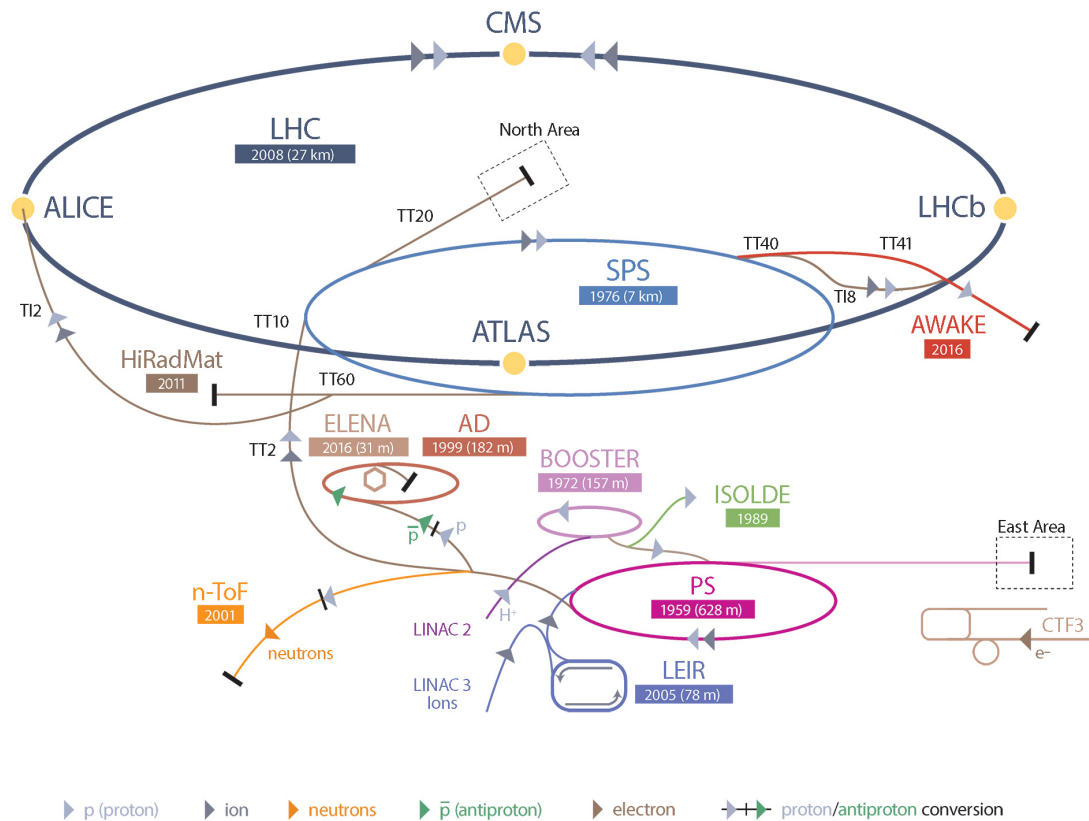


Fig. 2: Schematic view of CERN's accelerator complex.

its constituents (partons).

The LHC hosts four large, ultra-sophisticated experiments among which the general purpose detectors ATLAS and CMS, as well as ALICE and LHCb dedicated mainly (but not only) to heavy-ion and flavour physics studies, respectively, requiring optimisation for low transverse momentum physics. There are additional smaller-scale experiments dedicated to forward physics. The design of the ATLAS detector emphasises excellent jet and missing transverse momentum resolution, particle identification, flavour tagging, and standalone muon measurement. CMS features excellent electron/photon energy and track (muon) momentum resolution, and flavour tagging. Both detectors are highly hermetic with very few acceptance holes allowing to precisely determine missing transverse momentum. ALICE has highly efficient track reconstruction in busy heavy-ion environment and particle identification. LHCb is a forward spectrometer with a trigger for fully hadronic  $B$  and  $D$  hadron events, excellent low-momentum track resolution, and particle identification (pion/kaon separation).

The particle detectors measure particles produced as debris from the proton–proton collisions through interaction with active material. Different concentric detector layers measure different properties. The innermost parts of the detectors measure tracks of charged particles in layers of semi-conductors or straw tubes that respond to traversing charges. The momentum and charge of these particles is measured through their immersion in a homogeneous magnetic field. Outside the tracking volume are thick calorimeters that absorb most particles and measure their energy. Additional tracking chambers behind the calorimeters identify and measure muons, which are minimum ionising in most of their momentum range and thus penetrate the calorimeter layers. Neutrinos do not interact with the detector and therefore induce missing energy. Due to the non-zero longitudinal momentum of the collisions in the laboratory frame and the missing acceptance coverage along the beam line, only the transverse missing momentum

in a collision is conserved. Hence, the reconstructed missing transverse momentum is used to detect neutrinos or any unknown non or weakly interacting particles. In addition to stable particles, the experiments also reconstruct jets from near-by calorimeter depositions or tracks. Jets are narrow cones of hadrons and other particles produced by the hadronisation of a quark or gluon due to QCD confinement. Reconstruction of long-lived states in a jet allows to tag jets originating from  $b$  or  $c$  quarks. The ensemble of measurements in a given proton–proton bunch crossing makes up an event. It contains more than the original hard parton scattering due to underlying event interaction and additional soft proton–proton interactions (dubbed “pileup”).

LHC computing represents “big data”: the LHC experiments started more than a decade ago with large scale computing, which is now present everywhere. The ATLAS managed data volume of roughly 150 petabyte (dominated by simulated data) is of similar order as the Google search index or the content uploaded to Facebook every year. Unlike these companies, however, the LHC has to manage its data volume with a public science budget.

## 4 Experimental methods

We will review in this chapter a few (basic) experimental concepts at the LHC.

### 4.1 Luminosity

Besides energy, luminosity is the single most important quantity in collider physics. The instantaneous luminosity of the beam collision, expressed in units of  $\text{s}^{-1}\text{cm}^{-2}$ , is a function of the LHC beam parameters as follows

$$L = \frac{f_{\text{rev}} \cdot n_{\text{b}} \cdot N_p^2}{4\pi \cdot \sigma_x \cdot \sigma_y} \cdot F(\theta_c, \sigma_x, \sigma_z), \quad (1)$$

where  $f_{\text{rev}} = 11\,245.5$  Hz is the bunch revolution frequency determined by the size of the LHC (27 km) and the speed of light,  $n_{\text{b}} = 1, \dots, 2\,808$  is the number of proton bunches in the machine (2 808 is the maximum number of possible 25 ns slots; the theoretical maximum of 3 564 bunches cannot be reached due to space needed between bunch trains and for the beam dump kicker magnets (abort gap)),  $N_p \approx 1.15 \cdot 10^{11}$  is the number of protons per bunch (the bunch intensity), and  $\sigma_{x,y} = 12, \dots, 50 \mu\text{m}$  is the transverse beam width characterising the beam optics. The factor  $F(\theta_c, \sigma_x, \sigma_z)$  accounts for luminosity reduction due to the beam crossing angle  $\theta_c$ , roughly given by  $(1 + (\sigma_z/\sigma_x)^2 \cdot (\theta_c/2)^2)^{-1/2}$ , the hourglass effect leading to a varying transverse bunch size in the collision point because of the several cm longitudinal bunch extension, and other effects.

Luminosity drives the statistical precision of any measurement and our ability to observe low cross section processes as

$$N_{\text{events}}^{\text{obs}} = \text{cross section} \times \text{efficiency} \times \int L \cdot dt, \quad (2)$$

where the cross section is given by Nature, the efficiency of detection is optimised by the experimentalist, and the integrated luminosity is delivered by the LHC. There are several options to maximise the luminosity of the machine as outlined below.

- **Maximise the total beam current.** The cryogenic system limits the maximum beam current leading to an anticorrelation between  $N_p$  and  $n_{\text{b}}$ . Improvements in beam collimation, cryogenics vacuum, and background protection allow to extend that limit.

- **Maximise brightness and energy, minimise  $\beta^*$ .** The transverse beam size is given by  $\sigma(s) = \sqrt{\beta(s)\varepsilon_n/\gamma}$ , where  $\sigma^* = \sigma(s=0) \approx 17 \mu\text{m}$  at the collision point. The value  $\beta^* \approx 60 \text{ cm}$  is the longitudinal distance from the focus point where the transverse beam size grows twice as wide. The emittance  $\varepsilon \cdot \pi$  is the area in phase space occupied by the beam, and  $\varepsilon_n \approx 3.8 \mu\text{m}$  is the normalised emittance, where the Lorentz  $\gamma$  factor is taken out. To reduce (“squeeze”)  $\beta^*$  one needs to respect the quadrupole aperture limit. The beam brightness,  $N_p/\varepsilon_n$ , is limited by beam–beam interactions which have a quadrupole de-focusing effect, and by space-charge tune shift and spread (the tune spread is limited by resonances).
- **Compensate reduction factor.** The crossing angle is required to avoid parasitic long-range beam encounters. The hour glass effect may be reduced by shorter bunches, at the expense of a higher longitudinal pileup density.

The LHC group offers an excellent tool [6] to study the dependence of the expected LHC luminosity under various parameter settings.

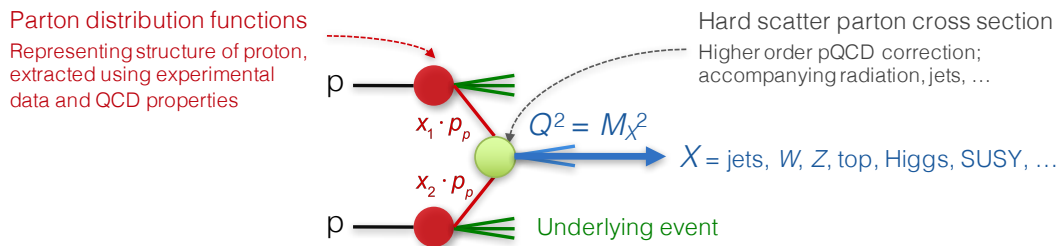
The instantaneous luminosity is measured by the experiments [7–9] with dedicated detector systems that are calibrated with the use of so-called van-der-Meer beam-separation scans [10]. Such scans are performed in specific low-intensity LHC fills during which the beams are separated from each other by an increasing distance in both  $x$  and  $y$  directions. The dedicated luminosity detectors count for each scan point the hits they receive from inelastic minimum-bias events. From the Gaussian profile of the hit counts versus the beam separation one can determine the transverse beam profiles entering Eq. (1) via the convolution relation  $\sigma_{x,y} = \Sigma_{x,y}/\sqrt{2}$ , where it is assumed that both LHC beams have the same width. The knowledge of  $L$  from the measured beam currents and beam widths in the specific LHC fill allows to extract the visible cross section,  $\sigma_{\text{vis}}$ , for any dedicated luminosity detector (this may include, for example, cluster counting in the Pixel detector). During normal (ie, high-luminosity) data taking, the counts measured in that detector together with the calibrated visible cross section from the van-der-Meer scan allow to extract the luminosity via  $L = N_{\text{counts}}/\sigma_{\text{vis}}$ . Note that this method does not require to know the acceptance nor the efficiency of the luminosity detector, which are included in  $\sigma_{\text{vis}}$ .

The precision of the luminosity calibration depends on many factors. Systematic uncertainties arise from correlations between the  $x$  and  $y$  transverse beam positions during the scan, beam–beam corrections, beam orbit drifts and position jitters, stability (reproducibility) of the results, and instrumental effects such as the absolute length scale calibration for the separated beams, beam backgrounds and noise, the reference specific luminosity, the measurement of the beam currents, and the extrapolation of the calibration from the low to high-luminosity regimes as well as the run-by-run stability of the luminosity detectors. The best precision on the integrated luminosity achieved by the experiments undercuts 2%.

During Run-1 of the LHC, spanning the years 2010–2012 of data taking, the peak luminosity achieved was  $L_{\text{peak}} = 7.7 \cdot 10^{33} \text{ cm}^{-2}\text{s}^{-1}$  and an integrated luminosity during the year 2012 of  $23 \text{ fb}^{-1}$  at 8 TeV proton–proton centre-of-mass energy was delivered to the experiments. With 50 ns bunch distance a maximum number of 1380 colliding bunches was reached. At  $L_{\text{peak}}$  the LHC produced every 45 minutes a  $H \rightarrow \gamma\gamma$  event, and typically two  $160 \text{ pb}^{-1}$  fills were needed to produce one  $H \rightarrow 4\ell$  ( $\ell = e, \mu$ ) event.

The high luminosity of the LHC comes to the price of additional inelastic proton–proton pileup interactions within a bunch crossing. An average of  $\langle \mu \rangle = 21$  (maximum  $\langle \mu \rangle = 40$ ) interactions per crossing occurred during 2012, with a similar or slightly higher rate in 2016. The LHC design pileup value at 14 TeV is obtained as follows

$$\langle \mu \rangle = \frac{\sigma_{\text{inel}} \cdot L}{f_{\text{rev}} \cdot n_b} \approx \frac{80 \text{ mb} \cdot 10 \text{ nb}^{-1}\text{s}^{-1}}{11\,245 \text{ s}^{-1} \cdot 2\,808} \approx 25, \quad (3)$$

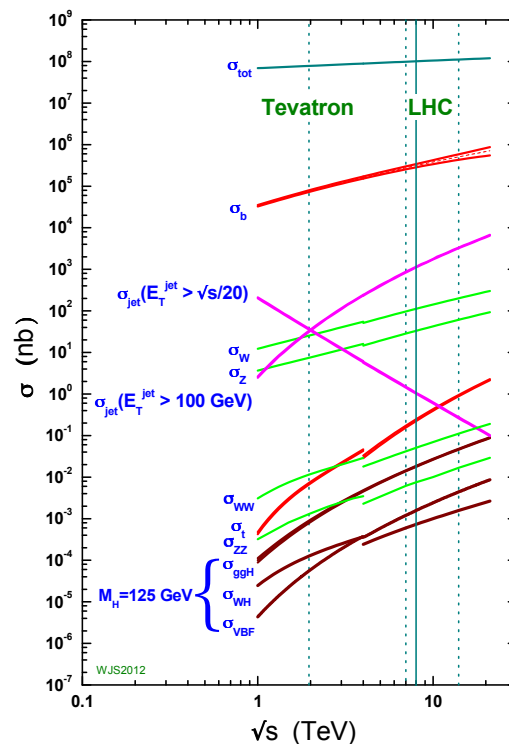


**Fig. 3:** Simplified view of a proton–proton collision.

where we used  $10^{34} \text{ cm}^{-2}\text{s}^{-1} = 10 \text{ nb}^{-1}\text{s}^{-1}$ . When the detector response integrates over several bunch crossings, as for example the calorimeter pulse shape, pileup occurring in the recorded proton collision (in-time pileup) as well as that in neighbouring collisions (out-of-time pileup) affect the event reconstruction. Most analyses are fairly insensitive to pileup at the rates experienced so far. Mitigation methods have been developed to further improve the robustness of the physics object reconstruction and analyses. Pileup does, however, affect the trigger requiring higher thresholds, which impacts the low transverse momentum physics programme of the experiments. It also increases the stored event size and CPU time needed for track reconstruction.

## 4.2 Proton–proton collisions

Owing to factorisation (see for example [12]), the cross section of a proton–proton collision can be computed as the convolution of parton density functions<sup>4</sup> (PDF) with the parton scattering matrix element (cf. Fig. 3). The PDFs are universal distributions containing the long-distance structure of the proton (or hadrons in general) in terms of valence and sea quarks and gluons. They are related to parton model distributions at leading order, but with logarithmic scaling violations (DGLAP<sup>5</sup>). Since precise Lattice QCD predictions are not yet available, the PDFs are extracted versus the parton momentum fraction  $x$  and the momentum transfer  $Q^2$  using experimental data and exploiting QCD evolution properties. The centre-of-mass energy-squared of the parton collision,  $\hat{s}$ , is given by the product of the momentum fractions of the colliding partons,  $x_{1,2}$ , times the proton centre-of-mass energy:  $\hat{s} = x_1 \cdot x_2 \cdot s$ . The production of the 125 GeV Higgs boson thus occurs at an average momentum fraction  $\langle x \rangle \sim 0.01$  at  $\sqrt{s} = 13 \text{ TeV}$ .



**Fig. 4:** Cross sections of selected proton–(anti-)proton processes versus centre-of-mass energy [11].

<sup>4</sup>Parton density functions were introduced 1969 by Feynman in the parton model to explain Bjorken scaling in deep inelastic scattering data.

<sup>5</sup>In analogy with a running coupling strength, one can vary the factorisation scale and obtain the renormalisation group equation for PDFs. The DGLAP equations [13] describe the  $Q^2$  dependence of the PDFs.



The parton density functions rise dramatically towards low  $x$  in particular at high  $Q^2$  and most notably for the gluon density. The consequences are: the cross section of a given process increases with increasing proton–proton collision energy, more luminosity allows to reach higher parton collision energy, and the low- $x$  regime is dominated by gluon–gluon collisions. Hence, although the parton-level cross sections falls with centre-of-mass energy, the cross-sections of proton–proton processes rise due to the convolution with increasing PDFs. This is depicted in Fig. 4: all cross sections rise with centre-of-mass energy, where gluon initiated processes have a steeper slope than quark initiated ones because of the strong enhancement of the gluon PDF towards lower  $x$ . One notes that when requiring a centre-of-mass-energy dependent minimum transverse momentum for jets, the jet cross section decreases with centre-of-mass energy as expected for a parton–parton cross section. The inclusive cross section is dominated by inelastic scattering (also denoted minimum bias events), the interesting jet and boson physics processes have many orders of magnitude lower rates.

### 4.3 The experimental data path in a nutshell

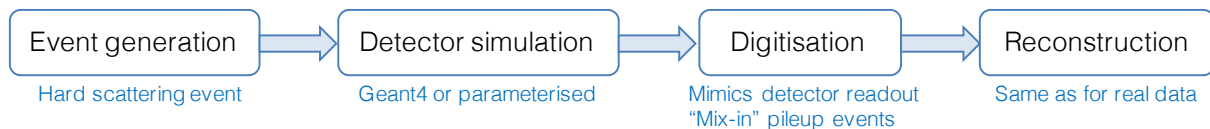
The LHC detectors cannot record events at the filled proton bunch crossings rate of approximately 30 MHz<sup>6</sup> as each event has an approximate raw size of 2 MB, requiring to store 60 TB per second. This is not only impractical given the computing resources, but also unnecessary in view of the LHC physics goals as most events contain only soft minimum bias interactions. Instead, online custom hardware and software triggers reduce that rate by filtering out events with a million and more times smaller cross sections than minimum bias events.

The data path of an LHC experiment can be described in a nutshell as follows.

1. LHC bunches collide every 25 n (50 ns during Run-1), but not all bunches are filled with protons.
2. LHC detectors record the detector response in pipelined on-detector memories that are time-stamped (synchronised) to the LHC collisions they belong to. The events are kept during the latency of the first level trigger decision (2–3  $\mu$ s).
3. Level-1 hardware and high-level software triggers filter the interesting events that are written to disk. The level-1 trigger system reduces the initial bunch crossing rate to up to 100 kHz. The high-level trigger further reduces this rate to about 1 kHz that are kept. A trigger menu is a large collection of physics and monitoring triggers. Among these are low-threshold single lepton triggers, single missing transverse momentum and jet triggers, and lower threshold di-object and topological triggers. The online system also provides detailed data quality monitoring.
4. The recorded data are subject to prompt offline calibration and refined monitoring, followed by the prompt reconstruction (mostly) at CERN.
5. The reconstructed data are distributed to computing centres world-wide from where standardised derived datasets are produced for physics and performance analysis.
6. Large amounts of Monte Carlo events using the same reconstruction software as used for data are also produced and distributed for analysis.
7. Performance groups provide standard physics objects with calibrations and uncertainties, unified in analysis releases. Analysis groups build physics analyses on top of this ground work.

---

<sup>6</sup>The LHCb phase-1 upgrade is preparing for exactly that!



**Fig. 5:** Monte Carlo event simulation chain.

#### 4.4 Monte Carlo event simulation

A crucial ingredient to any physics and performance analysis is Monte Carlo (MC) event simulation. MC events mimic the physics processes, which allows to isolate specific processes by subtracting simulated background processes, to evaluate the acceptance and efficiency of signal processes, to optimise signal selection, and to evaluate systematic uncertainties by varying MC parameters.

The MC generation path is sketched in Fig. 5. A matrix element generator calculates the hard parton-parton scattering event and stores it in a common data format. The event is passed through the detector simulation which simulates the interactions of the stable particles with the active and passive detector material. The simulation may use Geant4 [14] or a parametrised fast simulation. The output of that process is subject to the digitisation step during which the detector response and readout is mimicked. After this step the simulated data have the same format as real detector data except for the so-called truth information which records the information about the generated particle types, decay chains and four-momenta. The following event reconstruction is identical to that of real data (also format wise).

The physics modelling with event generators proceeds as follows. The hard scattering matrix element calculation including initial and final-state radiation (ISR/FSR) is convolved with the parton density functions. Decays of the hard subprocesses, and multiple parton interactions (and their ISR/FSR) are also generated. Matrix elements are used as much as possible, but one cannot fully avoid phenomenological description of nonperturbative effects such as parton showers, hadronisation, and the underlying event. State-of-the-art event generation includes next-to-leading order (NLO) matrix elements up to two partons, leading order (LO) matrix elements up to 5 partons, parton shower matching. Nonperturbative and electroweak corrections are sometimes applied. Fixed-order calculations are known to higher order (NNLO or even 3NLO). The physics modelling, including PDFs, has the largest systematic uncertainty in many analyses and is therefore a critical step that requires care, extensive validation and sometimes correction with data.

There has been significant progress in both generator developments and fixed-order calculations. Fixed-order predictions have seen an NNLO revolution with about 20 new results during the last two years. Some of the resulting NNLO to NLO K-factors, in particular for diboson production, were not covered by the uncertainties assigned to the NLO calculations as obtained from the canonical factors two and one-half variations of the renormalisation scale [15].

#### 4.5 Cross section measurement and data unfolding

The measurement of the cross section of a given process requires to isolate this process via selection of its final state, followed by the subtraction of contributing background processes. The remaining number of events needs to be corrected for resolution effects, the selection efficiency and, in case of the total cross section, the acceptance:

$$\sigma_{pp \rightarrow X}^{\text{tot}} = \frac{1}{A_X} \sigma_{pp \rightarrow X}^{\text{fid}} = \frac{1}{A_X} \left( \frac{N_{\text{obs}} - N_{\text{bkg}}}{L \cdot C_X} \right). \quad (4)$$

Here,  $N_{\text{obs}}$  and  $N_{\text{bkg}}$  are the total observed and the estimated number of background events,  $L$  the integrated luminosity,  $A_X = N_{\text{gen, fid}}/N_{\text{gen}}$  is the acceptance factor, given by the fraction of generated  $pp \rightarrow X$  events falling into the fiducial acceptance defined close to the final state selection,  $C_X = N_{\text{reco, sel}}/N_{\text{gen, fid}}$  is a correction factor that corrects for the detector resolution and inefficiency of events generated in the fiducial region, and  $\sigma_{pp \rightarrow X}^{\text{fid}}$  is the fiducial cross section.

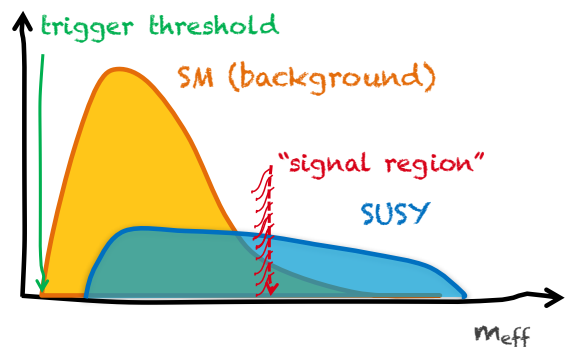
The acceptance factor is computed entirely from theory using the best available fixed order calculation. The correction factor depends on the particle detector and is partly or fully determined from MC simulation. It is therefore the main interest of the experimentalist to determine the fiducial cross section, which is corrected for experimental effects and has minimal theory dependence. The acceptance correction to obtain the total inclusive cross section is left for theory.

The definition of the fiducial cross section should facilitate the comparison between theoretical predictions and experimental results and thus should have the least possible dependence on the MC event generators available at the time of the measurement. A suitable definition of the observables is based on the physical particles that enter the detector. This includes the stable particles which account for the majority of interactions with the detector material, and from which the measurements are ultimately made. A detailed discussion about the definition of particle level objects is provided in a dedicated ATLAS note [16].<sup>7</sup>

A differential cross section corresponds to a binned fiducial cross section. It requires the application of unfolding due to bin-to-bin correlations. Unfolding is a mathematically unstable inversion problem that requires careful regularisation.

#### 4.6 Background determination

The subtraction of backgrounds as in Eq. (4) usually relies on MC simulation. In case of new physics searches, which often select extreme phase space regions such as a very large effective mass<sup>8</sup> as illustrated in Fig. 6, the MC predictions may suffer from large and difficult to estimate modelling uncertainties. More robust and reliable background estimates normalise the MC predictions of the main backgrounds in phase space regions (dubbed “control regions”) close to the signal region. Sometimes further data driven corrections for the transfer of the MC normalisation from the control region to a signal region are required.



**Fig. 6:** Sketch illustrating signal and background distributions of a discriminating variable. A requirement is applied to enrich the signal.

For a typical search for supersymmetry looking for jets, possibly  $b$ -jets and leptons, and missing transverse momentum, the main background sources stem from top production,  $W$  + jets and  $Z(\rightarrow \nu\nu)$  + jets

<sup>7</sup>The recommendations in Ref. [16] about how to suitably define an event topology at particle level on MC are: 1. Select the stable particles. 2. Select prompt leptons ( $e$ ,  $\nu_e$ ,  $\mu$ ,  $\nu_\mu$ ) and associate photons (not from hadrons) to electrons and muons to define dressed-level charged leptons. 3. Define particle-level jets by clustering all stable particles excluding the particles found in step (2). 4. Assign the jet flavour based on heavy-flavour hadrons ghost-matched to jets. 5. Sum all prompt neutrinos defined in (2) to form the missing transverse momentum. 6. Resolve lepton-lepton and jet-lepton overlap following a procedure close to that used at the detector level. 7. Define other particle-level observables in complex event topologies based on the particle-level objects defined above.

<sup>8</sup>The effective mass,  $m_{\text{eff}}$ , is defined by the scalar sum of the transverse momenta of all selected objects and the missing transverse momentum.

events,  $WW$ ,  $WZ$ ,  $ZZ$  diboson production, and rare processes such as  $tt + W$  or  $tt + Z$ . All these backgrounds may produce true missing transverse momentum due to decays to neutrinos. Additional backgrounds may arise from misreconstruction of multijet events, or the misidentification of non-prompt leptons or jets as prompt leptons. Such backgrounds are usually determined from data using control regions or so-called ABCD sideband methods exploiting two or more none or only weakly correlated variables. The former backgrounds are often called irreducible, and the latter due to misreconstruction are denoted reducible backgrounds. MC simulation is mainly used to predict irreducible backgrounds, in particular if they are sub-dominant.

#### 4.7 Basic physics objects

All ATLAS and CMS physics analyses are built upon basic physics objects. These correspond to single stable particles, ensembles of particles or event properties.

- **Tracks and vertices** are measured in the inner tracking systems. Their precise measurement requires an accurate detector alignment which is obtained from data by minimising hit residuals with respect to fitted tracks. Also important is a precise mapping of the inner tracker geometry and material, which is made with the help of survey data, and from collision data using reconstructed vertices from hadronic interactions, photon conversions to electron–positron pairs, track extensions, and long-lived hadrons.
- **Electrons and photons** are reconstructed as energetic clusters in the electromagnetic calorimeter associated or not with an inner detector track. Due to significant amount of active and passive tracker material (between 0.4 and 2.4 radiation lengths depending on  $|\eta|$ ),<sup>9</sup> roughly 40% of the photons convert to electron–positron pairs and hence are reconstructed as one or two displaced electron tracks. The electron efficiency, energy scale and resolution are precisely calibrated in data using  $Z \rightarrow ee$ ,  $J/\psi \rightarrow ee$  and  $W \rightarrow e\nu$  events. Photons are calibrated using MC and radiative  $Z \rightarrow ee(\mu\mu) + \gamma$  events, as well as  $\pi^0 \rightarrow \gamma\gamma$  decays at low energy.
- **Muons** are reconstructed in the inner tracker and the outer muon systems. Combined tracking improves the momentum resolution for high transverse momentum ( $p_T$ ) muons, which are dominated in precision by alignment uncertainties. Muons are calibrated using  $Z \rightarrow \mu\mu$ ,  $J/\psi \rightarrow \mu\mu$  and  $\Upsilon(1S) \rightarrow \mu\mu$  events.
- **Hadronic  $\tau$  decays** ( $\tau_h$ ) to a narrow jet of charged and neutral pions or kaons are reconstructed in the inner tracker and the electromagnetic and hadronic calorimeters. Multivariate analysers (and particle flow in CMS, see next item) are used to combine the available detector information and improve the efficiency and purity of the selection as well as the energy measurement (using multivariate regression). Taus are calibrated using  $Z \rightarrow \tau_h\tau_h$  decays and  $E/p$  for the hadronic tracks.
- **Jets** are formed by clustering particles using the infrared and collinear safe<sup>10</sup> anti- $k_r$  algorithm [17] (for which the distance between clustered particles is defined using negative  $p_T$  power) via a pairwise successive aggregation of proto-jets. Jet particles are reconstructed in the electromagnetic

<sup>9</sup>The “rapidity” of a particle is defined by  $y = \frac{1}{2} \ln[(E + p_z)/(E - p_z)]$ , where  $E$  denotes the particle’s energy and  $p_z$  the particle’s momentum along the beam direction ( $z$ ). Differences in rapidity are Lorentz invariant under a boost along  $z$ . The “pseudorapidity” is defined by  $\eta = \frac{1}{2} \ln[(p + p_z)/(p - p_z)] = -\ln(\tan \frac{\theta}{2})$ . The azimuthal angle  $\phi$  is measured in the plane transverse to the beam direction and the polar angle  $\theta$  is measured with respect to the beam direction. Rapidity and pseudorapidity are equal for massless particles.

<sup>10</sup>Infrared safety requires that a jet remains unaffected when adding a particle with  $|p_T| \rightarrow 0$  to it. Collinear safety requires that a jet remains unaffected when replacing a particle  $i$  with four-momentum  $p_i$  by two particles  $j$  and  $k$  with four-momenta  $p_j + p_k = p_i$  such that  $|\vec{p}_i - \vec{p}_j| = 0$ , where  $\vec{p} = (y, \phi)$ .

and hadronic calorimeters in ATLAS, and with the use of a particle flow algorithm in CMS. The particle flow algorithm aims at identifying and reconstructing all the particles from the collision by combining the information from the tracking and calorimeter devices. The algorithm results in a list of particles, namely charged hadrons, neutral hadrons, electrons, photons and muons, which are used to reconstruct jets and missing transverse momentum (see next item), and to reconstruct and identify hadronic  $\tau$  decays. In ATLAS, tracks are used via a multivariate algorithm to identify low transverse momentum jets from pileup interactions. Neutral energy contributions from pileup are corrected by subtracting from the calorimeter jet energy a contribution equal to the product of the jet area and the median energy density of the event. The jet energy scale and resolution are calibrated using the constraint from transversely balanced dijet and multijet events, photon plus jet and  $Z$  plus jet events, and  $E/p$  together with test beam results to extrapolate the absolute calibration to large transverse momenta with insufficient data coverage.

- **Missing transverse momentum** is computed as the negative vector sum of the transverse momenta of all identified objects (leptons, photons, jets, ...), and a contribution denoted soft term from objects originating from the primary event vertex that are not associated to any identified object. ATLAS uses a track-based soft term and CMS uses the particle flow algorithm. The missing transverse momentum magnitude is denoted  $E_T^{\text{miss}}$ .
- **Flavoured jets** containing a  $b$  or  $c$  hadron are identified in the inner tracking detector as a property of a reconstructed jet. The characteristics of (long-lived) weakly decaying heavy flavour hadrons include a displaced secondary vertex, large impact parameter, a large hadron mass, and semi-leptonic decays in 30–40% of the cases. A multivariate algorithm combines the available information to tag jets containing a heavy-flavour hadron (and hence originating from a heavy-flavour quark). The efficiency of  $b$ -tagging is calibrated from data using  $t\bar{t}$  events, muons from heavy flavour decays in dijet events, and using MC simulation to extrapolate the calibration to high transverse momenta. Charm tagging is calibrated using  $W + c$  or  $D^* \rightarrow D^0(\rightarrow K\pi)\pi$  events. Mistag rates are obtained from tracks with negative impact parameters or secondary vertices with negative decay lengths.

#### 4.8 Boosted objects

The high centre-of-mass energy of the LHC can produce highly boosted  $W$ ,  $Z$ ,  $H$  bosons or top quarks so that their hadronic decays are merged into a single jet. This would occur in particular in presence of hypothetical heavy states that decay to massive bosons or top quarks. The identification and reconstruction of such merged objects requires a jet substructure analysis. Boosted signatures can also be used to enhance the signal-to-background ratio in some analyses such as  $H \rightarrow \tau\tau$  and  $H \rightarrow bb$ . Boosted signatures originating from a very hard ISR jet can be used to render visible to the trigger and data analysis collisions with soft final state activity (eg., WIMPs, compressed spectra in supersymmetry or other new physics models).

The average transverse distance between two bodies originating from the decay of a resonance with mass  $m$  and transverse momentum  $p_T$  can be approximated by  $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \approx 2m/p_T$ . A  $W$  boson with  $p_T = 200$  GeV (400 GeV) has  $\Delta R = 0.8$  (0.4). To ensure that all final state objects are fully contained the experiments usually employ so-called “fat jets”, which are jets with radius parameter  $R = 1$  or 1.2, compared to standard anti- $k_t$  jets [17] of  $R = 0.4$ . There exist many strategies to reconstruct the substructure in a fat jet (eg., jet mass), and to correct for pileup effects (jet grooming), see, eg. [18, 19] and references therein.

## 4.9 Systematic uncertainties

Systematic uncertainties are the evil (see figure on the right) in every measurement. Well designed experiments minimise systematic uncertainties by achieving maximum phase space coverage, high measurement precision, response homogeneity and linearity, high calorimeter depth, sufficient longevity of the detector components including resistance against irradiation, etc. The understanding, evaluation and reduction of systematic uncertainties is often the main analysis challenge. A high quality analysis stands out by its thoroughness on all relevant sources of systematic uncertainty. It is thereby important to distinguish relevant from irrelevant sources, where in doubt a source should be considered relevant. For many uncertainty sources, in particular theoretical ones, estimating a “one-sigma” error is very difficult or simply impossible. In such cases conservative uncertainties should be chosen where possible.



(Reasonably) conservative uncertainty estimates are a must! It is of no use to the scientific endeavour to make over-aggressive statements that one cannot fully trust.

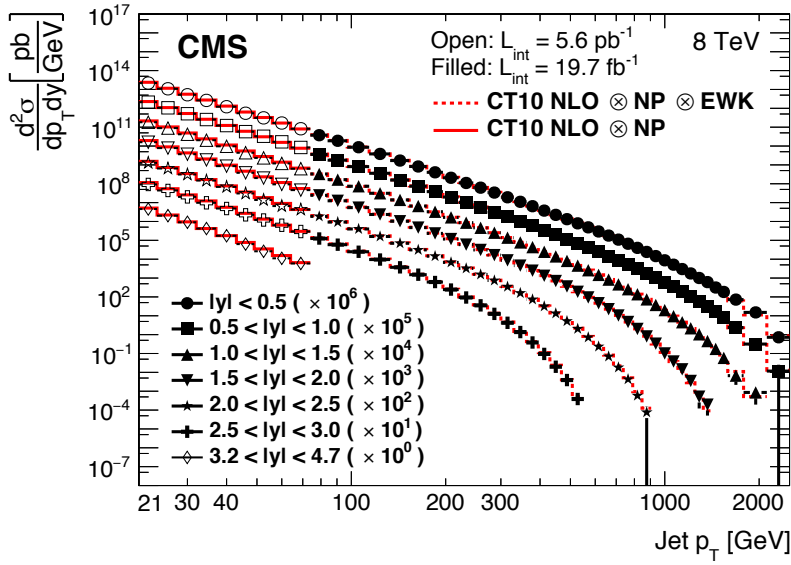
## 5 Physics highlights from the LHC Run-1

The LHC Run-1 featured proton–proton collisions at 7 and 8 TeV with datasets corresponding to approximately 5 and 20 fb<sup>-1</sup> integrated luminosity for ATLAS and CMS, and a total of 3 fb<sup>-1</sup> for LHCb. There are numerous physics highlights published in altogether more than a thousand physics papers. Only a small subset of these are recollected here.

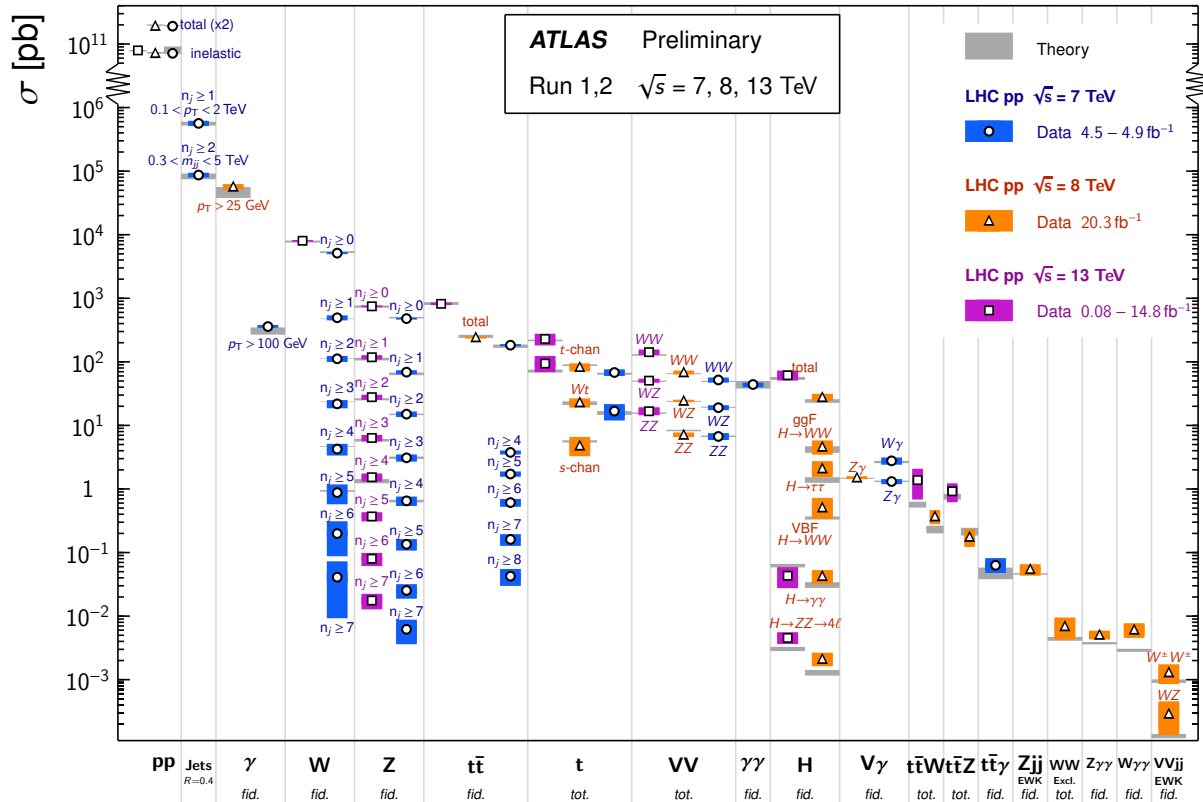
### 5.1 Standard Model and top-quark physics

We should praise the extraordinary match between a plethora of total, fiducial and differential cross-section measurements of all known proton–proton scattering processes and their theoretical predictions, confirming the predictive power of the SM. An example for the measurement of double-differential jet cross-sections by CMS compared to theory prediction is shown in Fig. 7 [20]. Figure 8 gives a summary of ATLAS Run-1 and Run-2 cross-section measurements witnessing the large variety of channels and cross section magnitude, as well as the agreement with the SM predictions. There are many subtleties in this comparison that are not represented in such a summary plot. For example, diboson cross sections exhibit some discrepancy with the NLO SM predictions, which are resolved by moving to NNLO and by taking into account soft-gluon resummation corrections that are needed in case of phase space cuts sensitive to such effects (as, eg., a low- $p_T$  jet veto). A compilation like Fig. 8 delivers a strong statement about the depth of the understanding of hadron collider physics at the highest centre-of-mass energies. It gives confidence that new physics searches, which depend on a good understanding of SM processes, can be reliably performed. We should stress that Run-1 analysis is not over yet: it represents a high-quality, extremely well understood data sample for precision measurements.

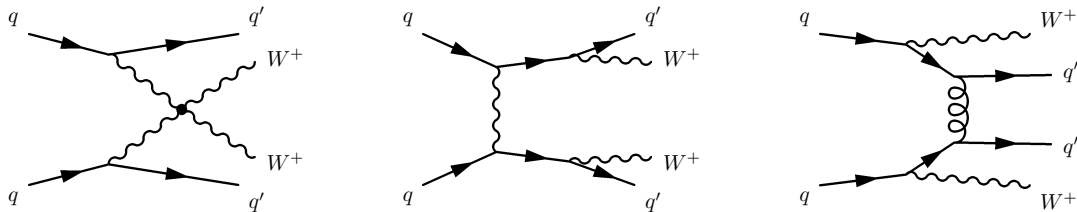
The analysis of Run-1 data allowed first critical electroweak studies of vector-boson scattering (VBS). In electroweak theory the Higgs boson acts as “moderator” to unitarise high-energy longitudinal vector boson scattering. Indeed, if only  $Z$  and  $W$  bosons are exchanged, the amplitude of (longitudinal)  $W_L W_L$  scattering,  $A_{Z,\gamma} \sim v^{-2} \cdot (s+t)$ , rises with the centre-of-mass energy and violates unitarity. Higgs-boson exchange regularises this amplitude via the negative term  $A_H \sim -(m_H^2/v^2) \cdot (s/(s-m_H^2) + t/(t-m_H^2))$ , if  $m_H^2/v^2$  is not too large, which is the case for  $m_H = 125$  GeV. That mechanism can be tested by, eg., measuring same-charge  $W^\pm W^\pm + 2\text{jets}$  production at the LHC (see graphs in Fig. 9). Requiring same-sign  $W^\pm W^\pm$  production greatly reduces strong production (see right-hand graph in Fig. 9) due to the lack of



**Fig. 7:** Double-differential inclusive jet cross sections as function of jet transverse momentum measured by CMS (dots with error bars) in 8 TeV data [20]. The red lines indicate the SM predictions using NLO perturbative QCD and applying nonperturbative (low- $p_T$ ) and electroweak corrections.



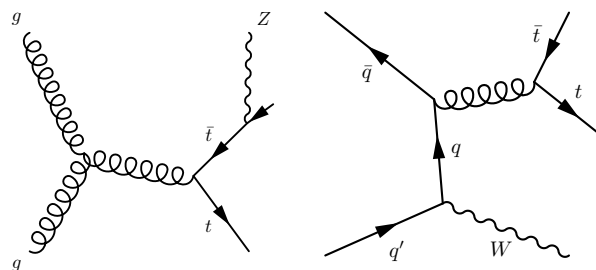
**Fig. 8:** Summary of ATLAS Run-1 and Run-2 cross-section measurements (status August 2016).



**Fig. 9:** Feynman graphs for same-charge  $W^\pm W^\pm + 2\text{jets}$  production. Top: electroweak VBS; middle: electroweak non-VBS; bottom: gluon exchange.

contributions from two initial gluons or one quark and one gluon. It also suppresses the s-channel Higgs amplitude, but moderation through t-channel Higgs exchange remains. The two electroweak processes in Fig. 9 cannot be separated in a gauge-invariant way. Contributions from electroweak VBS to this process can be separated from non-VBS electroweak and strong processes by requiring a large dijet invariant mass and a rapidity gap for hadronic activity. Evidence for electroweak production at the  $3.6\sigma$  ( $2.0\sigma$ ) level was found by ATLAS (CMS) [21, 22].

Strong top-quark pair production has been studied with unprecedented experimental precision at the LHC. Inclusive cross sections are best measured in the dilepton  $e\mu$  final state that is very pure and can be isolated with a minimal set of selection requirements. The measurements provide precise tests of NNLO QCD including leading-logarithmic resummation.<sup>11</sup> In addition, many top properties (mass, charge, charge asymmetry, polarisation, spin correlations, suppressed flavour-changing neutral currents (FCNC), etc.) were measured or probed.



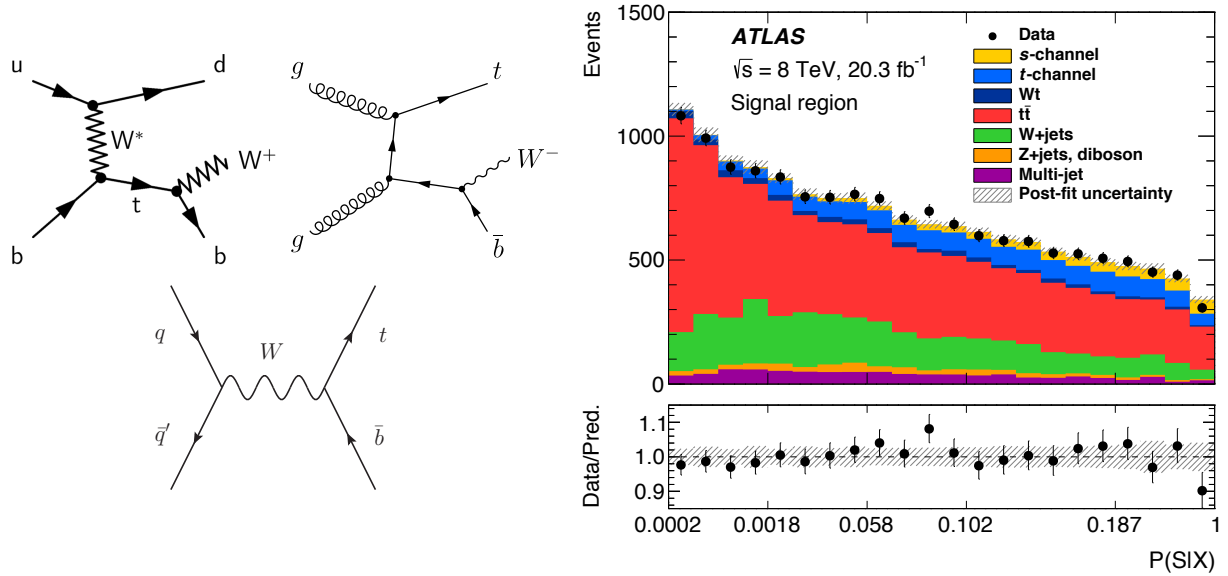
**Fig. 10:** Feynman graphs for leading order  $ttZ$  (left) and  $ttW$  production.

The large luminosity and high centre-of-mass energy also allowed to observe the rare  $tt + W$  and  $tt + Z$  production (see Feynman graphs in Fig. 10) with more than  $7.1\sigma$  combined significance for both modes [23, 24]. The neutral-current  $tZ$  coupling is directly probed in  $tt + Z$ .

Run-1 also allowed detailed studies of electroweak single top production and property measurements (see Fig. 11 for representative leading-order diagrams). Single top cross sections are enhanced at the LHC compared to the Tevatron: at 8 TeV LHC centre-of-mass energy, factors of 42 (t-channel), 31 ( $Wt$ ), but only 5 for s-channel production so that the signal to background ratio is worse for the latter channel at the LHC. Production of t-channel single top has been studied in great differential detail already [25, 26]. The separate measurement of  $tq$  and  $\bar{t}q$  production provides sensitivity to  $u$  and  $d$  quark PDFs. Inclusive  $Wt$  channel production was clearly observed by both ATLAS and CMS [27, 28]. Production via an s-channel process (see bottom diagram in Fig. 11) was recently observed by the Tevatron experiments with  $6.3\sigma$  combined significance in agreement with the SM prediction [30]. ATLAS reported an observed evidence of  $3.2\sigma$  (for  $3.9\sigma$  expected significance), also in agreement with the SM prediction [29].

<sup>11</sup>The cross section of soft gluon emission is infrared divergent (eikonal factor). The divergence is cancelled by virtual corrections up to logarithmic leftover terms  $\sigma(a \rightarrow b) \rightarrow \sigma \ln^2(1 - m_b^2/s)$ , which need to be resummed. Several resummation strategies denoted “threshold resummation”, “transverse momentum resummation”, or “high-energy resummation” exist in the literature.





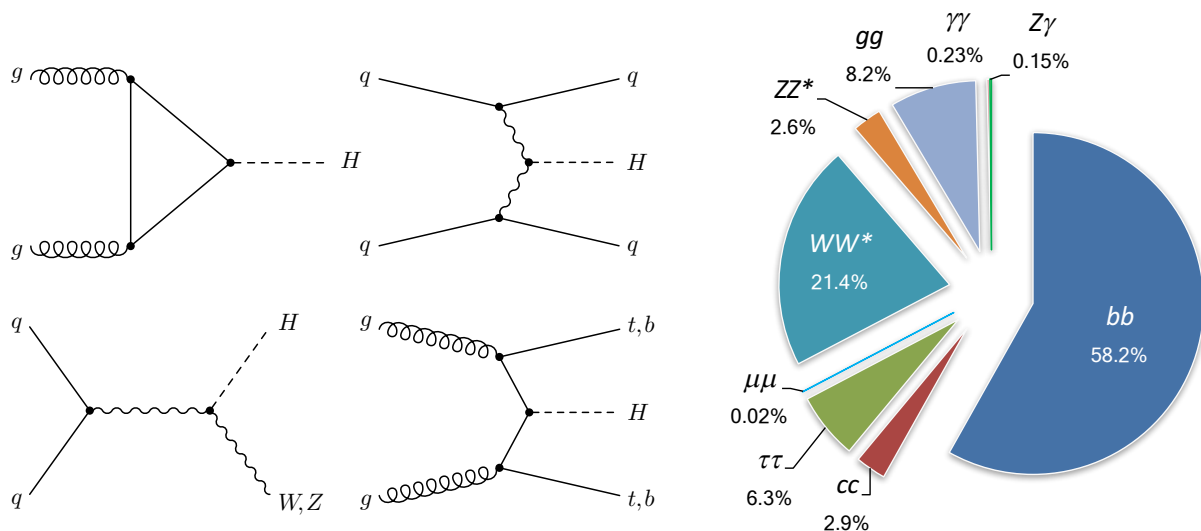
**Fig. 11:** Left: leading-order Feynman graphs for t-channel (top left),  $Wt$ -channel (top right), and s-channel (bottom) single top production at the LHC. Right: distribution of a discriminant variable in an ATLAS search for s-channel single top production [29].

## 5.2 Higgs boson physics

Among all the Run-1 physics results, the discovery of the Higgs boson is (so far) the magnum opus [31, 32] (articles that each have collected about 6 700 citations to date). The Higgs boson had been vainly searched for at many accelerators. The most stringent non-LHC limits came from the Large Electron–Positron Collider (LEP at CERN, 1989–2000) and the proton–antiproton collider Tevatron (Fermilab, 1990–2011) excluding at 95% confidence level  $m_H < 114$  GeV [33] and  $149 < m_H < 182$  GeV [34], respectively. Global fits to electroweak precision data (see Section 6) constrained the Higgs boson mass via logarithmic corrections and excluded about  $m_H > 160$  GeV at 95% confidence level [35].

At the LHC the Higgs boson is dominantly produced via gluon fusion with a cross section of 19.3 pb at 8 TeV for  $m_H = 125$  GeV [36] (see also [37] for a recent review on Higgs boson physics). The cross section steeply falls with the Higgs boson mass. Additional production modes are weak boson fusion (VBF) with 1.6 pb, associated production with a weak boson (also denoted Higgs-strahlung) with 0.70 pb (0.42 pb) for  $WH$  ( $ZH$ ), and associated production with a  $t\bar{t}$  or  $b\bar{b}$  pair ( $ttH$ ,  $bbH$ ) with 0.13 pb and 0.20 pb, respectively (cf. Fig. 12 for the corresponding Feynman diagrams). The uncertainties in the predictions are larger (7~14%) for the gluon initiated processes than for the quark initiated ones (~3%, dominated by PDF uncertainties). The inclusive 8 TeV Higgs cross section amounts to 22 pb. In total, about 470 thousand SM Higgs bosons of 125 GeV were produced in 2012 at 8 TeV in each ATLAS and CMS.

Because of the coupling to the mass of the decay particles ( $\propto m_V^2, m_f$ ) the Higgs boson decays with preference to the heaviest particles allowed. It does not couple directly to photons and gluons but proceeds via loops involving preferentially heavy particles (eg., top,  $W$  boson). The branching fractions predicted for an SM Higgs boson of mass 125 GeV are shown on the right panel of Fig. 12. The theoretical uncertainty in these predictions ranges from 3% to about 12%. The leptonic ( $\ell = e, \mu$ ) and photon final states provide the best discovery sensitivity. The decays  $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$  provide the best mass resolution (1–2% for  $m_H = 125$  GeV). The decay  $H \rightarrow WW^{(*)} \rightarrow 2\ell 2\nu$  (~20% mass resolution due to the neutrinos in the final state) has a good trigger, a sustainable background level, and large branching



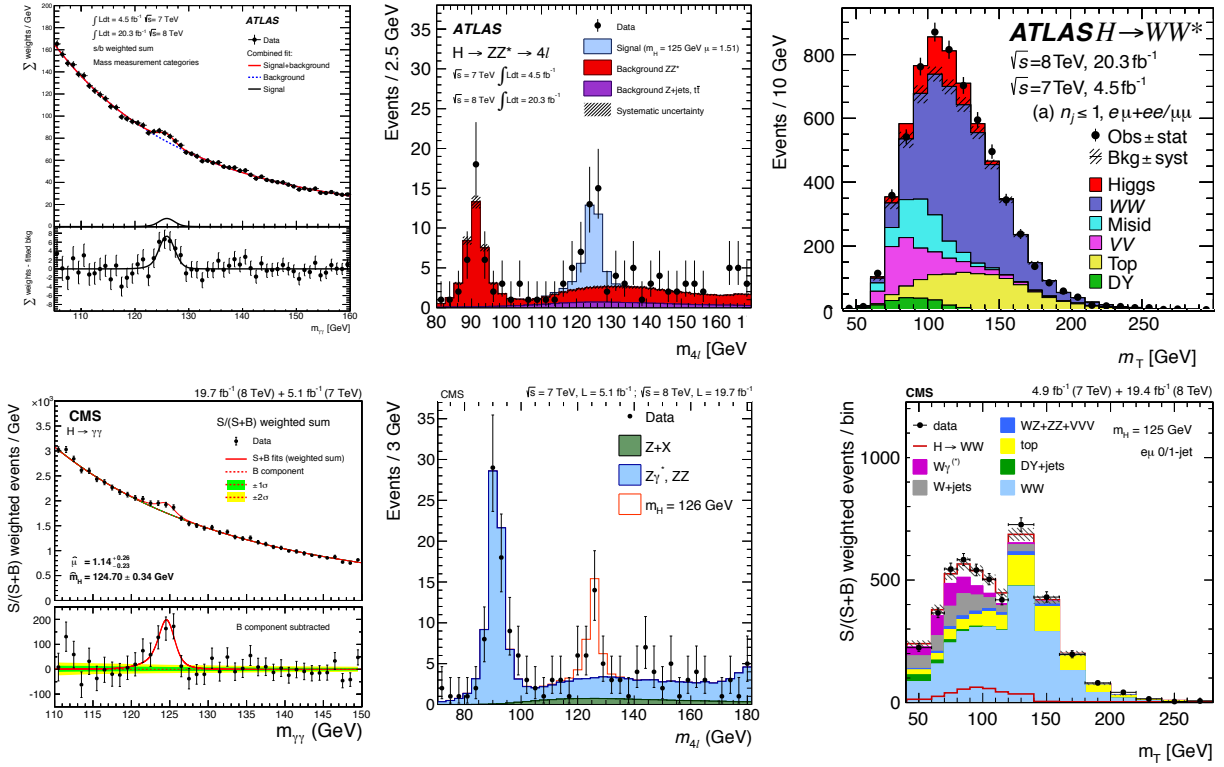
**Fig. 12:** Left: Feynman graphs for the dominant Higgs production channels: gluon fusion via (dominantly) a top quark triangle (top left), weak boson fusion (top right), associated production with a weak boson (bottom left) and with a heavy quark pair (bottom right). Right: branching fractions predicted for an SM Higgs boson of mass 125 GeV [36]. Considering only leptonic decays to  $e, \mu$ , the exploitable branching fractions to  $WW^*$  and  $ZZ^*$  are 1.1% and 0.012%, respectively.

fraction. The fermionic modes  $H \rightarrow \tau\tau$  and  $H \rightarrow bb$  have mass resolutions of about 10% and 15%, respectively, and are more challenging to detect due to large backgrounds. The decays  $H \rightarrow \mu\mu$  and  $H \rightarrow Z(\rightarrow \ell\ell)\gamma$  have excellent mass resolution but too low branching fractions to be in reach with the current datasets.

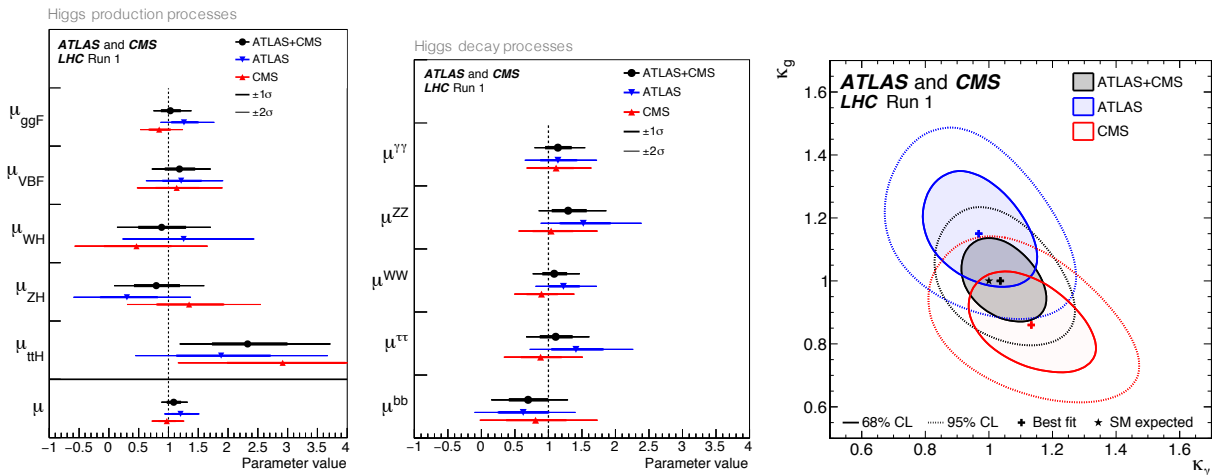
It is fortunate that at  $m_H = 125$  GeV many decays of the Higgs boson are experimentally accessible. The phenomenological aspects of that mass might appear less appealing as we will see later. The dominant  $H \rightarrow bb$  mode is only exploitable in association with  $W/Z$  or  $tt$ . Their leptonic decays provide a trigger signal and help to reduce the overwhelming background from strong interaction  $bb$  continuum production,  $\sigma(bb) \sim \mathcal{O}(100 \mu\text{b})$ . A boost of the Higgs boson helps to improve the signal purity at the expense of reduced efficiency.

There is no doubt about the discovery of the Higgs boson. Each of the most sensitive bosonic channels  $H \rightarrow \gamma\gamma$ ,  $H \rightarrow 4\ell$  and  $H \rightarrow 2\ell 2\nu$  from ATLAS and CMS have achieved an independent observation (cf. Fig. 13) [38–43]. The combination of ATLAS and CMS mass measurements gives  $m_H = 125.09 \pm 0.21_{\text{stat}} \pm 0.11_{\text{syst}}$  GeV [44]. There are very different experimental challenges in each Higgs channel. All analyses have constantly increased their sensitivity during Run-1 owing to improved understanding of lepton reconstruction and calibration, as well as improved background modelling and signal against background discrimination.

In addition to sophisticated individual analyses, ATLAS and CMS have joined forces and combined their Higgs mass and coupling measurements [44, 45]. These combinations represent the full picture of what the experiments have learned in a framework that consistently treats all processes in terms of production mechanism and decay. Figure 14 shows as an example the ratios of measured to predicted signal strengths per production process (left panel, assuming the Higgs decays to proceed according to the SM), and vice versa per decay channel (middle panel, assuming SM Higgs production) [45]. The overall signal strength, assuming an overall scale for all individual signal strengths, is measured to be  $\mu = 1.09 \pm 0.11$ . The right hand panel of Fig. 14 shows the results of a fit of leading order coupling modifiers [46] to



**Fig. 13:** Reconstructed inclusive Higgs candidate masses in the bosonic decay channels  $H \rightarrow \gamma\gamma$  (left),  $H \rightarrow 4\ell$  (middle), and  $H \rightarrow 2\ell 2\nu$  (right, shown is the transverse mass) for the ATLAS (top row) and CMS (bottom row) Run-1 analyses [38–43].



**Fig. 14:** Left: Higgs production signal strengths for ATLAS and CMS and their combination. Also shown is the measurement of the global signal strength. Middle: Higgs decay signal strengths for ATLAS, CMS and their combination. Right: 68% and 95% confidence level contours in the Higgs-to-gluon versus Higgs-to-photon coupling modifiers for ATLAS, CMS and their combination. The SM prediction is  $\kappa_g = \kappa_\gamma = 1$ . See text for the assumptions underlying these plots. The figures are taken from [45].

the combined Higgs boson data, where for a given production process or decay mode, denoted  $j$ , the coupling modifier  $\kappa_j$  is defined such that  $\kappa_j^2 = \sigma_j/\sigma_j^{\text{SM}}$ . Shown in the figure are the coupling modifiers  $\kappa_g$  versus  $\kappa_\gamma$  of the Higgs-to-gluon and Higgs-to-photon couplings, respectively. The fit was performed by constraining all the other coupling modifiers to their SM values and assuming no non-SM decays of the Higgs boson. The resulting agreement with the SM is remarkable as these couplings proceed through loops involving heavy fermions and also, in the photon case, bosons ( $W$ ). It is a powerful probe for new heavy degrees of freedom. For example, the result allows to reject a theory with heavy fermions with SM-like Yukawa couplings.<sup>12</sup> The ATLAS and CMS Higgs coupling combination exhibits agreement among the two experiments. It yields sufficient significance for the observation of the Higgs decay to fermions,  $H \rightarrow \tau\tau$  [47, 48], and of VBF production. The  $t\bar{t}H$  process [49, 50] comes out a bit large with a relative signal strength of  $\mu = 2.3$  and a combined observed significance of  $4.4\sigma$  (for  $2.0\sigma$  expected). With respect to the signal strengths shown in Fig. 14 we note that the least model-dependent observables at the LHC are coupling ratios rather than absolute coupling measurements [45].

The Higgs boson has been suggested to possibly act as a “portal” to new physics responsible for dark matter. In such models, a massive dark matter particle couples only weakly (or not at all) with the SM particles, except for the Higgs boson.<sup>13</sup> If the dark matter particle is not too heavy, the Higgs decays invisibly to it and is searched for via, eg., a VBF topology where the forward jets are used to trigger and select the events [51, 52]. Limits of about 25% are currently set for an invisible Higgs boson decay. In general, owing to its low mass and consequently narrow width of 4.1 MeV compared to the widths of the  $W$ ,  $Z$  or top quark of 2.1 GeV, 2.5 GeV and 1.3 GeV, respectively, the Higgs boson has good sensitivity to new physics as even small couplings to new states (if light enough) can measurably impact its branching fractions (see [53] for an analysis of constraints on new physics from the measurements of the Higgs couplings and invisible decays). It is therefore important to continue to measure the Higgs couplings, including the invisible one, with highest possible precision.

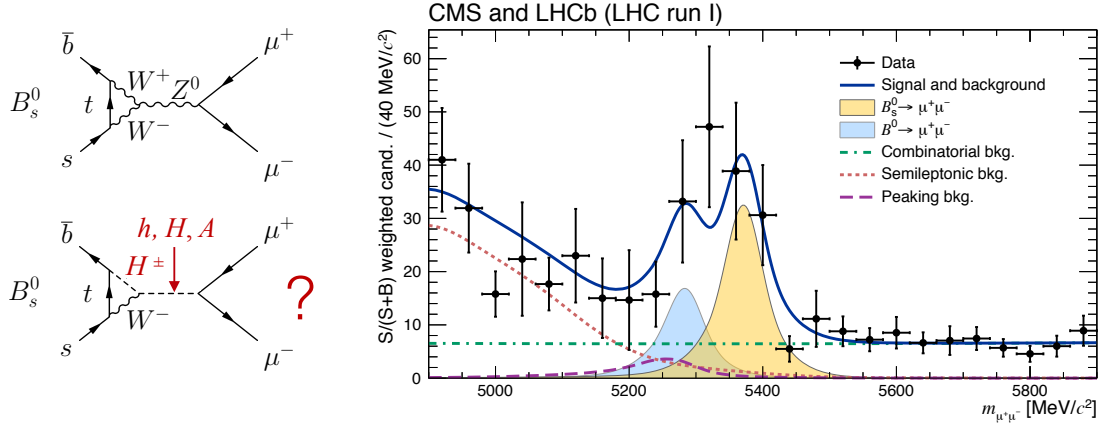
### 5.3 Heavy flavour physics

There have been beautiful flavour and low- $p_T$  physics measurements at the LHC. The LHCb experiment has produced a flurry of important results among which the observation, together with CMS [54], of the very rare decay  $B_s \rightarrow \mu\mu$  at a branching fraction of  $2.8_{-0.6}^{+0.7} \cdot 10^{-9}$  in agreement with the SM prediction of  $3.7 \pm 0.2 \cdot 10^{-9}$  [55]. The left panel of Fig. 15 shows representative SM and BSM Feynman graphs, and the right panel shows the combined CMS and LHCb data and the result of a simultaneous signal and background fit. The  $B_s \rightarrow \mu\mu$  decay proceeds through a loop as there is no tree-level FCNC in the SM. It is in addition CKM and helicity suppressed, thus the low branching fraction. The decay is sensitive to additional scalar bosons as, eg., predicted in supersymmetry. ATLAS recently published the Run-1 result giving a branching fraction value in agreement with CMS and LHCb and approximately  $2\sigma$  below the SM prediction [56].

Another high-priority flavour result from the LHC is the measurement of the mixing-induced CP violation parameter  $\phi_s$  in a flavour-tagged, time-dependent  $B_s \rightarrow J/\psi\phi$  analysis. That measurement represents one of the most sensitive CP-violation tests of the SM as  $\phi_s$  is small and predicted with negligible theoretical uncertainty within the CKM paradigm. ATLAS [57], CMS [58] and LHCb [59] have measured simultaneously  $\phi_s$  and  $\Delta\Gamma_s$ , the width difference of the two  $B_s$  mass eigenstates, with LHCb exhibiting the best precision. It found the combined result  $\phi_s = -0.010 \pm 0.039$  rad in agreement with the SM. Fig-

<sup>12</sup>Naively an additional heavy fermion generation would increase the gluon fusion Higgs cross section by a factor of nine with respect to the SM prediction due to the quadratic fermion form-factor dependence of the cross section.

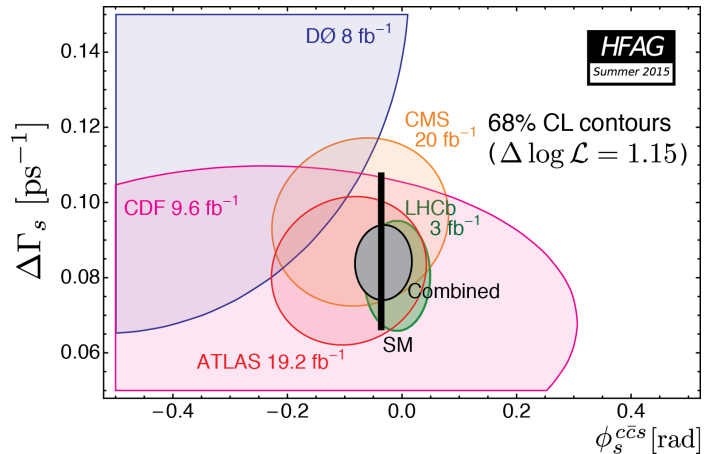
<sup>13</sup>For example in a Dirac neutrino case the massive right-handed neutrinos would transform as singlets under the SM gauge interactions, but would couple to the Higgs boson.



**Fig. 15:** Left: Feynman graphs for the decay  $B_s \rightarrow \mu\mu$  in the SM (top) and beyond the SM (bottom). Right: weighted distribution of the dimuon invariant mass for all CMS and LHCb measurement categories. The yellow region indicates the fit result for  $B_s \rightarrow \mu\mu$ , while the blue region shows that for the doubly CKM suppressed decay  $B_d \rightarrow \mu\mu$ . The dashed and dotted lines indicate the various background contributions as obtained from the fit (the abundance of the peaking background is predicted) and the solid line their sum [54].

ure 16 shows the various measurements as well as their combination as 68% confidence level contours in the  $\Delta\Gamma_s$  versus  $\phi_s$  plane. The SM prediction is indicated by the black vertical bar.

LHCb also contributed significantly to the long-term effort to overconstrain the CKM matrix in what is known as the unitarity triangle, a triangle given in the  $\bar{\rho}-\bar{\eta}$  CKM parameter plane, where  $\bar{\eta} \neq 0$  stands for CP violation in the SM. LHCb has engaged in a vigorous programme to determine the unitarity triangle angle  $\gamma \sim \arg(-V_{ub}^*)$ . It can be measured through interference of  $b \rightarrow u$  with  $b \rightarrow c$  tree transitions where hadronic amplitude parameters are determined simultaneously with  $\gamma$  from the data. A combined fit [60], dominated by the measurements from charged  $B^+$  to charm decays, gives  $\gamma = 70.9_{-8.5}^{+7.1}$  deg, which is in agreement with the prediction from the CKM fit (not including the direct  $\gamma$  measurements) of  $68 \pm 2$  deg [61]. LHCb also measured the ratio  $|V_{ub}/V_{cb}|$  from  $\Lambda_b \rightarrow p\mu\nu$  (a baryon decay!) with 5% precision [62]. The result is closer to the exclusive  $B$ -factory numbers for  $|V_{ub}|$ , which exhibit a tension with the larger inclusive numbers. Furthermore LHCb obtained the world's best single  $\Delta m_d$  measurement [63]  $0.5050 \pm 0.0021 \pm 0.0010$   $\text{ps}^{-1}$  (the  $B$ -factories have a combined uncertainty of  $0.005$   $\text{ps}^{-1}$ ), a  $\sin(2\beta)$  measurement [64] of  $0.731 \pm 0.035 \pm 0.020$  that approaches the precision of the  $B$ -factories, the world's best constraints on CP violation in  $B_{(s)}^0$  mixing ( $\alpha_{\text{sl}}^s, \alpha_{\text{sl}}^d$ ) in agreement with the SM (D0 sees a  $3.6\sigma$  deviation), and a search for CPT violation [67] (difference in mass or width) in the  $B_{(s)}^0$  systems together with the measurement of sidereal phase dependence of the CPT violating parameter.



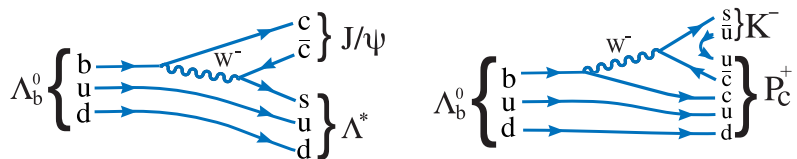
**Fig. 16:** Contours of 68% confidence level for  $\Delta\Gamma_s$  versus the mixing induced CP-violation parameter  $\phi_s$ , [57–59]. The SM prediction is indicated by the black vertical bar.

It is interesting to speculate about the “relevance” of the CKM phase. So far, all CP violating effects

measured in particle physics can be reduced to just that phase. On the other hand, there seems to be consensus of opinion among theorists that the CKM induced CP violation in the quark sector is too small by many orders of magnitude to generate the observed baryon asymmetry in the universe (non-zero CKM CP violation requires non-zero and non-degenerate quark masses, so the baryogenesis could only be generated during the electroweak phase transition at critical temperature of  $T_c \sim 100$  GeV). So is the CKM phase only an ‘‘accident of Nature’’? Because there are three quark generations, there is a phase in the quark mixing matrix,<sup>14</sup> and so that phase has ‘‘some’’ value? What would happen to the universe had we a dial to change that value [68]?

Several measurements in the flavour sector exhibit non-significant but interesting anomalies with respect to theory predictions. A prominent example is given by angular coefficients describing the transition  $b \rightarrow s\mu^+\mu^-$ , the prediction of which, however, are plagued by hadronic uncertainties. Theoretically robust are universality tests. Such tests were performed at the per-mil level at LEP and other  $e^+e^-$  colliders not showing any significant discrepancy from the expectation of universal lepton coupling. The  $B$ -factory experiments and LHCb have measured ratios of semileptonic  $B$  decays among which [69–72]  $R_{D^{(*)}} = \mathcal{B}(B^0 \rightarrow D^{(*)}\tau\nu)/\mathcal{B}(B^0 \rightarrow D^{(*)}\ell\nu)$  and  $R_K = \mathcal{B}(B^+ \rightarrow K^+\mu^+\mu^-)/\mathcal{B}(B^+ \rightarrow K^+e^+e^-)$ . The Heavy Flavour Averaging Group (HFAG) has combined the experimental results giving [73]  $R_{D^*} = 0.316 \pm 0.016 \pm 0.010$ , which is  $3.3\sigma$  away from the SM prediction  $0.252 \pm 0.003$  [74]. The two-dimensional combination with  $R_D$  increases the deviation to  $4.0\sigma$ . For  $R_K$  LHCb measures at low  $q^2$  (given by the invariant mass of the dimuon or dielectron system) the value  $0.745^{+0.090}_{-0.074}(\text{stat}) \pm 0.036(\text{syst})$ , which differs by  $2.6\sigma$  from the expected unity [75].

An intriguing observation in hadron spectroscopy was announced by LHCb in summer 2015 in a paper [76] that collected over 250 citations since. It is the observations of exotic structures in the  $J/\psi p$  channel, consistent



**Fig. 17:** Feynman graphs for  $\Lambda_b^0 \rightarrow J/\psi\Lambda^*$  (left) and  $\Lambda_b^0 \rightarrow P_c^+K^-$  (right).

with pentaquark-charmonium states occurring in  $\Lambda_b^0 \rightarrow J/\psi K p$  decays (see Fig. 17 for representative  $\Lambda_b^0$  decay diagrams). Analysing the full Run-1 data sample and performing an intricate three-body amplitude analysis, the observed structures could only be described by adding two resonances, one with mass and width of  $4380 \pm 8 \pm 29$  MeV and  $205 \pm 18 \pm 86$  MeV, respectively, and the other (narrower) with mass and width of  $4449.8 \pm 1.7 \pm 2.5$  MeV and  $39 \pm 5 \pm 19$  MeV. LHCb dubs these two states  $P_c(4380)^+$  and  $P_c(4450)^+$ . The binding mechanism for pentaquarks is not clear at present. They may consist of five quarks tightly bound together, but it is also possible that they are more loosely bound and consist of a three-quark baryon and a two-quark meson interacting relatively weakly in a meson-baryon molecule.

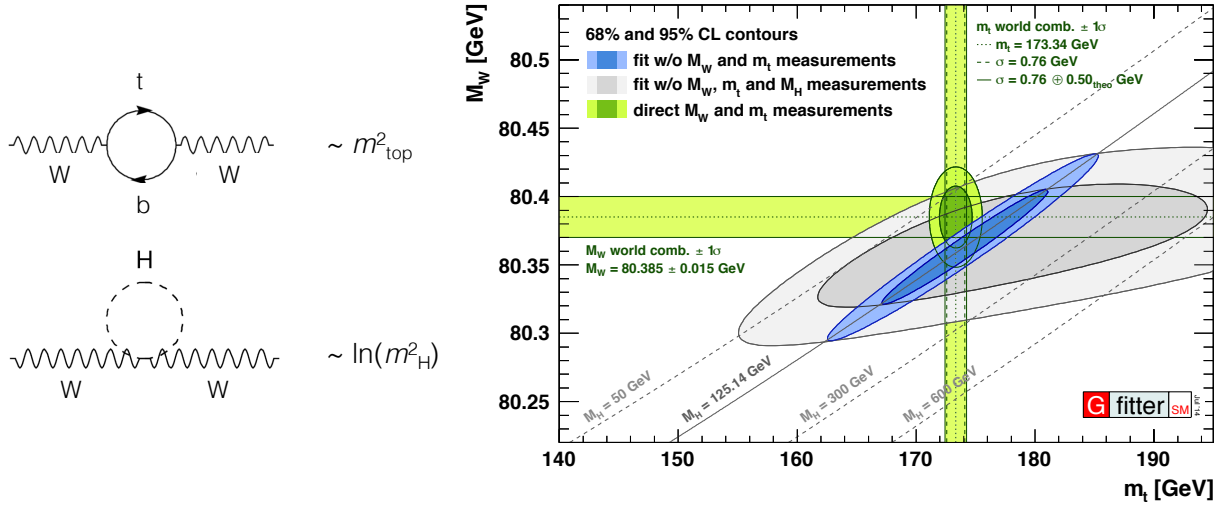
## 6 Digression on electroweak precision measurements

The global electroweak fit relating observables of the electroweak SM to each other by incorporating precise theoretical predictions of radiative corrections was a masterpiece of the LEP/SLC era. It led to the prediction of the top-quark mass prior to its discovery, provided a strong (logarithmic) constraint on the mass of the Higgs boson, predicting it to be light, and allowed to exclude or constrain models beyond the SM.<sup>15</sup> The discovery of the Higgs boson overconstrains the fit and dramatically improves its

<sup>14</sup> There would be zero (three) phases in a two (four) generations SM, as  $n_{\text{phases}} = (n_{\text{gen}} - 1)(n_{\text{gen}} - 2)/2$ .

<sup>15</sup>For example, it allowed to exclude the simplest technicolour models [77–79]. Technicolour invokes the existence of strong interactions at a scale of the order of a TeV and induces strong breaking of the electroweak symmetry. In the original form of





**Fig. 18:** Left: Feynman graphs of radiative corrections contributing to the  $W$  boson mass. The upper (lower) graph introduces a quadratic (logarithmic) top-quark (Higgs-boson) mass dependence. The right panel shows 68% and 95% confidence level contours obtained from scans of fits with fixed variable pairs  $m_W$  versus  $m_t$ . The narrower blue and larger grey allowed regions are the results of the fit including and excluding the  $m_H$  measurement, respectively [80]. The horizontal bands indicate the  $1\sigma$  regions of the  $m_W$  and  $m_t$  measurements (world averages).

predictability. The fit has thus turned into a powerful test of the SM.

Figure 18 shows the Feynman graphs of radiative corrections contributing to the  $W$  boson mass. They introduce a quadratic top-quark and logarithmic Higgs-boson mass dependence of the correction parameter  $\Delta r$  occurring in the relation

$$m_W^2 = \frac{m_Z^2}{2} \left( 1 + \sqrt{1 - \frac{\sqrt{8}\pi\alpha(1+\Delta r)}{G_F m_Z^2}} \right), \quad (5)$$

owing to electroweak unification. Similarly, the effective weak mixing angle,  $\sin^2\theta_{\text{eff}}^\ell$ , for lepton flavour  $\ell$  depends on  $m_W$  and  $m_Z$  and, via radiative corrections and by replacing  $m_W$ , on the top-quark and Higgs-boson masses. The current predictions of the observables that most benefit from the known Higgs mass, split into the various uncertainty terms, are [80]

$$\begin{aligned} M_W &= 80.3584 \pm 0.0046_{m_t} \pm 0.0030_{\delta_{\text{theo}} m_t} \pm 0.0026_{M_Z} \pm 0.0018_{\Delta\alpha_{\text{had}}} \\ &\quad \pm 0.0020_{\alpha_S} \pm 0.0001_{M_H} \pm 0.0040_{\delta_{\text{theo}} M_W} \text{ GeV}, \\ &= 80.358 \pm 0.008_{\text{tot}} \text{ GeV}, \end{aligned} \quad (6)$$

and

$$\begin{aligned} \sin^2\theta_{\text{eff}}^\ell &= 0.231488 \pm 0.000024_{m_t} \pm 0.000016_{\delta_{\text{theo}} m_t} \pm 0.000015_{M_Z} \pm 0.000035_{\Delta\alpha_{\text{had}}} \\ &\quad \pm 0.000010_{\alpha_S} \pm 0.000001_{M_H} \pm 0.000047_{\delta_{\text{theo}} \sin^2\theta_{\text{eff}}^\ell}, \\ &= 0.23149 \pm 0.00007_{\text{tot}}. \end{aligned} \quad (7)$$

Their total uncertainties of 8 MeV and  $7 \cdot 10^{-5}$ , respectively, undercut the world average experimental errors of 15 MeV and  $16 \cdot 10^{-5}$  [4, 81].

technicolor, the strong interactions themselves trigger electroweak symmetry breaking without the need of a Higgs boson.

The LHC experiments, as do CDF and D0 since long and continuing, are investing efforts into precision measurements of the electroweak observables  $m_W$ ,  $m_t$ , and  $\sin^2\theta_{\text{eff}}^\ell$ . All are extremely challenging.

### 6.1 Top-quark mass

There has been significant progress on the top-quark mass measurements at the LHC achieving similar precision as those performed by the Tevatron experiments. The currently most accurate LHC number is the CMS Run-1 combination of measurements, based on the kinematic top mass reconstruction and comparison with MC templates [82], giving  $m_t = 172.44 \pm 0.13 \pm 0.47$  GeV, where the first uncertainty is statistical and the second systematic. The ATLAS Run-1 combination, not yet including the lepton + jets, reads  $m_t = 172.84 \pm 0.34 \pm 0.61$  GeV [83]. The most recent Tevatron combination is  $m_t = 174.34 \pm 0.37 \pm 0.52$  GeV [84] that shows a tension of  $2.4\sigma$  or more with the CMS result.

While these kinematic mass measurements provide the best current precision on  $m_t$  and must be continued, it is also apparent that they approach a difficult systematic uncertainty regime from, mostly, the  $b$ -quark fragmentation. A way to improve could be to choose more robust observables with respect to the leading systematic effects at the possible price of losing statistical power. The dilepton kinematic endpoint is an experimentally clean observable, which has however large theoretical uncertainties [85]. More robust could be the selection of charmonium states [86] or charmed hadrons originating from a  $b$ -hadron produced in one of the  $b$ -jets. These provide a clean but rare signature.

ATLAS and CMS also indirectly determine the top mass from inclusive and differential cross-section measurements. These are promising approaches benefiting from theoretically well defined observables, which are however not yet competitive with the kinematic methods. They also strongly depend on the assumption that no new physics contributes to the measured cross sections. The currently best top pole mass determination from CMS [87] using a precise Run-1  $e\mu$ -based cross-section measurement is  $173.8_{-1.8}^{+1.7}$  GeV in agreement with the direct (kinematic) measurements.

### 6.2 Weak mixing angle

The CDF, D0 [91] and LHC experiments [88–90] have extracted the weak mixing angle from  $Z/\gamma^*$  polarisation measurements. The total uncertainty on  $\sin^2\theta_{\text{eff}}^\ell$  at the Tevatron is dominated by statistical effects, that of LHCb has similar statistical and systematic contributions, while for ATLAS and CMS parton density function (PDF) uncertainties are dominant. A data-driven “PDF replica rejection” method applied by CDF allows to reduce the sensitivity to PDF and update the measurement when improved PDF sets are available. Overall, these are complex measurements (in particular with respect to the physics modelling) that are important to pursue also in view of a better understanding of  $Z/\gamma^*$  production at hadron colliders. The precision obtained is however not yet competitive with that of LEP/SLC.

### 6.3 W-boson mass

The  $W$  boson was discovered at CERN’s SPS in 1983. A first measurement of its mass by the UA1 experiment in 1983 at centre-of-mass energy of 546 GeV gave  $m_W = 81 \pm 5$  GeV [65]. In 1992, at  $\sqrt{s} = 630$  GeV UA2 achieved  $80.35 \pm 0.37$  GeV using  $m_Z$  from LEP as reference calibration [66]. A factor of ten improvement in precision was obtained at LEP with the most recent combination giving  $80.376 \pm 0.033$  GeV. That precision has been undercut by the Tevatron experiments whose latest average, using proton–antiproton collision data taken at  $\sqrt{s} = 1.96$  TeV, is  $80.387 \pm 0.016$  GeV. The combination of the Tevatron and LEP results leads to the present world average  $m_W = 80.385 \pm 0.015$  GeV [4, 81].



While the LEP analyses are final, the Tevatron experiments are continuing to improve their precision and updated results can be expected in the future.

It likely came as a surprise to many in the particle physics community that such a precision measurement is now dominated by a hadron collider, which was not built with that goal in mind. The  $W$  boson mass is arguably the hardest measurement in high-energy physics, needing about seven years to be accomplished. Also the LHC was not built to measure the  $W$  boson mass, but to discover new particles. There is an unfavourable environment at the LHC compared to  $e^+e^-$  or proton–antiproton colliders. At the Tevatron,  $W$  boson production is dominated by the valence quarks of the proton. At the LHC on the contrary, sea and thus heavy quarks are much more important. This difference affects all aspects of the measurement: detector calibration, transfer from the  $Z$  to the  $W$  boson, PDF uncertainties,  $W$  polarisation, modelling of the  $W$  transverse momentum. It is thus a very challenging undertaking, but also a very interesting one: a lot can be learned on the way!

The measurement of the  $W$ -boson mass at the LHC using the leptonic  $W$  boson decay relies on an excellent understanding of the final state. The observables that probe  $m_W$  are the transverse momentum of the lepton ( $p_{T,\ell}$ ), the transverse momentum of the neutrino ( $p_{T,\nu}$ ), measured from the transverse recoil of the event, and the transverse mass of the lepton-neutrino system ( $m_T$ ). The measurement requires a high-precision momentum and energy scale calibration (including the hadronic recoil) obtained from  $Z$ ,  $J/\psi$  and  $\Upsilon$  data, and excellent control of the signal efficiency and background modelling. The biggest challenge is posed by the physics modelling. The production is governed by PDF and initial state interactions (perturbative and nonperturbative), that can be constrained by  $W^+$ ,  $W^-$ ,  $Z$ , and  $W + c$  data, and the use of NNLO QCD calculations including soft gluon resummation. The experimental  $m_W$  probes are very sensitive to the  $W$  polarisation (and hence to PDF, including its strange density). Electroweak corrections are sufficiently well known.

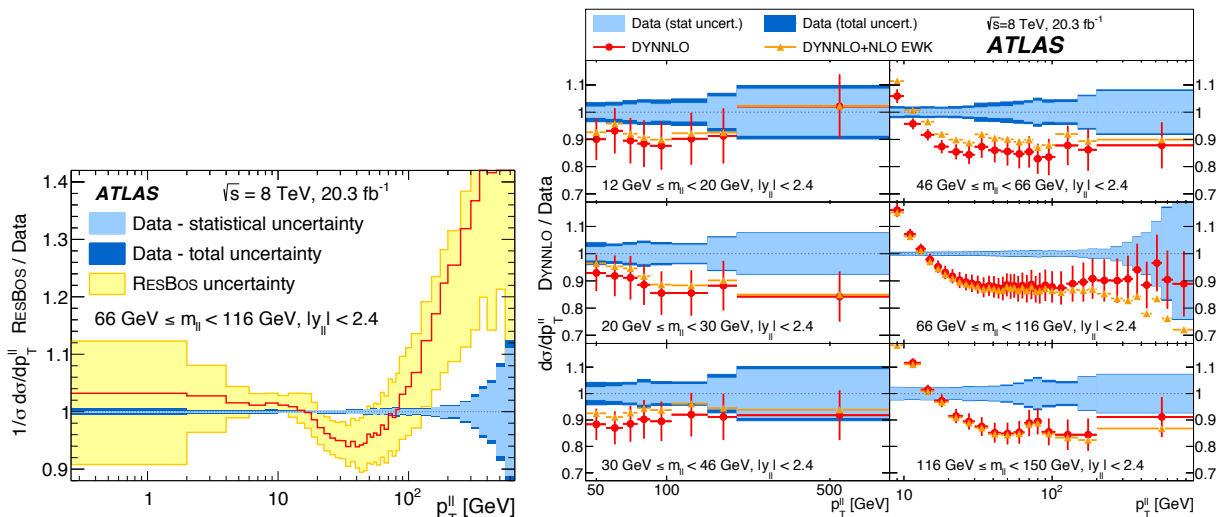
The experiments are thriving to address the above issues. Many precision measurements (differential  $Z$ ,  $W + X$  cross sections, polarisation analysis, calibration performance, etc.) are produced on the way with benefits for the entire physics programme. Theoretical developments are also mandatory. Altogether this is a long-term and iterative effort.

CMS presented for the first time a  $m_Z$  measurement using a  $W$ -like  $Z \rightarrow \mu^+\mu^-$  analysis where one muon is replaced by a neutrino that contributes to the missing transverse momentum in the event [92]. It represents a proof-of-principle, although differences with the full  $m_W$  analysis remain in the event selection, the background treatment and most of the physics modelling uncertainties. CMS used the 7 TeV dataset to take benefit from the lower number of pileup interactions. The momentum scale and resolution calibration for that measurement relies on  $J/\psi$  and  $\Upsilon$  data. Track-based missing transverse momentum is used and the  $W$  transverse recoil is calibrated using  $Z + \text{jets}$  events. The results for the different probes and the positive and negative  $W$ -like cases are found to agree with the LEP measurement. The uncertainties, depending on the probe used, are: statistical: 35–46 MeV, total systematic: 28–34 MeV, QED radiation:  $\sim 23$  MeV (dominant), lepton calibration: 12–15 MeV.

ATLAS and CMS use precise measurements of the  $Z$  boson  $p_T$  to tune the  $p_T$  modelling of the  $W$  boson, which relies on NNLO and NNLL/resummed calculations. But: different generators predict different transfers from  $Z$  to  $W$ . In addition, PDFs play different roles in  $Z$  and  $W$  production. Figure 19 shows normalised differential cross section ratios measured by ATLAS [93] of resummed NLO predictions from ResBos<sup>16</sup> [94] to data (left) and NNLO predictions using the DYNNLO programme<sup>17</sup> [95] without

<sup>16</sup>ResBos features ISR at approximate NNLO,  $\gamma^*-Z$  interference at NLO, NNLL soft-gluon resummation, no FSR or hadronic event activity, CT14 PDF set.

<sup>17</sup>DYNNLO features QCD production at NNLO, no soft-gluon resummation, CT10 PDF set.



**Fig. 19:** Left: ratio of ResBos predictions of the normalised differential  $p_T^Z$  cross section to ATLAS Born-level data [93]. Right: the same ratio for different  $Z$  rapidity intervals and by using the DYNNLO programme for the theoretical prediction.

soft gluon resummation to data (right). While resummation is needed to describe the low- $p_T$  data, NLO calculations and better are required in the high- $p_T$  regime.

## 7 The SM is complete

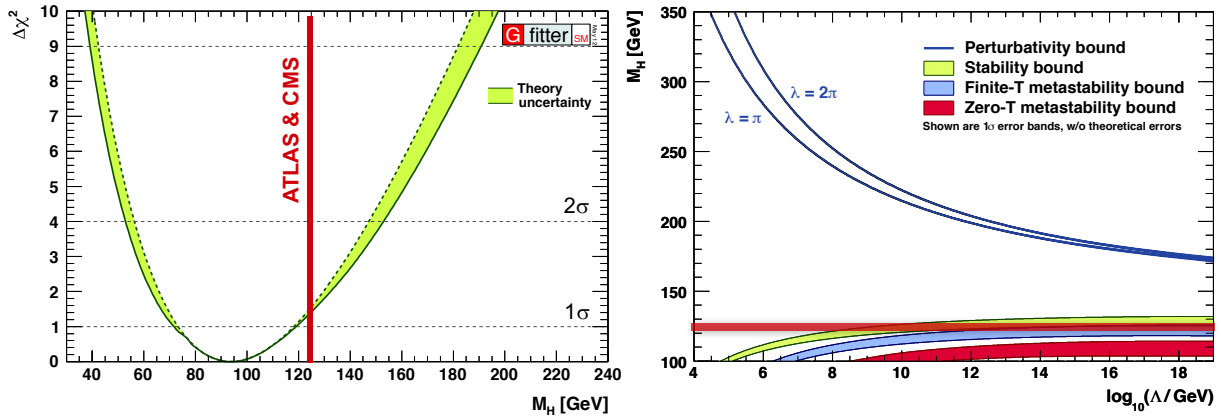
Since the LHC Run-1 the SM is a complete and self-consistent theory. The discovery of the Higgs boson is a triumph for the imagination and rigour of the scientific endeavour. It is also a triumph for the greatest experimental undertaking ever, at the frontier of accelerator and detector technologies, global data sharing, analysis and collaboration.

The Higgs mass of 125.1 GeV is in agreement with the prediction from the global electroweak fit [96] (cf. left panel of Fig. 20) and it lies marginally within the requirement for vacuum stability [97] (right panel of Fig. 20). The Higgs discovery does thus not come with a strict requirement for new physics below the Planck scale.

We have now two beautiful and extremely precise theories. On one hand the SM describing electroweak and strong interactions (though not their unification), predicting, eg., the anomalous magnetic moment of the electron to a relative precision of  $10^{-10}$  in agreement with experiment. On the other hand there is general relativity, the theory of gravitation. It has been tested to an accuracy of order  $10^{-5}$  (Cassini probe [98]). Unfortunately, the SM and general relativity do not work in regimes where both are important, that is at very small scales.

Indeed, many open questions not addressed by the SM remain as we have already alluded to in the introduction to these proceedings. We shall repeat some of them here.

- **Scalar sector.** Is there a single Higgs doublet or are there additional scalar states? Is the Higgs boson elementary or composite? What is the exact form of the scalar potential? What is the origin of the Yukawa couplings?
- **Quarks and leptons.** What is the origin of the fermion generations, mass, mixing, CP violation?



**Fig. 20:** Left:  $\chi^2$  curve obtained from the global electroweak fit in 2012 at the moment of the Higgs boson discovery [96]. Right: comparison of the observed Higgs mass with the lower limits from the vacuum stability constraint and the upper perturbativity limit [97].

How was the matter–antimatter asymmetry in the universe generated? What is the origin of baryon and lepton number conservation and what is the proton lifetime?

- **Neutrinos.** What is the nature of the neutrinos: Majorana or Dirac? Are there sterile neutrinos? What is the origin of neutrino mass and what are their values (and hierarchy)? Is there CP violation in the neutrino mixing?
- **Strong CP problem.** Why is there no noticeable CP violation in strong interactions albeit predicted by the SM?
- **Dark matter.** What is its composition: WIMPs, axions, sterile neutrinos, hidden sector particles, gravitational effect only? Is there a single or are there multiple sources?
- **Expansion of Universe.** Primordial expansion via inflation: which fields, and what is the role of the Higgs boson and of quantum gravity? Accelerated expansion today: cosmological constant problem.
- **High-scale physics.** Is there a solution to the hierarchy problem<sup>18</sup> and is there new physics at the TeV scale? Will there be grand unification of the forces? How does unification with gravity proceed? How is quantum gravity realised? Is everything just made of tiny strings?

Because the SM cannot be all there is, the LHC experiments have performed a large number of searches for new physics during Run-1, covering a vast space of possible signatures as witnessed in the exclusion plots of Fig. 21. Heavy resonances are excluded up to 3.5 TeV mass in some scenarios. Gluinos up to 1.3 TeV are excluded for light neutralinos (supersymmetry limits are usually lower than those of many other new physics scenarios because  $R$ -parity conservation requires pair production of supersymmetric particles).

<sup>18</sup>The term hierarchy problem stands for the apparent dependence of phenomena at the electroweak scale on a much higher (possibly the Planck) scale, as exemplified by the extreme ultra-violet sensitivity of the Higgs potential.

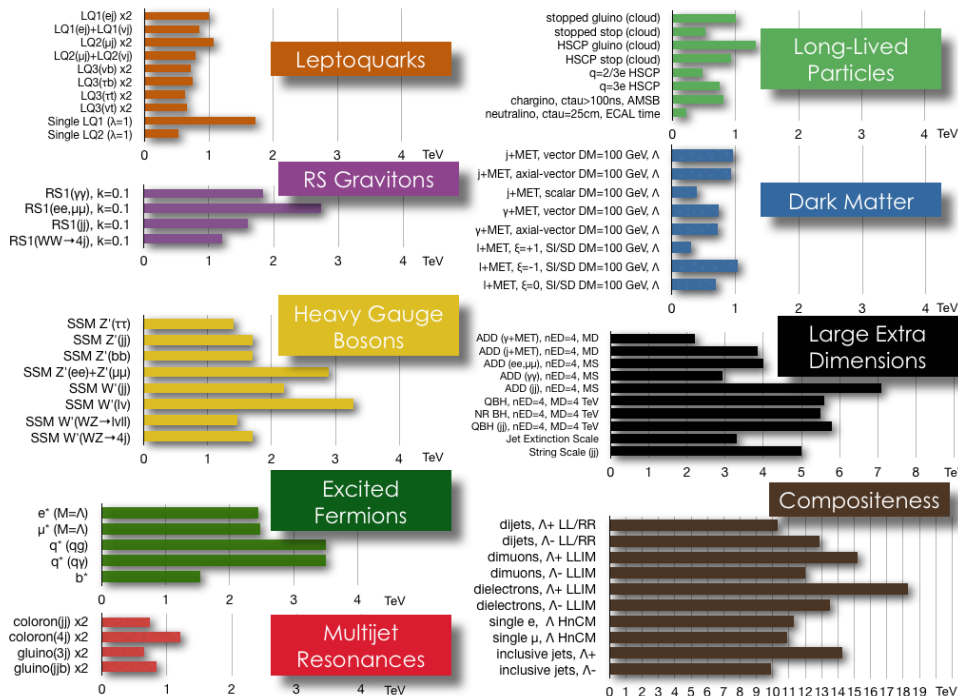
**ATLAS SUSY Searches\* - 95% CL Lower Limits**  
 Status: July 2015

ATLAS Preliminary  
 $\sqrt{s} = 7, 8 \text{ TeV}$

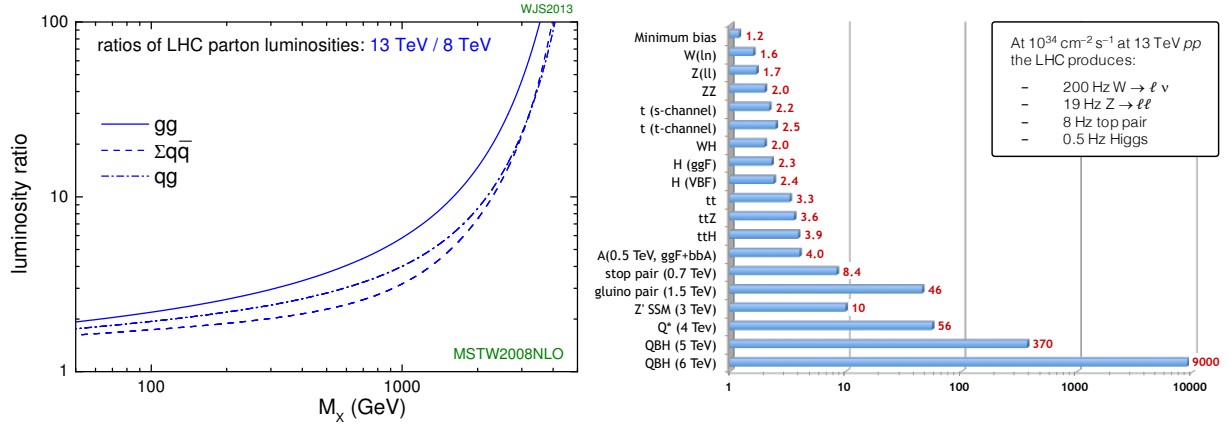
| Model  | $e, \mu, \tau, \gamma$  | Jets                  | $E_{miss}^T$   | $[L d] [\text{fb}^{-1}]$ | Mass limit                 |                                  | Reference  |                      |
|--|---|-----------------------|----------------|--------------------------|----------------------------|----------------------------------|--|----------------------|
|  |   |                       |                |                          | $\sqrt{s} = 7 \text{ TeV}$ | $\sqrt{s} = 8 \text{ TeV}$       |  |                      |
| Inclusive Searches                                 | MSUGRA/CMSSM  | 0-3 $e, \mu, \tau$    | 2-10 jets/3 b  | Yes                      | 20.3                       | $\tilde{g}, \tilde{q}$           | 1.8 TeV $m(\tilde{g})=m(\tilde{q})$  | 1507.05525           |
|  | $\tilde{g}\tilde{g}, \tilde{g}\tilde{q}$  | 0                     | 2-6 jets       | Yes                      | 20.3                       | $\tilde{g}$                      | 850 GeV $m(\tilde{g})=0 \text{ GeV}, m(\tilde{q})=m(\tilde{g})$                                    | 1405.7875            |
|  | $\tilde{g}\tilde{g}, \tilde{g}\tilde{q}$ (compressed)   | mono-jet              | 1-3 jets       | Yes                      | 20.3                       | $\tilde{g}$                      | 100-440 GeV $m(\tilde{g})=m(\tilde{q}) < 10 \text{ GeV}$   | 1507.05525           |
|  | $\tilde{g}\tilde{g}, \tilde{g}\tilde{q}$ ( $\tilde{g}\tilde{g}/\tilde{g}\tilde{q}$ )                                      | 2 $e, \mu$ (off-Z)    | 2 jets         | Yes                      | 20.3                       | $\tilde{g}$                      | 780 GeV $m(\tilde{g})=0 \text{ GeV}$   | 1503.03290           |
|  | $\tilde{g}\tilde{g}, \tilde{g}\tilde{q}$ ( $\tilde{g}\tilde{g}/\tilde{g}\tilde{q}$ )                                      | 0                     | 2-6 jets       | Yes                      | 20.3                       | $\tilde{g}$                      | 1.33 TeV $m(\tilde{g})=0 \text{ GeV}$  | 1405.7875            |
|  | $\tilde{g}\tilde{g}, \tilde{g}\tilde{q}$ ( $\tilde{g}\tilde{g}/\tilde{g}\tilde{q}$ )                                      | 0-1 $e, \mu$          | 2-6 jets       | Yes                      | 20.3                       | $\tilde{g}$                      | 1.26 TeV $m(\tilde{g})=300 \text{ GeV}, m(\tilde{q})=0.5(m(\tilde{g}))+m(\tilde{g})$               | 1507.05525           |
|  | $\tilde{g}\tilde{g}, \tilde{g}\tilde{q}$ ( $\tilde{g}\tilde{g}/\tilde{g}\tilde{q}$ )                                      | 2 $e, \mu$            | 0-3 jets       | Yes                      | 20.3                       | $\tilde{g}$                      | 1.32 TeV $m(\tilde{g})=0 \text{ GeV}$  | 1501.03555           |
|  | GMSB ( $\tilde{g}$ NLSP)  | 1-2 $\tau + 0-1 \ell$ | 0-2 jets       | Yes                      | 20.3                       | $\tilde{g}$                      | 1.6 TeV $\text{tan}\beta > 20$   | 1407.0603            |
|  | GGM (higgs NLSP)  | $\gamma$              | 1 b            | Yes                      | 20.3                       | $\tilde{g}$                      | 1.29 TeV $\tau(\text{NLSP}) < 0.1 \text{ mm}$  | 1507.05493           |
|  | GGM (higgsino-bino NLSP)  | $\gamma$              | 2 jets         | Yes                      | 20.3                       | $\tilde{g}$                      | 1.3 TeV $m(\tilde{g})=900 \text{ GeV}, \tau(\text{NLSP}) < 0.1 \text{ mm}, \mu < 0$                | 1507.05493           |
|  | GGM (higgsino NLSP)   | 2 $e, \mu$ (Z)        | 2 jets         | Yes                      | 20.3                       | $\tilde{g}$                      | 1.25 TeV $m(\tilde{g})=850 \text{ GeV}, \tau(\text{NLSP}) < 0.1 \text{ mm}, \mu > 0$               | 1503.03290           |
|  | Gravitino LSP   | 0                     | mono-jet       | Yes                      | 20.3                       | $\tilde{g}^{\text{Higgs}}$ scale | 865 GeV $m(\tilde{g})=1.8 \times 10^{-8} \text{ eV}, m(\tilde{g})=m(\tilde{g})=1.5 \text{ TeV}$    | 1502.01518           |
| $\tilde{g}\tilde{g}$ ann. $\tilde{g}$ infed.       | $\tilde{g}\tilde{g}, \tilde{g}\tilde{b}$  | 0                     | 3 b            | Yes                      | 20.3                       | $\tilde{g}$                      | 1.25 TeV $m(\tilde{g}) < 400 \text{ GeV}$  | 1407.0600            |
|  | $\tilde{g}\tilde{g}, \tilde{g}\tilde{t}$  | 0                     | 7-10 jets      | Yes                      | 20.3                       | $\tilde{g}$                      | 1.1 TeV $m(\tilde{g}) < 350 \text{ GeV}$   | 1308.1841            |
|  | $\tilde{g}\tilde{g}, \tilde{g}\tilde{q}$  | 0-1 $e, \mu$          | 3 b            | Yes                      | 20.3                       | $\tilde{g}$                      | 1.34 TeV $m(\tilde{g}) < 400 \text{ GeV}$  | 1407.0600            |
| $\tilde{g}\tilde{g}$ ann. $\tilde{g}$ direct prod. | $\tilde{g}\tilde{g}, \tilde{g}\tilde{b}$  | 0                     | 2 b            | Yes                      | 20.1                       | $\tilde{g}$                      | 100-620 GeV $m(\tilde{g})=90 \text{ GeV}$  | 1308.2831            |
|  | $\tilde{g}\tilde{g}, \tilde{g}\tilde{t}$  | 2 $e, \mu$ (SS)       | 0-3 b          | Yes                      | 20.3                       | $\tilde{g}$                      | 275-440 GeV $m(\tilde{g}) < 2 m(\tilde{t})$  | 1404.2500            |
|  | $\tilde{g}\tilde{g}, \tilde{g}\tilde{q}$  | 1-2 $e, \mu$          | 1-2 b          | Yes                      | 4.7/20.3                   | $\tilde{g}$                      | 310-167 GeV $m(\tilde{g}) = 2m(\tilde{t}), m(\tilde{t})=55 \text{ GeV}$                            | 1209.2102, 1407.0583 |
| $\tilde{g}\tilde{g}$ ann. $\tilde{g}$ natural GMSB | $\tilde{g}\tilde{g}, \tilde{g}\tilde{t}$  | 0-2 $e, \mu$          | 0-2 jets/1-2 b | Yes                      | 20.3                       | $\tilde{g}$                      | 90-191 GeV $m(\tilde{g})=1 \text{ GeV}$  | 1506.0816            |
|  | $\tilde{g}\tilde{g}, \tilde{g}\tilde{q}$  | 0                     | mono-jet/c-tag | Yes                      | 20.3                       | $\tilde{g}$                      | 90-240 GeV $m(\tilde{g})=m(\tilde{q})=85 \text{ GeV}$  | 1407.0608            |
|  | $\tilde{g}\tilde{g}, \tilde{g}\tilde{q}$  | 2 $e, \mu$ (Z)        | 1 b            | Yes                      | 20.3                       | $\tilde{g}$                      | 150-530 GeV $m(\tilde{g})=150 \text{ GeV}$   | 1403.5222            |
| EW direct  | $\tilde{g}\tilde{g}, \tilde{g}\tilde{t}$  | 2 $e, \mu$            | 0              | Yes                      | 20.3                       | $\tilde{g}$                      | 90-325 GeV $m(\tilde{g})=0 \text{ GeV}$  | 1403.5294            |
|  | $\tilde{g}\tilde{g}, \tilde{g}\tilde{q}$  | 2 $e, \mu$            | 0              | Yes                      | 20.3                       | $\tilde{g}$                      | 140-465 GeV $m(\tilde{g})=0 \text{ GeV}, m(\tilde{t}, \tilde{b})=0.5(m(\tilde{t}))+m(\tilde{t})$   | 1403.5294            |
|  | $\tilde{g}\tilde{g}, \tilde{g}\tilde{q}$  | 2 $\tau$              | -              | Yes                      | 20.3                       | $\tilde{g}$                      | 100-350 GeV $m(\tilde{g})=0 \text{ GeV}, m(\tilde{t}, \tilde{b})=0.5(m(\tilde{t}))+m(\tilde{t})$   | 1407.0350            |
| Long-lived particles                               | Stable, stopped $\tilde{g}$ R-hadron  | 0                     | 1-5 jets       | Yes                      | 19.9                       | $\tilde{g}$                      | 832 GeV $m(\tilde{g})=m(\tilde{t}), m(\tilde{t})=160 \text{ MeV}, \tau(\tilde{g})=0.2 \text{ ns}$  | 1310.3675            |
|  | Stable $\tilde{g}$ R-hadron   | trk                   | -              | Yes                      | 27.1                       | $\tilde{g}$                      | 482 GeV $m(\tilde{g})=m(\tilde{t}), m(\tilde{t})=160 \text{ MeV}, \tau(\tilde{g}) < 15 \text{ ns}$ | 1506.05332           |
|  | GMSB, stable $\tilde{g}, \tilde{q}$   | 1-2 $\mu$             | -              | Yes                      | 19.1                       | $\tilde{g}$                      | 537 GeV $m(\tilde{g})=100 \text{ GeV}, 10^{-8} \mu\text{s} < \tau(\tilde{g}) < 1000 \text{ s}$     | 1310.6584            |
| RPV  | GMSB, $\tilde{g}\tilde{g} \rightarrow \tilde{g}\tilde{g}$ , long-lived $\tilde{g}$  | 4 $e, \mu$            | -              | Yes                      | 20.3                       | $\tilde{g}$                      | 435 GeV $10^{-10} \text{ tan}\beta < 50$   | 1411.6795            |
|  | GGM $\tilde{g}\tilde{g}, \tilde{g}\tilde{q}$  | displ. ee/μμ/μμ       | -              | Yes                      | 20.3                       | $\tilde{g}$                      | 1.0 TeV $2 < \tau(\tilde{g}) < 3 \text{ ns}, \text{SPSB model}$                                    | 1504.5542            |
|  | GGM $\tilde{g}\tilde{g}, \tilde{g}\tilde{q}$  | displ. vtx + jets     | -              | Yes                      | 20.3                       | $\tilde{g}$                      | 1.0 TeV $7 < \tau(\tilde{g}) < 740 \text{ ms}, m(\tilde{g})=1.3 \text{ TeV}$                       | 1504.05162           |
| Other  | LFV $\tilde{g}\tilde{g} \rightarrow \tilde{g}\tilde{g} + X, \tilde{g}\tilde{g} \rightarrow \tilde{g}\tilde{g} + \tau\tau$ | 0                     | -              | Yes                      | 20.3                       | $\tilde{g}$                      | 1.7 TeV $A_{12} = -0.11, A_{12313131} = 0.07$  | 1503.04430           |
|  | Bilinear RPV CMSSM  | 2 $e, \mu$ (SS)       | 0-3 b          | Yes                      | 20.3                       | $\tilde{g}, \tilde{q}$           | 1.35 TeV $m(\tilde{g})=m(\tilde{q}), \tau_{\tilde{g}} < 1 \text{ mm}$                              | 1404.2500            |
|  | GMSB, $\tilde{g}\tilde{g} \rightarrow \tilde{g}\tilde{g}$ , long-lived $\tilde{g}$  | 4 $e, \mu$            | -              | Yes                      | 20.3                       | $\tilde{g}$                      | 750 GeV $m(\tilde{g})=0.2 m(\tilde{t}), A_{1212} = 0$  | 1405.5086            |

\*Only a selection of the available mass limits on new states or phenomena is shown. All limits quoted are observed minus 1 $\sigma$  theoretical signal cross section uncertainty.

**CMS Preliminary** Status: Moriond 2015



**Fig. 21:** Exclusion bounds on mass scales from searches for supersymmetry (top, ATLAS) and other new physics phenomena (bottom, CMS Exotica).



**Fig. 22:** Left: parton luminosity ratio of 13 TeV to 8 TeV proton–proton collisions [11]. Right: cross section ratios for selected processes (the heavy flavour cross section scales roughly linearly with centre-of-mass energy).

## 8 The LHC Run-2

A huge milestone was achieved in 2015 when the new record proton–proton collision energy of 13 TeV was reached. After a rocky start, the LHC delivered  $4.2 \text{ fb}^{-1}$  integrated luminosity to ATLAS and CMS. That amount of data already surpassed the Run-1 new physics sensitivity of many searches. During 2016 a peak luminosity of  $1.4 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  was reached and a total of  $39 \text{ fb}^{-1}$  integrated luminosity delivered, which exceeded expectations.

The new centre-of-mass energy increases the cross-section of all LHC processes. Figure 22 gives the 13 TeV to 8 TeV parton luminosity ratios for gluon–gluon, quark–gluon and quark–quark scattering (left panel) and the resulting proton–proton cross-section ratios (right). The parton luminosity as a function of the hard scattering  $Q^2 = M_X^2$  (cf. Fig. 3) is defined by the convolution integral

$$\frac{\partial \mathcal{L}_{ab}}{\partial M_X^2} = \frac{1}{s} \int_{\tau}^1 \frac{dx}{x} f_a(x, M_X^2) f_b(\tau/x, M_X^2), \quad (8)$$

where  $\tau = M_X^2/s$ . There is a larger parton luminosity increase with energy for gluon initiated processes than for quark ones. Owing to the important cross section rise at large  $M_X$  the early Run-2 analyses put their emphasis on searches.

Most of the results from ATLAS and CMS presented at the 2016 summer conferences contained data up to approximately  $15 \text{ fb}^{-1}$ . CMS used different software releases and thus did not merge the 2015 and 2016 datasets, but in selected cases provided a statistical combination. ATLAS performed a reprocessing of the 2015 data and MC allowing it to treat both years as a single coherent dataset. LHCb performed luminosity levelling leading to an approximately ten times smaller dataset in terms of integrated luminosity. The uncertainty on the luminosity values from ATLAS, CMS and LHCb, for the summer 2016 results were 2.9%, 6.2% and 3.8%, respectively. The amount of pileup interactions with an average  $\mu$  (cf. Eq. 3) of 23 interactions was similar to that in 2012. LHCb observed 1.7 pileup interactions in average.

### 8.1 Standard Model and top-quark physics

Along increasing scattering momentum transfer, SM processes the LHC can be split the as follows.

- **Soft QCD:** study of particle spectra. The transverse momenta are typically smaller than a few GeV. More than 99.999% of the proton–proton collisions belong to that type. Measurements of

soft QCD processes serve to probe LO matrix elements, parton shower models, generator tunings, and for pileup modelling.

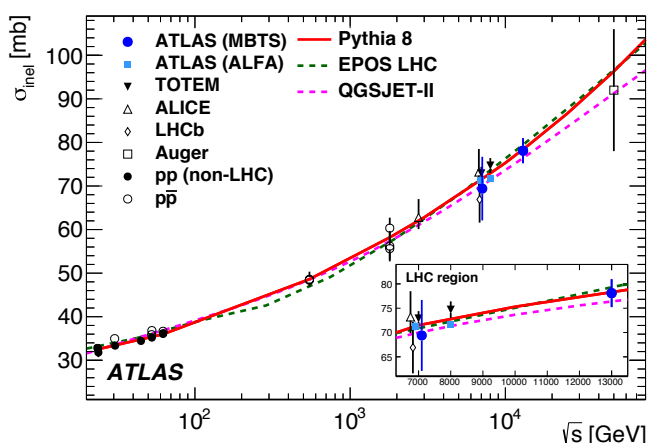
- **Hard QCD:** study of jets. Typical jet  $p_T$  greater than tens of GeV up to the TeV scale; approximately  $10^{-5}$  of the collisions belong to that category. The measurements probe NLO QCD, the running  $\alpha_s$ , PDFs, parton showers, etc.
- **Hard QCD and electroweak processes:**  $W$ ,  $Z$ ,  $H$ , top decaying to stable identified particles. Typical  $p_T$  scale of greater than tens of GeV ; a fraction of  $10^{-6}$  and less of the collisions belong to this category. Measurements probe NLO, NN(N)LO QCD, soft gluon resummation, PDFs, electroweak physics, etc.

Standard Model and Higgs precision measurements are key to the LHC programme up to the High-Luminosity LHC (HL-LHC). Michelangelo Mangano at the SEARCH 2016 workshop [99] summarised the importance of these measurements as follows.

- **Scientific perspective.** No matter what BSM the LHC will unveil in the next years, improving the knowledge of Higgs properties is a must, which by itself requires and justifies the largest possible LHC statistics so that stopping after  $300 \text{ fb}^{-1}$  would not be satisfying.
- **Pragmatic perspective.** Higgs and SM physics are the only guaranteed deliverables of the LHC programme. Need to exploit this part of the programme to its maximum extent!
- **Utilitarian perspective.** Elements of the SM, besides the Higgs, require further consolidation, control and improved precision, both in the EW and QCD sectors. They hold a fundamental value (eg. the precise determination of parameters of nature and to better understand detailed scattering dynamics), or are critical to fully exploit the BSM search potential (eg. the knowledge of backgrounds, production rates and production dynamics).
- **Spinoffs.** The study of SM processes at colliders is typically more complex than the search for BSM signatures and throughout the years it has been the main driver of fundamental theoretical innovation.

### 8.1.1 Inelastic proton–proton cross section

A key initial measurement is the inclusive inelastic cross-section at 13 TeV. While the most precise total cross section measurement is obtained via elastic scattering and the optical theorem ( $\sigma_{\text{tot}}(pp \rightarrow X) \propto \text{Im}f_{\text{elastic}}(t \rightarrow 0)$ , where  $f_{\text{elastic}}(t \rightarrow 0)$  is the elastic scattering amplitude extrapolated to the forward direction, and  $t$  is the Mandelstam momentum transfer variable) using dedicated forward devices (such measurements have achieved better than 1% precision in Run-1, dominated by the luminosity uncertainty [100,101]); it is also possible to determine  $\sigma_{\text{tot}}(pp \rightarrow X)$  from a measurement of inelastic scattering cross section if the extrapolation between fiducial to total acceptance is not too large. This can be achieved using forward detectors such as scintillators installed in ATLAS within  $2.07 < |\eta| < 3.86$ .



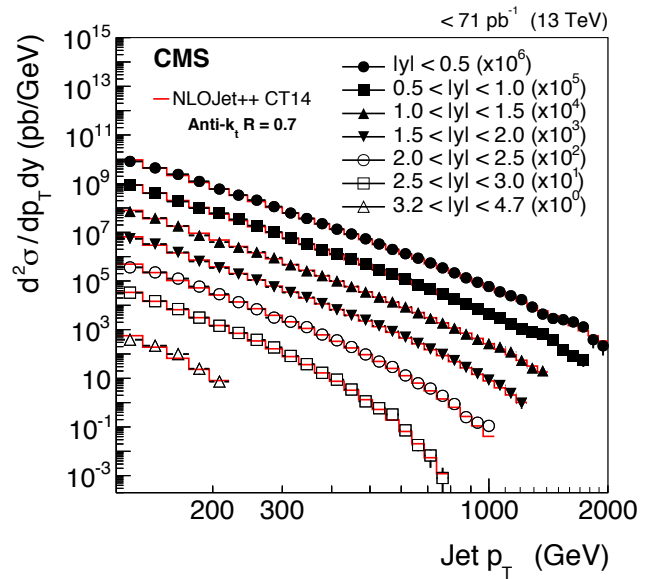
**Fig. 23:** Inelastic proton–proton cross section versus centre-of-mass energy [102].



The ATLAS measurement [102] was performed in the fiducial region  $\xi = M_X^2/s > 10^{-6}$ , where  $M_X$  is the larger invariant mass of the two hadronic (proton-dissociation) systems separated by the largest rapidity gap in the event. In this  $\xi$  range the scintillators have high efficiency. When extrapolated to the full phase space, a cross-section of  $\sigma_{\text{tot}}(pp \rightarrow X) = 78.1 \pm 0.6 \pm 1.3 \pm 2.6$  mb is obtained, where the first uncertainty is experimental, the second due to the luminosity, and the third and dominant one from the extrapolation to full phase space. The result is consistent with the expectation from phenomenological models (cf. Fig. 23, where also measurements from other hadron collider experiments and from the Pierre Auger experiment are shown, see references in [102]).

### 8.1.2 Jet production

Moving up in transverse momentum, ATLAS and CMS measured jet production. Figure 24 shows the double differential inclusive jet cross section as measured by CMS. The unfolded data points are compared to predictions from NLOJet++ based on the CT14 PDF set and corrected for the nonperturbative and electroweak effects (line in figure). It is interesting to compare Fig. 24 at 13 TeV to the 8 TeV result shown in Fig. 7 on page 15. For a given rapidity interval, the relative drop in cross section between high and low  $p_T$  is less pronounced at 13 TeV, as expected from the parton luminosities. Indeed, taking the ratio between the 13 TeV and 8 TeV cross sections approximately reproduces the left panel of Fig. 22 for gluon–gluon scattering.



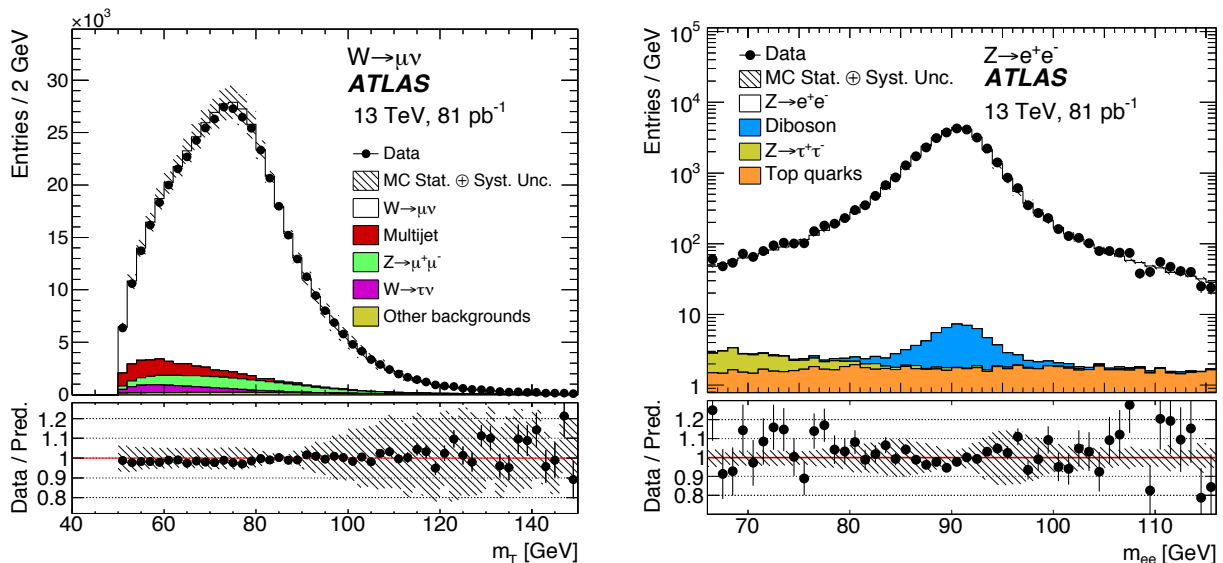
**Fig. 24:** Double-differential inclusive jet cross section versus jet  $p_T$  as measured by CMS at 13 TeV [103] and compared to a theoretical prediction.

### 8.1.3 Weak boson production

The inclusive  $W$  and  $Z$  boson production cross sections are expected to rise at 13 TeV over 8 TeV centre-of-mass energy by factors of 1.7 and 1.6, respectively, to 19.7 nb and 1.9 nb for the decays to muons. Leptonic  $W$  and  $Z$  decays are very pure channels as can be seen from Fig. 25, which shows the transverse and invariant dilepton mass distributions for 13 TeV  $W \rightarrow \mu\nu$  (left panel) and  $Z \rightarrow ee$  candidates (right). The transverse mass-squared is defined by  $m_T^2 = 2p_{T,\ell}E_T^{\text{miss}}(1 - \cos\Delta\phi_{\ell,\nu})$ , where  $\Delta\phi_{\ell,\nu}$  is the azimuthal angle difference between lepton and missing transverse momentum. The dilepton invariant mass-squared is given by  $m_{\ell_1\ell_2}^2 = 2p_{T,\ell_1}p_{T,\ell_2}(\cosh\Delta\eta_{12} - \cos\Delta\phi_{12})$ .

Apart from the intrinsic interest in precise cross section measurements, leptonic  $W$  and  $Z$  decays also serve the experiments as standard candles to calibrate the electron and muon reconstruction performance via mass constraints and so-called tag-and-probe efficiency measurements. Tag-and-probe methods [104] are used to select, from known resonances, unbiased samples of electrons or muons (probes) by using strict selection requirements on the second object produced from the particle's decay (tags). The efficiency of a requirement can then be determined by applying it directly to the probe sample after accounting for residual background contamination.

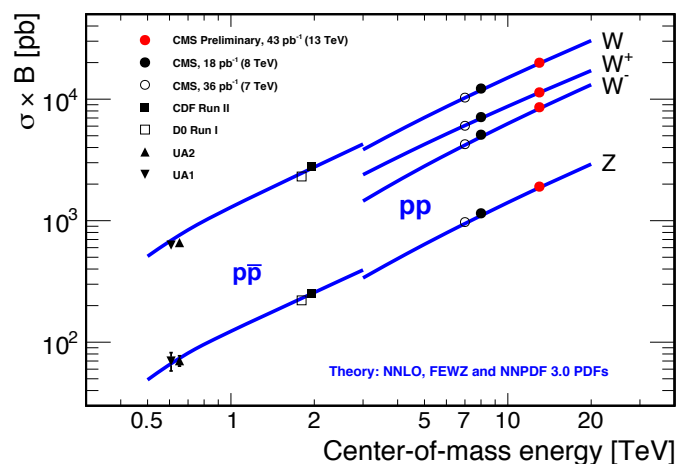
Both ATLAS and CMS measured fiducial and inclusive cross sections for  $W$  and  $Z$  boson production as well as their ratios using partial 2015 datasets [105, 106]. The fiducial cross sections are dominated by



**Fig. 25:** Transverse mass (left) and invariant dilepton mass (right) for  $W \rightarrow \mu\nu$  and  $Z \rightarrow ee$  candidates, respectively [105, 106]. The predicted signal distributions are normalised to the measured cross sections.

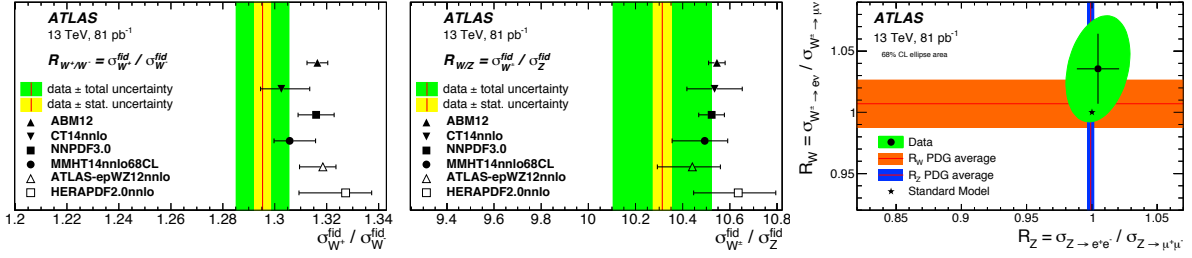
the luminosity uncertainty of 2.1% (ATLAS). Comparisons of the measured cross-sections with NNLO QCD and NLO EW Drell-Yan predictions show good agreement within uncertainties. Figure 26 shows the energy dependence of the measured inclusive  $W$  and  $Z$  boson cross sections compared to theoretical predictions. LHCb has measured the 13 TeV  $Z$  boson cross section in the fiducial acceptance  $2.0 < \eta < 4.5$  and found agreement with the SM prediction [107].

Ratios of cross sections already achieve precision of better than 1–2% owing to a cancellation of systematic uncertainties. They represent powerful tools to constrain PDFs: the  $W^+/W$  ratio is sensitive to the low- $x$   $u$  and  $d$  valence quarks, and the  $W^\pm/Z$  ratio constrains the strange quark PDF, in particular when also using the rapidity distributions. Figure 27 shows the measured and predicted ratios. Fair agreement between the data and most PDF sets is seen. An increased strange quark contribution [108] (towards SU(3) flavour symmetry of sea squarks in the proton) would likely improve the agreement. The right panel in Fig. 27 shows tests of the universality of the first and second generation leptonic couplings to the weak bosons. Lepton universality in the charged current was measured to the 0.14% level at LEP in  $\tau$  lepton decays, however at low energy (off-shell), so with less sensitivity to new physics in loops.

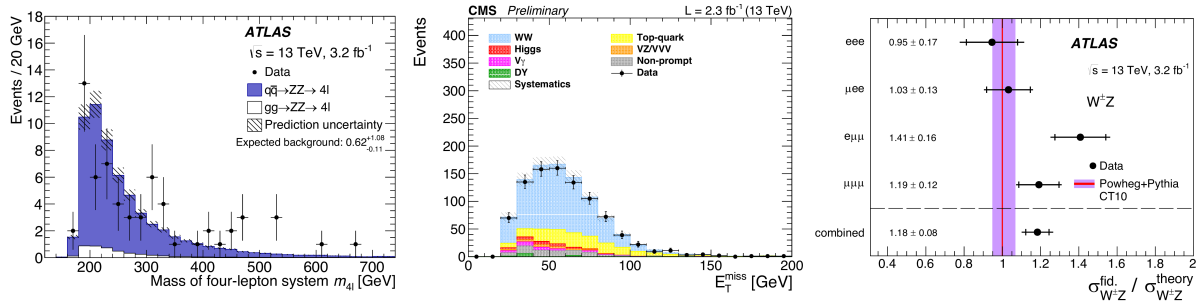


**Fig. 26:** Cross sections of proton–(anti)-proton production of inclusive  $W$  and  $Z$  bosons versus centre-of-mass energy.





**Fig. 27:** Ratios of fiducial cross sections compared to various PDF predictions (left and middle panels), and  $W$  and  $Z$  lepton universality tests (right panel) [111].

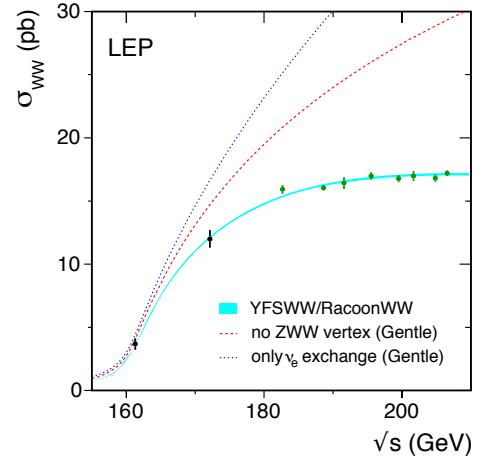


**Fig. 28:** Distributions from 13 TeV diboson selections. Left: four-lepton invariant mass in the  $ZZ \rightarrow 4\ell$  analysis [109]; middle: missing transverse momentum in  $WW \rightarrow 2\ell 2\nu$  [110]; right: ratio of measured over predicted (NLO QCD) fiducial cross sections measured in  $WZ \rightarrow 3\ell\nu$  [111].

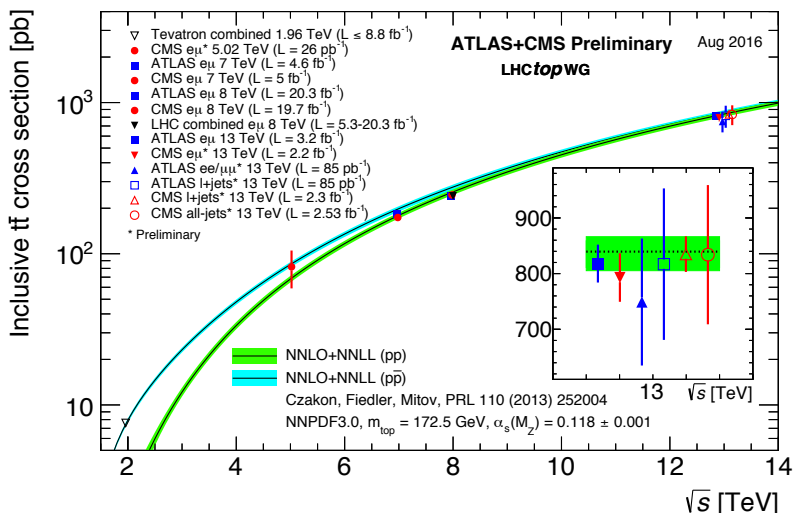
### 8.1.4 Diboson production

The production of boson pairs is a highly important sector of LHC physics that is intimately related to electroweak symmetry breaking. In the  $s$ -channel (Drell-Yan), via photon,  $Z$  or  $W$  exchange, diboson production is sensitive to anomalous triple gauge boson couplings (aTGC). Triple and also quartic gauge boson couplings, the latter vertices involving the scattering among four gauge bosons, are predicted by the SM as the electroweak gauge bosons carry weak charge (non-Abelian structure of EW theory).

The production of  $WW$  and  $ZZ$  events was studied at LEP versus the  $e^+e^-$  centre-of-mass energy resulting in a famous plot that showed the moderation of the  $WW$  cross section versus energy by TGC processes as predicted by the SM (see Fig. 29) [112]. The Tevatron experiments studied a multitude of diboson production processes. ATLAS and CMS performed inclusive, fiducial and differential cross-section analyses at 8 TeV. First fiducial and total cross section measurements at 13 TeV are also available (see Fig. 28 for a selection of representative plots). Inclusive diboson  $WW$ ,  $WZ$  and  $ZZ$  events are reconstructed through the leptonic decays of the weak bosons, leading to two-lepton, three-lepton and four-lepton final states, where the former two channels are accompanied by  $E_T^{miss}$ . Hadronic weak boson decays are not competitive for inclusive cross section measurements, but are interesting for aTGC searches at high diboson mass where possible new physics effects are expected to show up first



**Fig. 29:**  $W$  pair production cross section measured at LEP compared to the SM predictions [112].



**Fig. 30:** Summary of LHC and Tevatron measurements of the inclusive top-pair production cross section versus centre-of-mass energy. The bands correspond to predictions with uncertainties from NNLO QCD calculations including NNLL soft gluon resummation. Measurements and theory calculations assume  $m_t = 172.5$  GeV.

(see [113] for a recent review). These results show that NNLO QCD is needed to match the data, and, in case of  $WW \rightarrow 2\ell 2\nu$  measured in an exclusive zero-jet channel, higher-logarithmic-order (NNLL) soft gluon resummation.

### 8.1.5 Top production

Top–antitop production at the LHC is dominated by gluon–gluon scattering in the initial state. A factor of 3.3 cross-section increase at 13 TeV centre-of-mass energy compared to 8 TeV is expected. The inclusive  $pp \rightarrow t\bar{t} + X$  cross section can be robustly measured using dilepton events selecting different lepton flavours to suppress Drell-Yan background. A method applied successfully during Run-1 allows to simultaneously determine the  $t\bar{t}$  cross section and  $b$ -tagging efficiency from data [114]. We shall briefly discuss it here for the corresponding 13 TeV measurement as it is an instructive example for a straightforward experimental approach relying where possible on data.

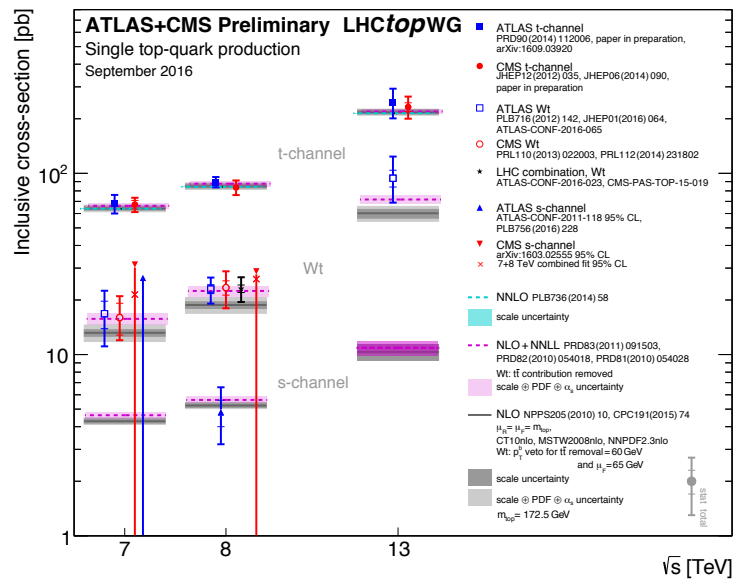
The method employs an exclusive selection of  $e\mu$  events with one and two  $b$ -tags. The observed number of events is given by  $N_1 = L \cdot \sigma_{t\bar{t}} \cdot \epsilon_{e\mu} \cdot 2\epsilon_b \cdot (1 - C_b \cdot \epsilon_b) + N_1^{\text{bkg}}$  and  $N_2 = L \cdot \sigma_{t\bar{t}} \cdot \epsilon_{e\mu} \cdot C_b \cdot \epsilon_b^2 + N_2^{\text{bkg}}$ , where  $N_{1(2)}$  is the number of observed events with one (two)  $b$ -tags,  $L$  the integrated luminosity of the analysed data sample,  $\epsilon_{e\mu}$  the combined  $t\bar{t} \rightarrow e\mu + X$  selection acceptance and efficiency determined from MC,  $\epsilon_b$  the probability to  $b$ -tag  $q$  from  $t \rightarrow Wq$  determined from data ( $\epsilon_b$  includes the selection acceptance and efficiency), and  $C_b = \epsilon_{bb}/\epsilon_b^2$  is a small non-factorisation correction ( $1.002 \pm 0.006$ ) determined from MC. The selection of  $t\bar{t} \rightarrow e\mu + X$  events is very pure with, for the ATLAS 2015 dataset ( $3.2 \text{ fb}^{-1}$ ),  $N_1 = 11958$ ,  $N_2 = 7069$ , and  $N_1^{\text{bkg}} = 1370 \pm 120$ ,  $N_2^{\text{bkg}} = 340 \pm 88$  event counts [115]. The background is dominated by the  $Wt$  single-top process. Solving the equations simultaneously for  $\sigma_{t\bar{t}}$  and  $\epsilon_b$  gives:  $\sigma_{t\bar{t}} = 828 \pm 8 \pm 27 \pm 19 \pm 12 \text{ pb}$ , where the first error is statistical, the second systematic, and the third (fourth) due to the luminosity (beam-energy) uncertainty.<sup>19</sup> The total relative uncertainty of 4.4% is already

<sup>19</sup>The LHC beam energy during the 2012 proton–proton run was calibrated to be  $0.30 \pm 0.66 \%$  below the nominal value of 4 TeV per beam. That estimate, dominated by systematic uncertainties, was made by measuring the revolution frequency, that is, the speed difference of protons and lead ions during proton–lead runs in early 2013 [116], taking advantage of the simultaneous presence of both particle types with the same orbits in the LHC. The measurement result agrees with the beam energy derived from the magnetic calibration curves of the dipole magnets that are used to generate the current settings of the

comparable with the 4.3% obtained at 8 TeV. The measurement is in agreement with the theoretical prediction of  $832^{+40}_{-46}$  pb, based on NNLO QCD including NNLL soft gluon resummation and of similar precision as the measurement. The systematic uncertainty affecting the measurement (total 3.3%) is dominated by theoretical sources in particular the modelling of nonperturbative effects related to parton showering and hadronisation. It is interesting (though not mandatory for the method to work) to observe that the resulting value for  $\epsilon_b = 0.559 \pm 0.004 \pm 0.003$  is in agreement with the value of 0.549 found in MC simulation. Figure 30 shows a summary of various LHC and Tevatron  $t\bar{t}$  cross section measurements versus centre-of-mass energy, and compared to theoretical predictions. The  $e\mu$  method provides the most precise inclusive results at all LHC centre-of-mass energies.

The experiments also performed first differential cross section measurements at 13 TeV which show reasonable modelling although deviations at large jet multiplicity and top  $p_T$  persist, similar to those seen in Run-1.

Electroweak single-top production (cf. Feynman graphs in Fig. 11) amounts, in case for the dominant t-channel, to about one third of the  $t\bar{t}$  production cross section and is expected to increase by a factor of 2.5 at 13 TeV compared to 8 TeV. A summary of the inclusive cross section measurements for all available centre-of-mass energies is displayed in Fig. 31 (see [117–119] for the 13 TeV results). Agreement with theoretical predictions based on pure NLO QCD, NLO QCD complemented with NNLL resummation, and NNLO QCD (t-channel only) is observed. In the t-channel, also the charge asymmetries are measured at 13 TeV and found in agreement with the SM prediction (the ratio of  $tq$  to  $\bar{t}q$  production is measured to be  $1.72 \pm 0.20$ ). The s-channel is challenging at the LHC and requires more data.



**Fig. 31:** Summary of ATLAS and CMS measurements of single-top production cross sections versus centre-of-mass energy. The measurements are compared to theoretical calculations (see text for details).

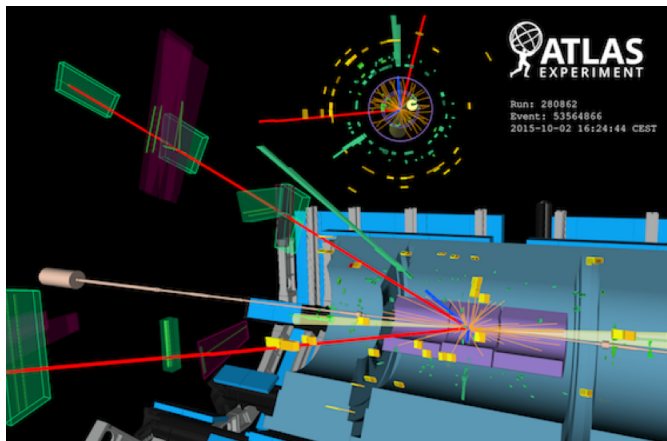
The opening of the phase space at 13 TeV allows to produce heavier final states, such as the associated production of  $t\bar{t}$  with a  $W$  or  $Z$  boson (cf. Feynman graphs in Fig. 10 on page 16). Because of different production mechanisms (dominantly gluon s-channel scattering in case of  $ttZ$  and t-channel quark–antiquark annihilation for  $ttW$ ) the 13 TeV to 8 TeV cross-section ratios are different for the two channels: 3.6 for  $ttZ$  compared to only 2.4 for  $ttW$ . Both experiments have produced first inclusive 13 TeV cross section results [120, 121] finding for  $ttW$ :  $1.5 \pm 0.8$  pb (ATLAS) and  $0.98^{+0.32}_{-0.28}$  pb (CMS), and for  $ttZ$ :  $0.9 \pm 0.3$  pb (ATLAS) and  $0.70^{+0.21}_{-0.19}$  pb (CMS). The corresponding SM predictions are  $0.60 \pm 0.08$  pb ( $ttW$ ) and  $0.84 \pm 0.09$  pb ( $ttZ$ ). Both, ATLAS and CMS are slightly on the high side for  $ttW$  reproducing a similar pattern already observed in the Run-1 data. Improving these results is important in its own rights, but also because the  $ttW/Z$  channels are important backgrounds to  $ttH$  in final states with multiple leptons, where in particular  $ttW$  is difficult to separate.

power converters which feed the magnets during beam operation. The magnetic calibration is expected to be accurate within an uncertainty of about 0.07%, which is significantly better than the measurement uncertainty based on proton–lead data.

## 8.2 Reobservation of the Higgs boson at 13 TeV

The expected 13 TeV to 8 TeV cross section ratios amount to  $2\sim 2.4$  for  $VH$ ,  $ggH$ , VBF, and 3.9 for  $ttH$  production. The combination of the 2015 and 2016 data available by the 2016 summer conferences should therefore already achieve similar or better significance and precision on Higgs boson production than in Run-1.

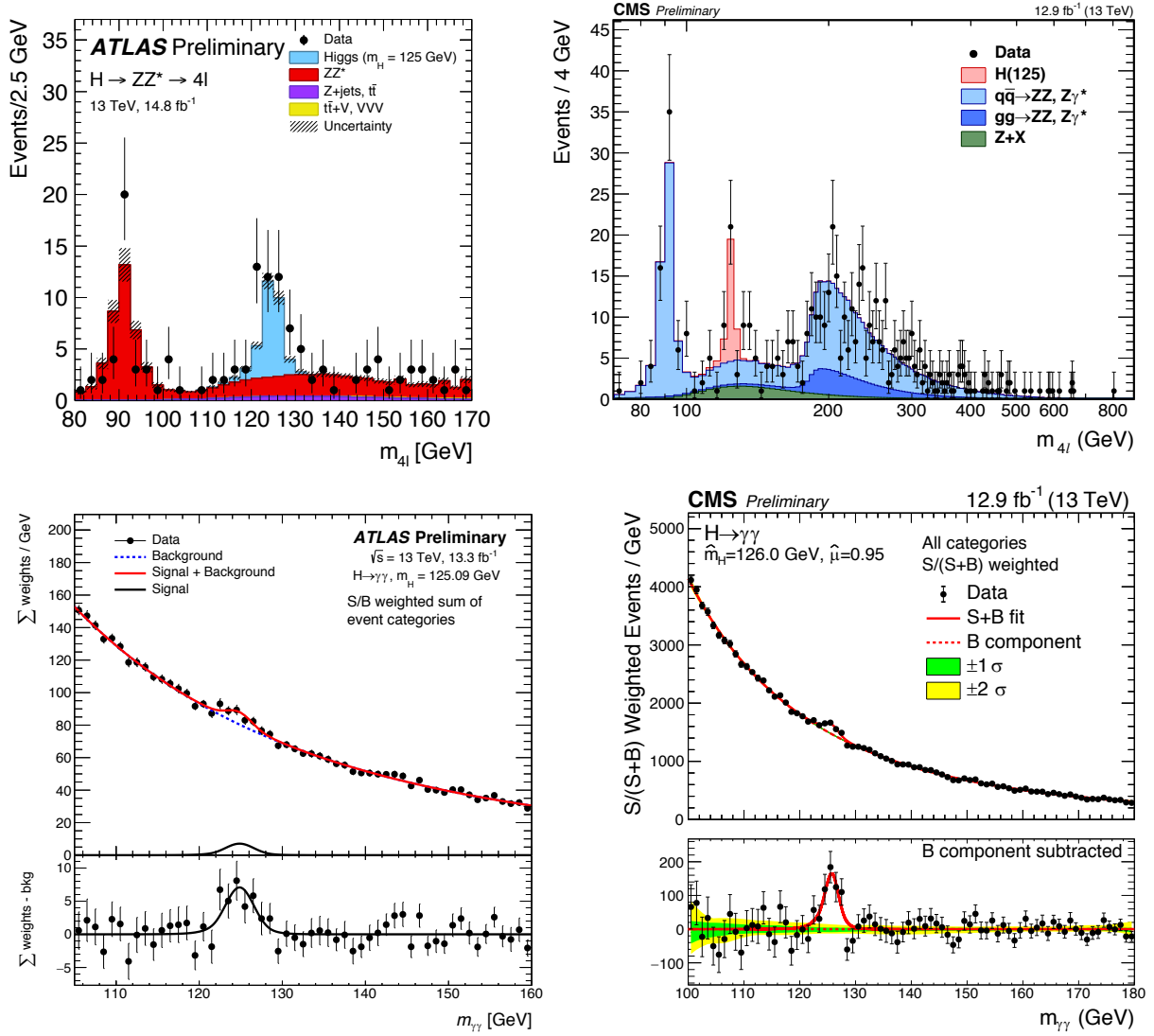
Figure 32 shows a rare and beautiful VBF  $H \rightarrow 4\ell$  candidate event. Such an event has large signal to background probability. Preliminary results for the cleanest bosonic channels  $H \rightarrow 4\ell$  and  $H \rightarrow \gamma\gamma$  were released by ATLAS and CMS for the 2016 summer conferences [122–125]. The Higgs boson was reobserved with high significance at the expected mass in both channels by either experiment (cf. Fig. 33 for the corresponding diphoton and four-lepton mass spectra). The extracted inclusive cross sections have still large uncertainties and are found in agreement with the SM expectations. In the four-lepton channel, ATLAS found a cross section of  $81^{+18}_{-16}$  pb compared to  $55 \pm 4$  pb expected. CMS measured a signal strength of  $\mu_{H \rightarrow 4\ell} = 0.99^{+0.33}_{-0.26}$  pb. CMS also measured the mass to be  $m_H = 124.50^{+0.48}_{-0.44}$  GeV in agreement with the Run-1 ATLAS and CMS combined value of  $125.09 \pm 0.24$  GeV. In the diphoton channel signal strengths of  $\mu_{H \rightarrow \gamma\gamma} = 0.85^{+0.22}_{-0.20}$  and  $0.91 \pm 0.21$  are measured by ATLAS and CMS, respectively. The measured cross sections versus centre-of-mass energy are shown in Fig. 34 for the inclusive cases (ATLAS, left) and fiducial measurements (CMS, right). ATLAS has combined the  $4\ell$  and  $\gamma\gamma$  results to perform a coupling analysis [126]. The combined inclusive cross section is also shown on the left panel of Fig. 34. The experiments also measured differential cross sections that are compared to NNLO plus parton shower predictions. No deviation from the SM prediction is found within yet large uncertainties.



**Fig. 32:** Display of  $H \rightarrow ee\mu\mu$  candidate from 13 TeV pp collisions measured by ATLAS. The event is accompanied by two forward jets with pseudorapidity difference of 6.4 and invariant dijet mass of 2 TeV. This event is consistent with VBF production of a Higgs boson decaying to four leptons.

ATLAS also released first preliminary studies of associated  $VH$  production with the decay  $H \rightarrow bb$  [127]. The channel is very challenging due to large backgrounds that need to be controlled with high precision in order to extract the signal. Run-1 had provided a signal strength slightly below the SM expectation [45, 128, 129]. The Run-2 yield was again low with  $\mu_{VH(\rightarrow bb)} = 0.21^{+0.51}_{-0.50}$  after combining the zero, one, and two charged lepton final states covering the  $ZH$  and  $WH$  modes.

ATLAS also looked into inclusive production of  $H \rightarrow \mu\mu$  [130] that has an expected branching fraction of 0.02%, but might be enhanced due to new physics effects. The sensitivity to that decay depends primarily on the dimuon mass resolution. The sensitivity can be improved, similarly to that of  $H \rightarrow \gamma\gamma$ , by splitting the event sample into categories with different mass resolution and/or signal signal-to-background ratios, such as low versus high  $p_T$ , central versus forward muons, ggF versus VBF, etc. The observed 95% confidence level limit for  $H \rightarrow \mu\mu$  is found at 4.4 times the SM prediction, reducing to 3.5 when combined with Run-1. About  $300 \text{ fb}^{-1}$  are needed to reach SM sensitivity. These results allow to exclude a universal Higgs coupling to fermions, as  $H \rightarrow \mu\mu$  would have been observed had it the same branching fraction as  $H \rightarrow \tau\tau$ .

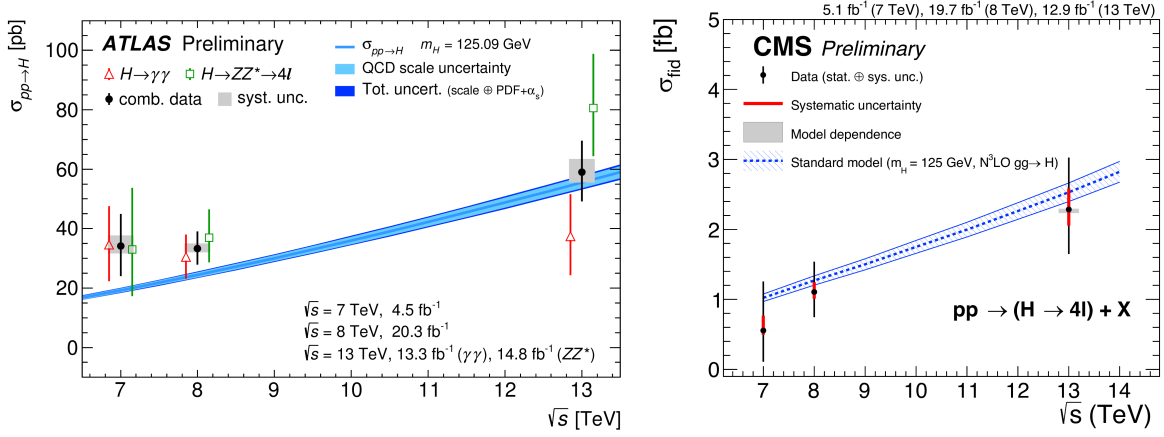


**Fig. 33:** Four-lepton (top row) and diphoton (bottom row) invariant mass distributions for ATLAS (left column) and CMS (right column) for the combined 2015 and 2016 datasets (ATLAS) and 2016 dataset (CMS) taken at 13 TeV proton–proton centre-of-mass energy [122–125]. The bottom plots show each event weighted by the signal-to-background ratio of the event category it belongs to.

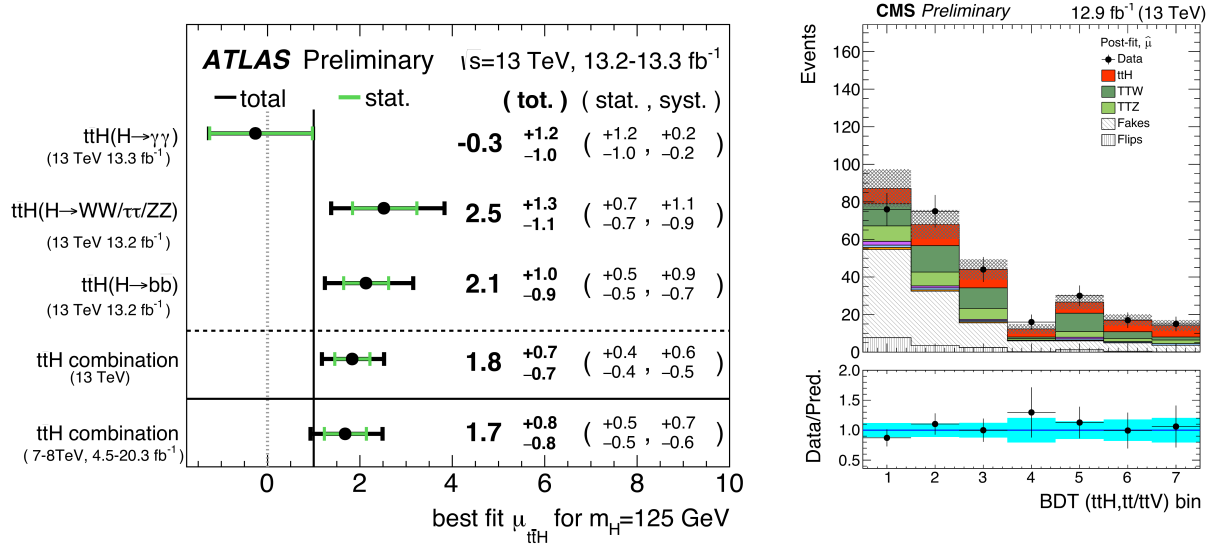
The Higgs production mode that most benefits from the increased centre-of-mass energy is  $ttH$  that was found a bit enhanced compared to the SM prediction in the Run-1 Higgs couplings combination (cf. Fig. 14, page 19). The motivation was thus large to look for that mode in Run-2.

The associated production of  $ttH$  is the only currently accessible channel that directly measures the top–Higgs coupling (cf. Feynman graph in Fig. 12 on page 18). All major Higgs decay channels,  $\gamma\gamma$ , multileptons, and  $bb$ , are analysed, where in particular the latter two channels represent highly complex analyses. The multilepton mode targets Higgs decays to  $\tau\tau$ ,  $WW \rightarrow 2\ell 2\nu$ , and  $ZZ \rightarrow 2\ell 2\nu$ ,  $4\ell$  together with at least one top quark decaying leptonically. It requires at least two leptons with the same charge, which greatly reduces SM backgrounds. The dominant remaining backgrounds are misidentified prompt leptons and  $ttV$  production in particular the difficult to separate  $ttW$  (cf. the right panel in Fig. 35 for the distribution of a boosted decision tree trained to distinguish  $ttW$  background from  $ttH$  signal). The  $H \rightarrow bb$  mode is analysed in the one and two lepton channels. Here the biggest challenge represents





**Fig. 34:** Total (ATLAS, left) and fiducial (CMS, right)  $pp \rightarrow H + X$  cross sections measured at different centre-of-mass energies and compared to SM predictions at up to 3NLO in QCD. The left plot shows the individual results of the  $4\ell$  and  $\gamma\gamma$  channels and their combination [126]. The right plot shows the  $4\ell$  fiducial measurements. Agreement with the SM predictions is observed within yet large uncertainties.



**Fig. 35:** Left: ATLAS summary of the  $\mu_{ttH}$  signal strength measurements from the individual analyses and their combination, assuming  $m_H = 125$  GeV [135]. Right: boosted decision tree output from CMS in the same-charge channel trained to separate  $ttW$  background from  $ttH$  signal [131].

background due to  $t\bar{t}$  production associated with heavy flavour quarks ( $c$  or  $b$ ) originating mostly from gluon splitting, which is poorly known and needs to be constrained from data simultaneously with the signal. CMS released preliminary 13 TeV results for  $ttH$  in all three Higgs decay categories finding for the relative signal strengths [131–133]:  $\mu_{ttH(\rightarrow\gamma\gamma)} = 3.8^{+4.5}_{-3.6}$ ,  $\mu_{ttH(\rightarrow\text{leptons})} = 2.0^{+0.8}_{-0.7}$ , and  $\mu_{ttH(\rightarrow bb)} = -2.0 \pm 1.8$ , with no significant excess observed. ATLAS also measured all three channels [134] and their statistical combination [135] that is shown in the left panel of Fig. 35. The combined preliminary signal strength is  $\mu_{ttH} = 1.7 \pm 0.8$ .

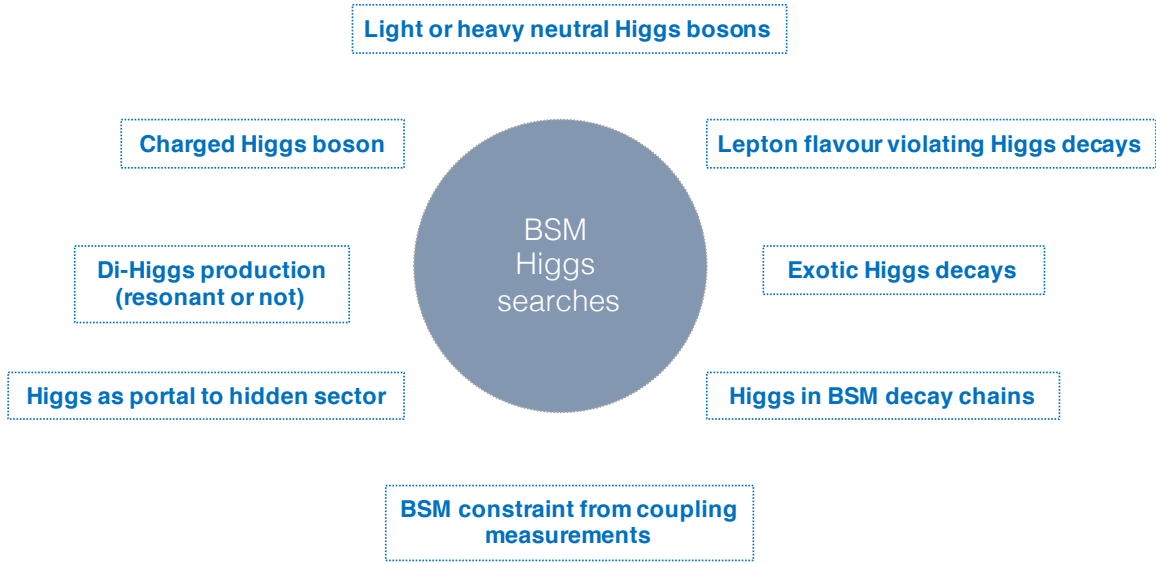


Fig. 37: Illustration of beyond the SM Higgs boson search areas.

### 8.3 Searches — a fresh start

Many of the high mass and higher cross section searches for new physics already benefited from the 2015 13 TeV data sample to extend their sensitivity, and all searches surpass their Run-1 limits with the 2016 datasets (see Fig. 36). Run-2 represents thus a fresh start in the quest for new physics after the negative searches from Run-1. The legacy of Run-1 also contained a small number of anomalies that needed to be verified in the Run-2 data. Only 13 TeV searches are discussed in the following.

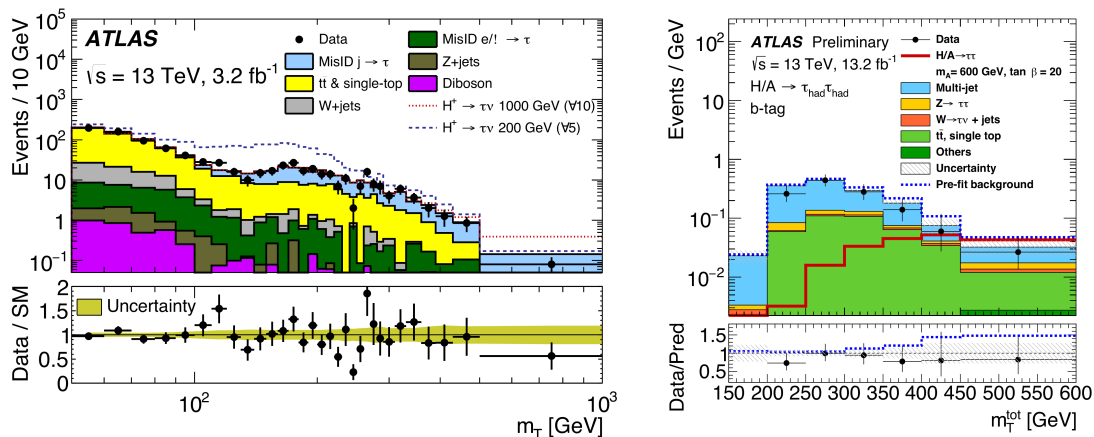


Fig. 36: The bulldozer (aka, LHC at 13 TeV) moving out of the way the Run-1 limits on beyond the SM searches.

#### 8.3.1 Additional Higgs bosons

The 125 GeV Higgs boson completes the four degrees of freedom of the SM BEH doublet. Nature may have, however, chosen a more complex scalar sector of, eg., two BEH doublets, which extends the sector by four additional Higgs bosons, of which two are neutral (one  $CP$ -even and one  $CP$ -odd) and the other two are charged. Searching for ancillary scalar bosons is thus one way to detect BSM physics in the scalar sector. Other ways are to look for non-SM decays of the Higgs boson such as decays to invisible particles where the Higgs boson acts as a portal to new physics responsible for dark matter. The Higgs boson could also be produced as a particle in the decay chain of new physics processes such as supersymmetry. New heavy resonances might decay to a pair of Higgs boson. A summary of BSM options around the Higgs boson is sketched in Fig. 37.

ATLAS and CMS have searched for additional Higgs bosons in Run-1 and Run-2. For  $H^\pm \rightarrow \tau\nu$  [136, 137] ( $H/A \rightarrow \tau\tau$  [138, 139]), the sensitivity of the new data exceeds that of Run-1 for masses larger than 250 GeV (700 GeV). The search for  $A \rightarrow Z(\rightarrow \ell\ell, \nu\nu)h_{125}(\rightarrow bb)$  features improved sensitivity beyond about 800 GeV [140]. Searches for  $H \rightarrow ZZ(\rightarrow \ell\ell qq, \nu\nu qq, 4\ell)$  and  $WW(\rightarrow \ell\nu qq)$



**Fig. 38:** Left: transverse mass in a search for a charged Higgs boson decaying to  $\tau\nu$  [136, 137]. Right: distribution of the reconstructed transverse mass in a search for a heavy neutral Higgs boson decaying to a tau tau pair, where both  $\tau$  leptons are reconstructed via their hadronic decay modes [138].

target the  $> 1$  TeV mass range where the bosons are boosted and their hadronic decays are reconstructed with jet substructure techniques. The search for a resonance decaying to  $hh_{125}(\rightarrow bb\gamma\gamma)$  had a small excess in Run-1 at about 300 GeV [141], which has been excluded at 13 TeV [142, 143]. Also performed were searches for resonant and non-resonant  $H_{125}H_{125} \rightarrow bb\tau\tau, bbVV_{V=Z/W}, bbbb$  production [144–147]. None of these many searches exhibits an anomaly so far in the 13 TeV data.

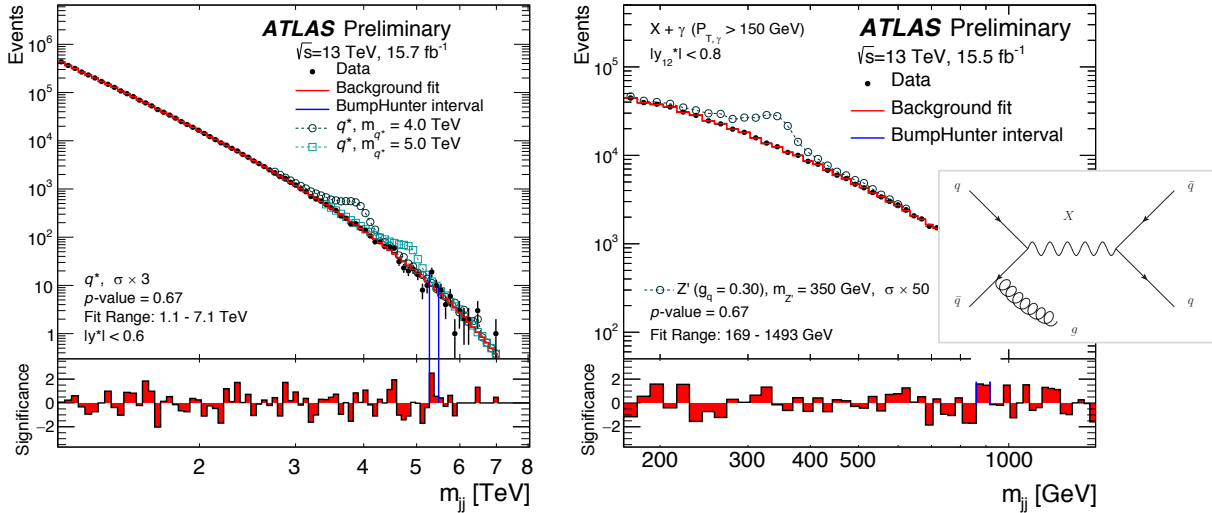
A slight Run-1 excess of  $2.4\sigma$  seen by CMS in the search for the lepton-flavour violating decay  $H \rightarrow \tau\mu$  [148] was not seen by ATLAS in Run-1 [149], and also not confirmed by CMS in an early Run-2 analysis [150].

### 8.3.2 New physics searches in events with jets

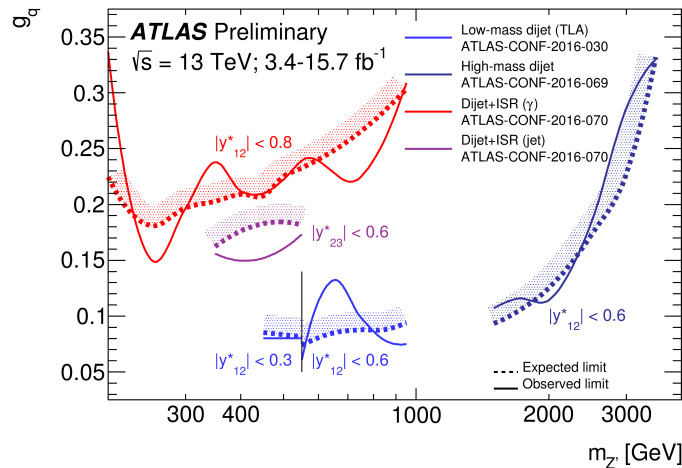
Among the first searches performed at any increase of collision energy are those for heavy strongly interacting new phenomena such as excited quarks due to quark substructure, or strong gravity effects. The signatures investigated are a dijet resonance and angular distributions, a resonance decaying to heavy-flavour quarks  $X \rightarrow b\bar{b}$  or  $t\bar{t}$  [156, 157], high- $p_T$  multijet events, high- $p_T$  lepton plus jets events, and a lepton-jet resonance as could occur in presence of heavy leptoquarks. None of these searches exhibited an anomaly.

Figure 39 shows dijet invariant mass spectra as measured by ATLAS [151, 152] (see [153] for the corresponding CMS analysis). The left panel shows the high-mass tail as obtained with standard unpre-scaled jet triggers. The right panel shows lower mass events obtained with the use of a hard ISR jet trigger (see Feynman graph in right panel). The measured spectra are compared to phenomenological fits using smoothly falling functions as expected from the QCD continuum. No significant deviation from these fits is seen in the data. In addition to the ISR “trick”, it is possible to reach the low mass dijet regime with high statistics by using high-rate trigger-level objects of events that are prescaled for offline analysis (this technique is denoted Data Scouting in CMS) [154, 155]. Figure 40 shows the combined exclusion plot obtained by ATLAS in the coupling-vs-mass plane for a hypothetical leptophobic  $Z'$  resonance. The low-mass region is covered by ISR-based searches both for ISR jets and photons. For intermediate masses the trigger-level analysis (TLA) provides the strongest bounds, and for high  $Z'$  masses the standard dijets search takes over, smoothly extending the TLA bound.





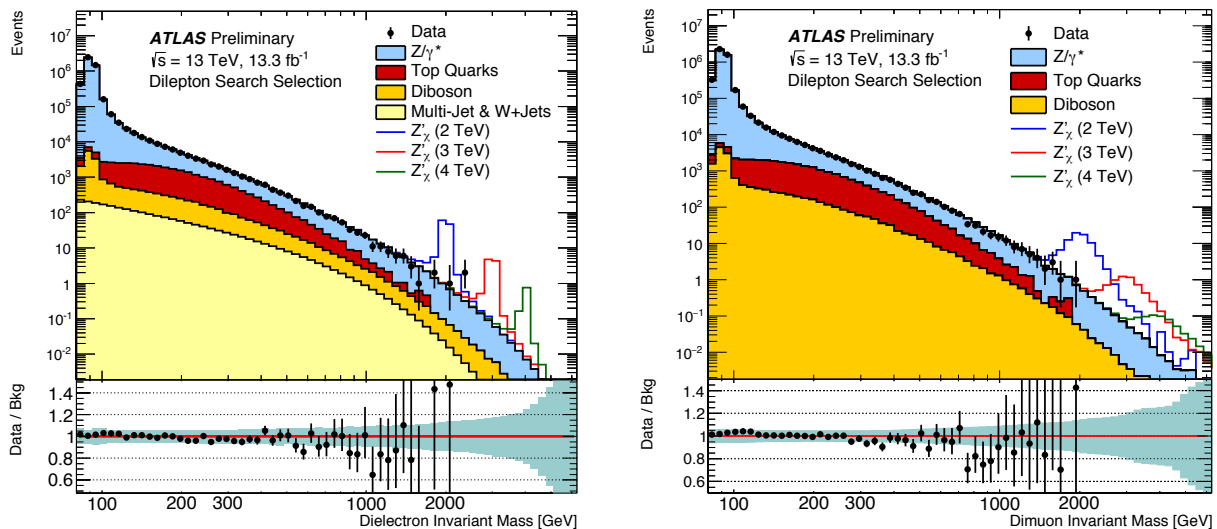
**Fig. 39:** Dijet invariant mass distributions measured by ATLAS for the high-mass resonance search [151] (left panel) and the low-mass search (right) using events with significant initial-state radiation (cf. Feynman graph in panel) [152].



**Fig. 40:** ATLAS bounds in the coupling-vs-mass plane on a leptophobic  $Z'$  model obtained from dijet searches.

### 8.3.3 Searches in leptonic final states

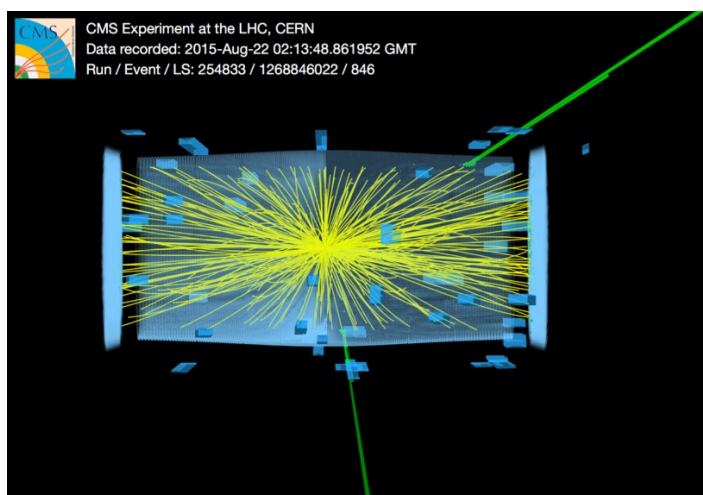
Canonical searches for new physics are performed in high-mass Drell-Yan production ( $Z' \rightarrow \ell\ell$ ,  $W' \rightarrow \ell\nu$ ) [158–161]. These searches require faithful SM Drell-Yan modelling that is tested using SM differential cross section measurements. High transverse momentum muons represent a challenge for the detector alignment, requiring, eg., down to 30  $\mu\text{m}$  relative alignment precision in the ATLAS muon spectrometer. The electron and muon channels have complementary strength: the electron energy resolution measured in the calorimeters being more precise than the muon track momentum resolution, the electron channel has better discovery sensitivity. On the other hand, there is almost no charge information from the electron tracks, so the muon channel is needed to measure the charge of a resonance if detected (cf. the panels in Fig 41). No anomaly was found in the measured spectra. Sequential SM  $Z' / W'$  benchmark limits are set at 4.1 / 4.7 TeV (compared to 2.9 / 3.3 TeV at 8 TeV). Figure 42 shows the highest-mass dielectron event measured by CMS in the early 2015 data. It has an invariant mass of 2.9 TeV. For comparison, the highest-mass Run-1 events have 1.8 TeV ( $ee$ ) and 1.9 TeV ( $\mu\mu$ ).



**Fig. 41:** Dielectron (left panel) and dimuon (right panel) reconstructed invariant mass distributions for data and the SM background estimates as well as their ratios. Benchmark  $Z'$  signals with masses of 2, 3 and 4 TeV are overlaid [158].

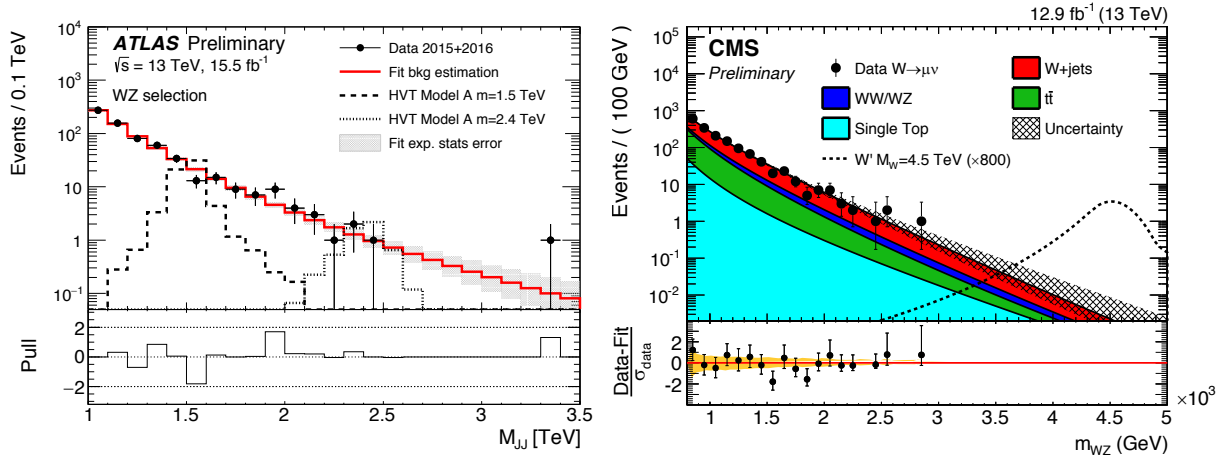
ATLAS and CMS also looked into high-mass  $e\mu$  production not accompanied by neutrinos that would violate lepton flavour conservation. The main background here are top–antitop events that are estimated from MC simulation.

The following table [163] gives a historical evolution of the 95% confidence level lower limits for selected leptonically and hadronically decaying benchmark resonances from Tevatron, via LHC up to the HL-LHC expectation (see [162] for the latter studies). The corollary from these numbers is that future improvement in reach will take more time.



**Fig. 42:** Display of a rare, colossal  $e^+e^-$  event with mass of 2.9 TeV measured by CMS. The electrons are azimuthally back-to-back.

| Model                               | 95% confidence level lower limits (TeV) |           |             |           |                                     |                                      |
|-------------------------------------|---|-----------|-------------|-----------|-------------------------------------|--------------------------------------|
|                                     | CDF                                     | Run-1 '12 | Moriond '16 | ICHEP '16 | 300 fb <sup>-1</sup><br>14 TeV $pp$ | 3000 fb <sup>-1</sup><br>14 TeV $pp$ |
| $Z' \rightarrow \ell\ell$           | 1.1                                     | 2.9       | 3.4         | 4.1       | 6.5                                 | 7.8                                  |
| $q^* \rightarrow qq$                | 0.9                                     | 4.1       | 5.2         | 5.6       | 7.4                                 | 8                                    |
| $Z' \rightarrow tt$<br>(1.2% width) | 0.9                                     | 1.8       | 2.0         | –         | 3.3                                 | 5.5                                  |



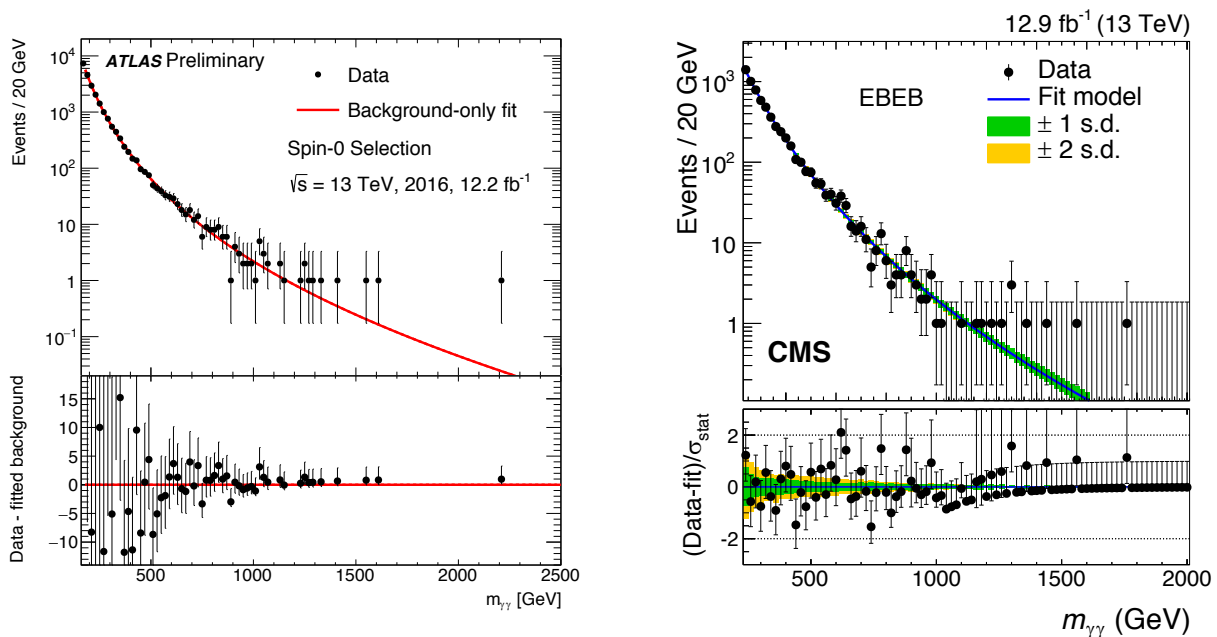
**Fig. 43:** Diboson mass in the fully hadronic channel (ATLAS [166], left panel) and semileptonic channel (CMS [167], right) for data and background expectations.

### 8.3.4 Searches for diboson resonances ( $VV$ , $Vh$ , $hh$ )

Diboson resonances occur in many new physics scenarios and also in extended scalar sector models. If the resonances are heavy, the high transverse momentum of the decaying bosons boosts the hadronic decay products into merged jets. Jet substructure analysis is used to reconstruct hadronically decaying bosons and to suppress strong interaction continuum backgrounds. Some excess of events with a (global) significance of  $2.5\sigma$  was seen by ATLAS in Run-1 around a mass of 2 TeV in fully hadronically decaying  $VV$  events (mostly  $WZ$ ) [164,165], which was however not observed in the other weak gauge boson decay channels of similar sensitivity. The excess is not confirmed in Run-2 [166,167] (cf. Fig. 43).

Searches for a new resonance in the diphoton mass spectrum were performed by ATLAS [168,169] and CMS [170] in Run-1 looking for a low to medium mass scalar resonance, or a medium to high mass spin-two resonance motivated by strong gravity models. Diphoton spectra were also analysed in view of high-mass tail anomalies due to new nonresonant phenomena. Searches involving at least three photons were used during Run-1 to look for new physics in Higgs or putative  $Z'$  decays [171].

Preliminary analyses of the 13 TeV diphoton data presented at the 2015 end-of-year seminars showed an excess of events at around 750 GeV invariant diphoton mass in ATLAS and, albeit weaker, also in CMS. In spring 2016, reanalyses of the 2015 data were published by ATLAS [172] and CMS [173] confirming the preliminary results. CMS also included  $0.6 \text{ fb}^{-1}$  of data taken without magnetic field requiring a dedicated reconstruction. The photons are tightly identified and isolated and have a typical purity of 94%. The background modelling uses empirical functions fit to the full invariant mass spectra (ATLAS uses a theoretical model to describe the background shape in the spin-2 case). ATLAS observed the lowest background-only p-value for a resonance at around 750 GeV with a natural width of about 45 GeV (6% with respect to the mass). The local and global significance was found to be  $3.9\sigma$  and  $2.1\sigma$ , respectively. The global significance was derived by running background-only pseudo-experiments, modelled according to the fit to data, and by evaluating for each experiment the mass and width that leads to the largest excess, that is, the lowest p-value. One then counts the fraction of experiments with a p-value lower than that in data. This procedure corrects the local p-value for the trials factor (also called “look-elsewhere effect”). Indeed, the local p-value corresponds to a non-normalised probability that does not have a well-defined interpretation. Only the global p-value defines a proper probability and is thus the correct reference value. CMS also found its lowest p-value at around 750 GeV at, however, a narrow width. Combining 8 TeV and 13 TeV data a global (local) significance of  $1.6\sigma$  ( $3.4\sigma$ ) was seen. These



**Fig. 44:** Diphoton mass spectrum in the summer 2016 dataset from ATLAS (left panel) and CMS (right). There is no noticeable excess at around 750 GeV.

results have prompted intense theoretical activity.

As is well known the first 12–13 fb<sup>-1</sup> data taken in 2016 did not reproduce the excess in neither experiment [174, 175] (cf. Fig. 44). The excesses in the 2015 data were thus the result of a statistical fluctuation which, given the global significance, is not that unlikely to occur. One should also take into consideration that the actual trials factor is larger than the global factor quoted for these analyses as there are many signatures probed by the experiments. This truly global significance of a local excess is hard or impossible to estimate in a thorough manner, but the additional trials factor should be kept in mind. In that respect, having a second experiment with a similar non-significant excess does not remove the trials factor if the results from both experiments are retained. Removing the 2015 data and looking solely at 750 GeV in the 2016 data does, however, properly remove any trials factor.

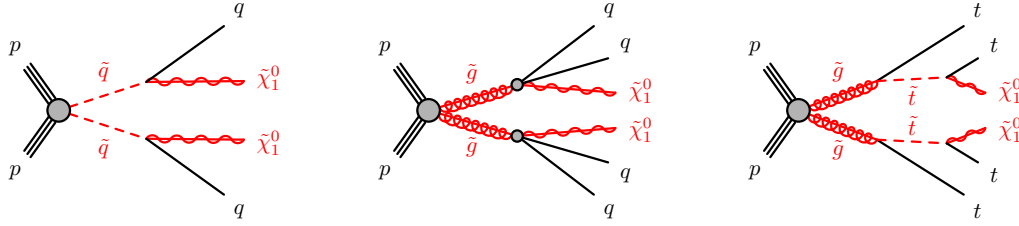
### 8.3.5 Supersymmetry

Supersymmetry (SUSY) is still among the most popular SM extensions owing to the elegance of the theoretical ansatz, and its phenomenological appeal by offering potential solutions to the hierarchy problem,<sup>20</sup> grand unification of the gauge couplings, and dark matter. However, if the SM is included in a supersymmetric theory with SUSY particles (sparticles) that differ by half-a-unit of spin from their SM partners, how is it possible that more than half the particles in the superworld have escaped our observations?

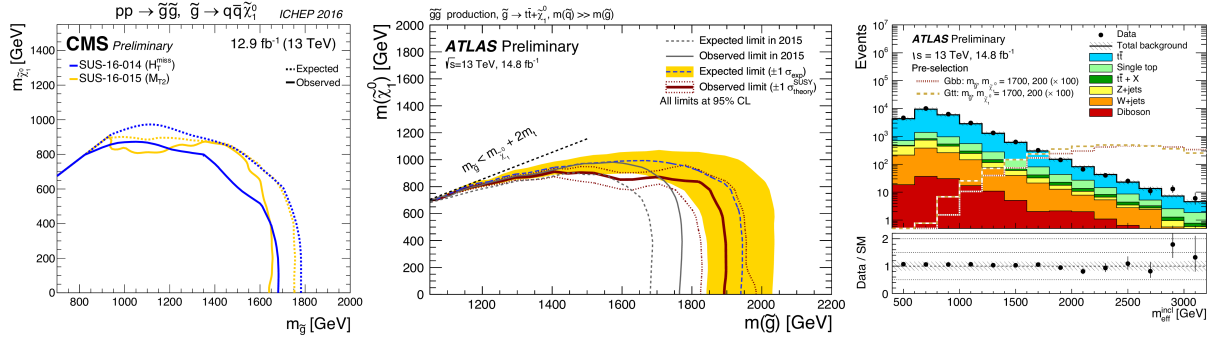
Due to SUSY breaking, allowing the sparticles to acquire large masses,<sup>21</sup> SUSY comes with very diverse signatures. Highest cross-section events produce gluino or squark pairs with decays to jets and missing

<sup>20</sup>As the SM, SUSY is a weakly coupled approach to electroweak symmetry breaking in which the Higgs boson remains elementary.

<sup>21</sup>In unbroken SUSY, fermionic  $|f\rangle$  and bosonic  $|b\rangle$  partner states, transformed into each other via the SUSY generator  $Q$ , have the same mass, as  $P^2 Q|b\rangle = P^2|f\rangle = m_f^2|f\rangle$ ,  $P^2 Q|b\rangle = Q P^2|b\rangle = Q m_b^2|b\rangle = m_b^2|f\rangle$ , and hence  $m_f^2 = m_b^2$ .  $P^2$  is the square of the energy-momentum operator which commutes with  $Q$ .



**Fig. 45:** Graphs for simplified models describing the pair production and decay of supersymmetric particles. Left: squark pair production and decay to two quark-jets and two neutralinos; middle: gluino pair production and decay to four quark-jets and two neutralinos; right: gluino pair production and decay to four top quarks and two neutralinos. The top quarks will each further decay to a  $W$  boson and  $b$  quark.



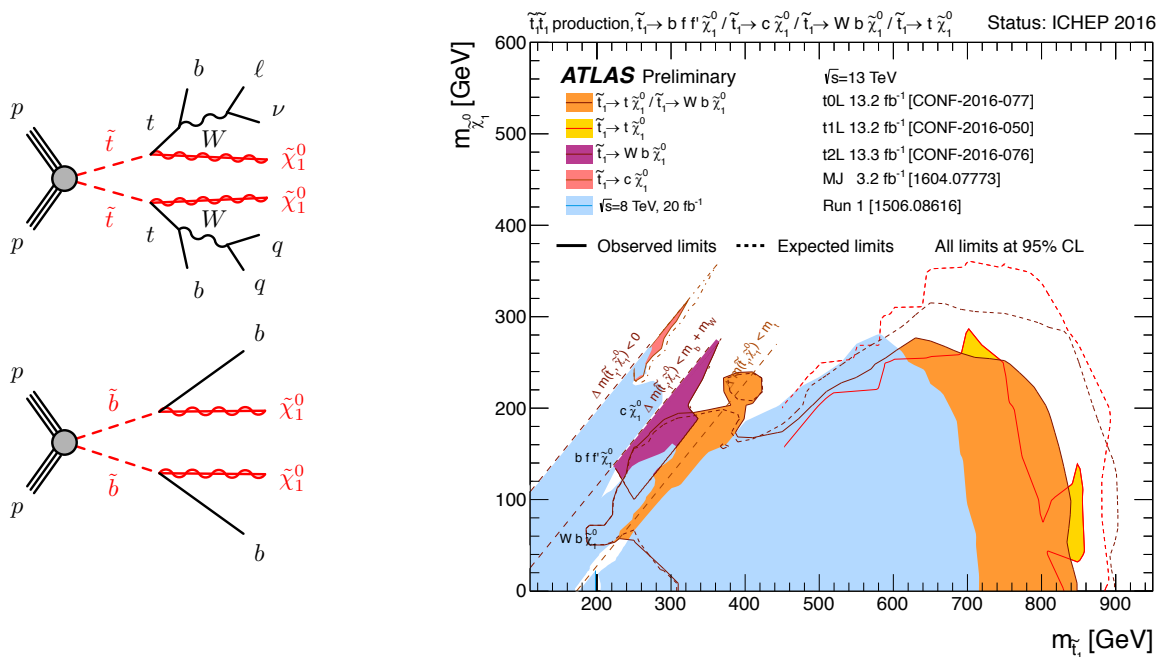
**Fig. 46:** Left: exclusion limits for strong gluino pair production and decay to four jets and two lightest neutralinos as obtained by CMS [178, 179]. Observed limits are shown with solid lines and expected with dashed lines. Middle: exclusion limits obtained for strong gluino pair production and decay via stop squarks into a four top quark final state and two neutralinos. Right: effective mass distribution obtained in the search for gluino mediated stop production [180] (cf. Footnote 8, page 11 for the definition of the effective mass).

transverse momentum if  $R$ -parity is conserved.<sup>22</sup> Naturalness suggests not too heavy SUSY top, weak and Higgs boson partners to effectively cancel the radiative corrections to the Higgs mass at high scale and hence provide a solution to the hierarchy problem.<sup>23</sup> It might thus occur that stop pair production, or gluino pair production and decay via stop and top to a four-top final state are the dominant SUSY processes at the LHC. If all strongly interacting SUSY particles are too heavy to be directly produced at the LHC, it could still be that the electroweak partners of the photon, weak bosons and five physical Higgs states are light enough so that SUSY would manifest itself through “electroweak-ino” production featuring final states with leptons (and/or photons) and  $E_T^{\text{miss}}$ . Finally, SUSY could also give rise to the existence of long-lived heavy particles, and, if  $R$ -parity is nonconserved, the lightest SUSY particle could decay to jets or leptons depending on the  $R$ -parity violating couplings.

To approach the search for SUSY in a systematic manner, a bottom-up approach through so-called simplified models is used by the experiments. These models correspond to simple signatures as those depicted in Fig. 45. While a simplified model cannot encompass the full SUSY phenomenology, an ensemble of simplified models and the corresponding searches have been shown to cover signatures of complete

<sup>22</sup> $R$ -parity, defined by  $R = (-1)^{3(B-L)+2S}$ , where  $B$ ,  $L$ ,  $S$  are the baryon, lepton numbers and spin, respectively, is assumed to be conserved in most SUSY models to avoid baryon and/or lepton number violation (and thus proton decay if both occur together).  $R$ -parity conservation is arbitrarily imposed and not enforced by any known symmetry. Its consequence is that SUSY particles must be produced in pairs and the lightest SUSY particle is stable.

<sup>23</sup>The top quark gives the largest contribution to the radiative corrections of the Higgs mass,  $\delta m_H^2$ , in presence of a high new physics scale  $\Lambda$ . If the stop is heavier than the top residual logarithmic contributions  $\delta m_H^2 \propto \ln(\Lambda^2/m_t^2)$  remain.



**Fig. 47:** Graphs for simplified models of stop and sbottom pair production and decay. The right panel shows exclusion limits in the neutralino versus stop mass plane as obtained by ATLAS with several dedicated analyses (see references in figure).

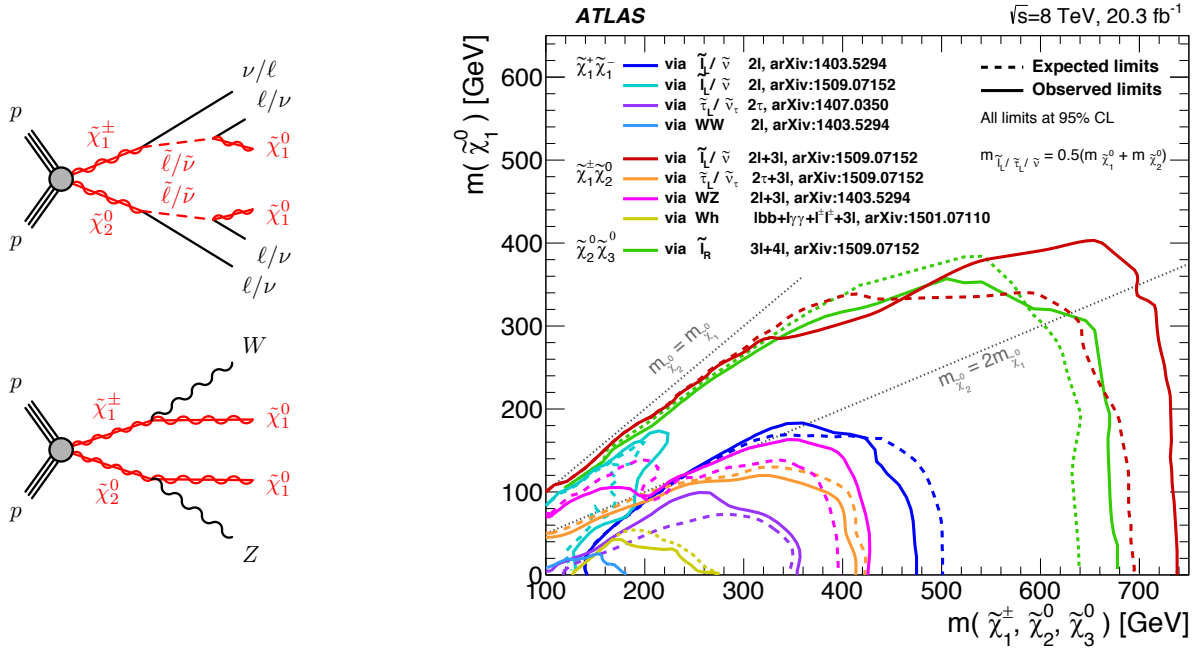
models such as the phenomenological minimal supersymmetric standard model (pMSSM) [176, 177].

Searches for strong SUSY production study events with jets and  $E_T^{\text{miss}}$  with or without leptons, photons, and  $b$ -jets. Up to ten jets are exclusively selected, which requires a data-driven background determination as MC cannot be trusted to reliably predict such large jet multiplicity. None of the searches have revealed a significant anomaly. Figure 46 shows on the left and middle plots exclusion limits in the lightest neutralino mass versus gluino mass planes. The analyses have the sensitivity to exclude gluinos of up to 1.8 TeV for low-mass neutralinos depending on the scenarios. In case of heavy neutralinos, the final states exhibit softer jets and less  $E_T^{\text{miss}}$ , which leads to reduced trigger efficiency and a more difficult background discrimination thus reducing the sensitivity. The right panel in Fig. 46 shows the effective mass distribution obtained in the search for gluino mediated stop production (four top quark final state). The distribution reaches beyond 3 TeV with the dominant background from top-quark production. No excess of events is seen in data compared to the background estimation.

If gluinos are too heavy to be produced in significant quantities, squark mixing could make third generation squarks lighter than the first and second generation squarks. Direct searches for stop and sbottom squark production have been the topic of intense efforts in both ATLAS and CMS since Run-1. The analyses are distinguished according to the number of identified leptons (0, 1, 2) and differently optimised signal regions target different stop/sbottom and neutralino mass regimes. In the stop case, the signatures also depend on whether the stop decays in a two-body signature to an on-shell top quark and the lightest neutralino, or off-shell via three or four body decays to the top decay products. The right panel in Fig. 47 shows the exclusion limits for simplified models of stop and sbottom pair production and decay obtained by ATLAS with several dedicated analyses. The analyses have sensitivity to exclude stop masses up to 900 GeV. As in the gluino and first-generation squark cases, the limits for heavy neutralinos are significantly worse.

Alternative models for new heavy quark partners introduce, for example, vector-like quarks, which are





**Fig. 48:** Graphs for simplified models of associated lightest chargino and next-to-lightest neutralino production and decay through sleptons (top graph) or, if sleptons are too heavy, through  $W$  and  $Z$  bosons. The right plot shows the exclusion bounds obtained by ATLAS for various electroweakino production scenarios (see references in figure).

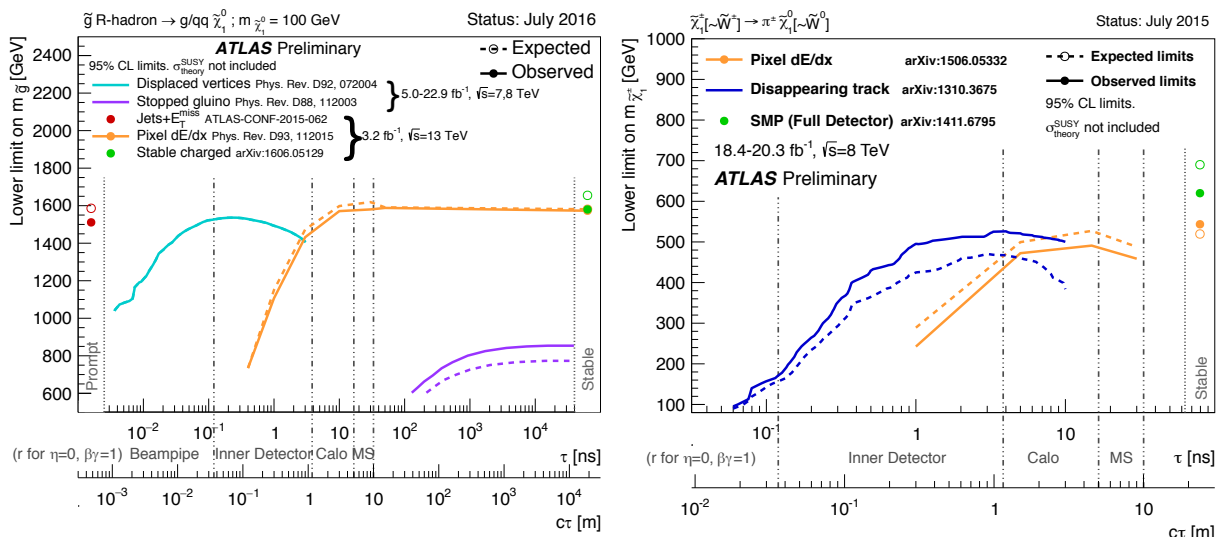
hypothetical fermions that transform as triplets under colour and who have left-handed and right-handed components with same colour and electroweak quantum numbers. Vector-like quarks can be singly or pair produced and decay to  $bW$ ,  $tZ$  or  $tH$ . Also exotic  $X_{5/3} \rightarrow tW$  processes may exist.

It could also be that all squarks and gluinos are beyond reach of the current LHC sensitivity and electroweakinos are the lightest fermions. They have low cross-sections, so that the present Run-2 luminosity just suffices to surpass the Run-1 sensitivity. Figure 48 shows graphs for simplified models of associated lightest chargino and next-to-lightest neutralino production and decay through sleptons (top graph) or, if sleptons are too heavy, through  $W$  and  $Z$  bosons. The right plot shows the exclusion bounds obtained by ATLAS for various electroweakino production scenarios. Electroweakino decays via sleptons are a favourable case due to the larger leptonic rate than in weak boson decays. In the models considered, chargino pair production has lower cross section than  $\tilde{\chi}_1^+ \tilde{\chi}_2^0$  production. The cross section depend on the mixing properties of the states: neutralinos can be bino, wino or higgsino like; charginos wino or higgsino like, depending on the dominant contribution.<sup>24</sup>

### 8.3.6 Search for massive long-lived massive particles

Massive long-lived heavy particles are predicted in many new physics models. They can occur due to large virtuality (such as predicted in split supersymmetry), low couplings (such as predicted in some gauge mediated SUSY breaking scenarios where the gravitino is the lightest SUSY particle), and mass

<sup>24</sup>There are a total of eight spin-half partners of the electroweak gauge and Higgs bosons: the neutral bino (superpartner of the  $U(1)$  gauge field), the winos, which are a charged pair and a neutral particle (superpartners of the  $W$  bosons of the  $SU(2)_L$  gauge fields), and the higgsinos, which are two neutral particles and a charged pair (superpartners of the Higgs field's degrees of freedom). The bino, winos and higgsinos mix to form four charged states called charginos ( $\tilde{\chi}_i^\pm$ ) and four neutral states denoted neutralinos ( $\tilde{\chi}_i^0$ ). Their indices  $i$  are ordered according to the increasing mass of the  $\tilde{\chi}_i$  state.



**Fig. 49:** Exclusion limits on the gluino mass versus lifetime (left panel) and chargino mass versus lifetime (right) as obtained by ATLAS (see references in figure). The dots on the left (right) of the plots indicate the limits obtained on promptly decaying (stable) gluinos/charginos. Varying searches cover the full lifetime spectrum.

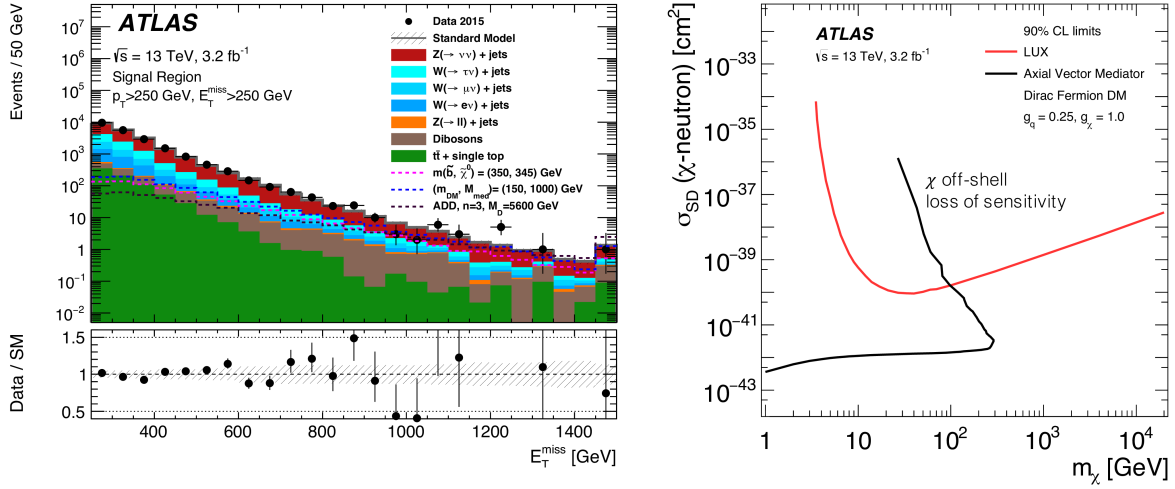
degeneracy in a cascade decay, eg., via a scale-suppressed colour triplet scalar from unnaturalness [181] or anomaly-mediated SUSY breaking scenarios with a wino-like lightest chargino [182]. The search for massive long-lived particles is a key part of the LHC search programme.

The LHC experiments search for massive long-lived particles using measurements of specific ionisation loss in the tracking detectors, the time-of-flight in the calorimeters and muon systems, and by reconstructing displaced vertices, kinked or disappearing tracks. Looking for calorimeter deposits outside of the colliding proton bunches makes it possible to look for very long-lived strongly interacting massive particles that were stopped in the calorimeter layers [183, 184]. Some signatures need dedicated triggers, most require novel analysis strategies to determine backgrounds from data. Figure 49 shows exclusion limits on the gluino mass versus lifetime (left panel) and chargino mass versus lifetime (right) as obtained by ATLAS. The dots on the left (right) of the plots indicate the limits obtained on promptly decaying (stable) gluinos/charginos. Varying searches cover the full lifetime spectrum. It is interesting to observe that the standard SUSY searches are not blind to scenarios with long-lived sparticles if their lifetime is short enough to still decay before the calorimeter.

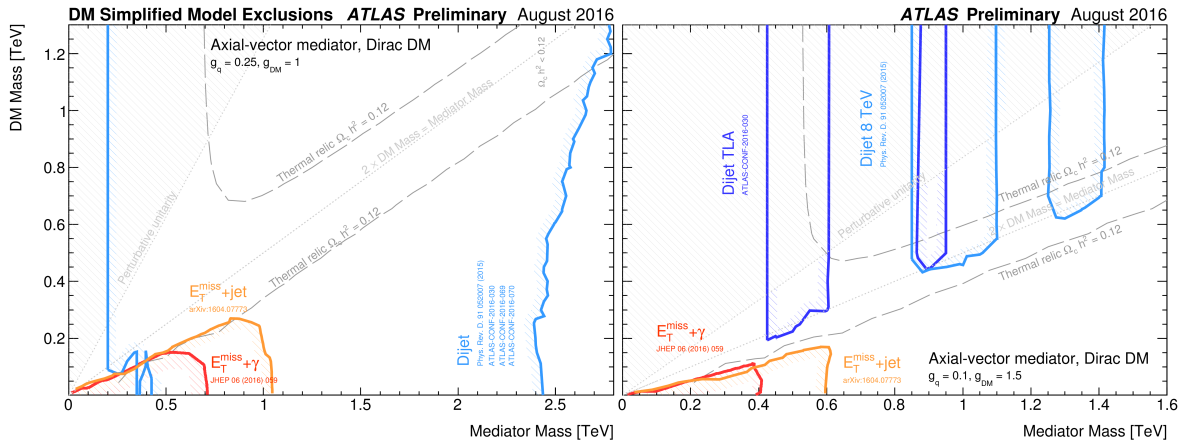
### 8.3.7 Searches for dark matter production

If dark matter particles (assumed to be weakly interacting and massive, WIMPs) interact with quarks and/or gluons they can be directly pair produced in the proton collisions at the LHC [185]. Since the WIMPs remain undetected, to trigger the events a large boost via initial state jet or photon radiation (or other recoiling particles) is needed leading to large missing transverse momentum from the recoiling WIMP pair. The final state signature depends on the unknown details of the proton–WIMP coupling. A set of “ $X + E_T^{\text{miss}}$ ” searches is therefore needed for full experimental coverage. The most prominent and among the most sensitive of these is the so-called “mono-jet” search, which extends to a couple of high- $p_T$  jets recoiling against the  $E_T^{\text{miss}}$  (cf. Fig. 50). Large irreducible SM backgrounds in this channel stem from  $Z(\rightarrow \nu\nu) + \text{jets}$  and  $W(\rightarrow \ell\nu) + \text{jets}$  events, where in the latter case the charged lepton is either undetected or a hadronically decaying tau lepton. These backgrounds are determined in data control regions requiring accurate input from theory to transfer the measured normalisation scale factors to the





**Fig. 51:** Left: distribution of missing transverse momentum measured by ATLAS at 13 TeV in a “mono-jet” search [186]. The dominant backgrounds stem from leptonic  $Z$  and  $W$  plus jets events. Also shown are distributions for new physics benchmark models. Right: exclusion limit on the spin-dependent WIMP–neutron scattering cross section versus the WIMP mass in the context of a  $Z'$ -like simplified model with axial-vector couplings. The result is compared with limits from the LUX experiment [195]. All limits are shown at 90% confidence level, which is the standard benchmark in direct dark matter detection experiments. The comparison to LUX is valid solely in the context of this model, assuming minimal mediator width and the coupling values  $g_q = 1/4$  and  $g_\chi = 1$  [186].

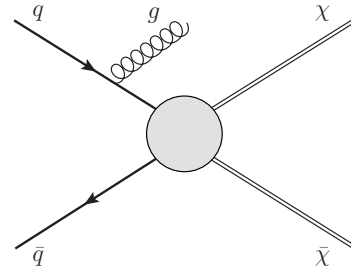


**Fig. 52:** Regions in a dark matter (DM) versus mediator mass planes excluded at 95% CL by a selection of ATLAS DM searches, for a possible interaction between the SM and DM, the lepto-phobic axial-vector mediator described in [185]. The left panel shows exclusion bounds for quark coupling  $g_q = 1/4$ , universal to all flavors, and dark matter coupling  $g_{DM} = 1$ . On the right panel  $g_q = 1/10$  and  $g_{DM} = 3/2$  are assumed. Shown are the results from the monojet, monophoton and dijet resonance searches. Dashed curves labelled “thermal relic” indicate combinations of DM and mediator mass that are consistent with the cosmological DM density and a standard thermal history. Between the two curves, annihilation processes described by the simplified model deplete the relic density. A dotted curve indicates the kinematic threshold where the mediator can decay on-shell into DM. Points in the plane where the model is in tension with perturbative unitary considerations are indicated by the shaded triangle at the upper left. The exclusion regions, relic density contours, and unitarity curve are not applicable to other choices of coupling values or model. See [196] for more information.

signal regions.

Numerous 13 TeV results have been released by ATLAS and CMS, including jets +  $E_T^{\text{miss}}$  [186, 187], photon +  $E_T^{\text{miss}}$  [188, 189], Z/W +  $E_T^{\text{miss}}$  [190, 191], and  $bb/tt + E_T^{\text{miss}}$  [192–194] signatures. None of these has so far shown an anomaly. Figure 51 shows the missing transverse momentum distributions measured by ATLAS in the monojet jets +  $E_T^{\text{miss}}$  search.

Since the mediator is produced via quark annihilation ( $g_q$ ) it can also decay to quarks and hence the dijet resonance search is sensitive to it. Figure 52 shows for a specific benchmark model and two different coupling sets (see figure caption) ATLAS exclusion regions in the DM versus the model’s mediator mass plane as obtained from the jets +  $E_T^{\text{miss}}$  and photon +  $E_T^{\text{miss}}$  analyses as well as from the dijet resonance search. These searches have complementary sensitivity.

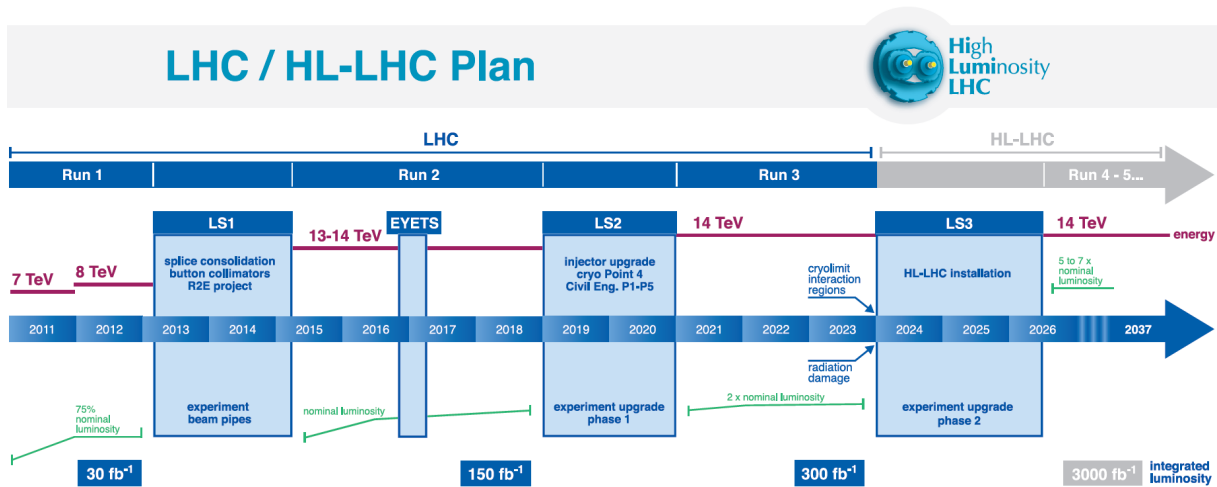


**Fig. 50:** Graph for WIMP pair production with initial-state radiation jet.

Finally, we note that even in the case of a signal in one of the LHC WIMP searches the LHC may not be able to prove that a signal is indeed dark matter because of insufficient constraints on the lifetime of the detected WIMPs.

### 9 The road to the future

The LHC experimental programme follows a well-defined suit of data taking periods followed by longer technical stops used to repair and upgrade the accelerator and experiments. With the approval of the HL-LHC project by the CERN Council in 2016 a roadmap for twenty more exciting years of physics with the LHC has been established. That roadmap is sketched in Fig. 53. The current Run-2 will continue until end of 2018 with a delivered integrated luminosity at 13 TeV (or higher) that may reach  $120 \text{ fb}^{-1}$ . The following two year long shutdown (LS2) will be used to upgrade the injector for an increased beam brightness (batch compression in the PS, new optics in the SPS, collimator upgrades). Also the experiments upgrade their detectors to prepare for the increased Run-3 luminosity. The following data taking period between 2021 and 2023 should allow the LHC to deliver a total of  $300 \text{ fb}^{-1}$  at 13–14 TeV proton–proton centre-of-mass energy. This is followed by the major HL-LHC

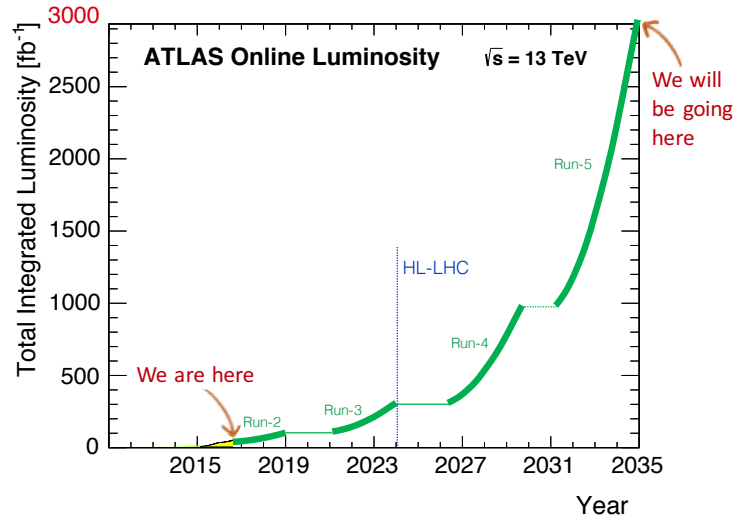


**Fig. 53:** Timeline of the LHC programme up to the high-luminosity LHC (HL-LHC).

upgrade during 2024 until 2026, featuring a new LHC triplet design (low- $\beta^*$  quadrupoles, crab cavities), and injector upgrades for luminosity levelling [197]. Here, also the experiments will undergo major upgrades to prepare for the high-luminosity phase [198, 199]. Collisions are expected to resume in 2026 allowing to deliver to each experiment (ATLAS and CMS)  $300 \text{ fb}^{-1}$  per year. The following table summarises some of the LHC beam parameters during Run-1, Run-2, and as expected for Run-3 and the HL-LHC.<sup>25</sup>

| Parameter  | LHC Run-1           | LHC Run-2 & 3                   | HL-LHC                       |
|--|---------------------|---------------------------------|------------------------------|
| Beam energy [TeV]  | 0.45–4              | 6.5–7                           | 7                            |
| Peak inst. luminosity [ $\text{cm}^{-2} \text{s}^{-1}$ ] | $0.8 \cdot 10^{34}$ | $(0.7\text{--}2) \cdot 10^{34}$ | $5 \cdot 10^{34}$ (levelled) |
| Bunch distance [ns]                                      | 50                  | 25                              | 25                           |
| Max. number of bunches                                   | 1380                | 2028~2748                       | 2748                         |
| $\beta^*$ [cm]   | 60                  | 40                              | 15                           |
| $\varepsilon_n$ [ $\mu\text{m}$ ]                        | 2.3                 | 2.5–3.5 (2.3 with BCMS)         | 2.5                          |
| Max. num. protons per bunch                              | $1.7 \cdot 10^{11}$ | $1.2 \cdot 10^{11}$             | $2.2 \cdot 10^{11}$          |
| Average pileup ( $\mu$ )                                 | 21                  | 21~50                           | 140                          |

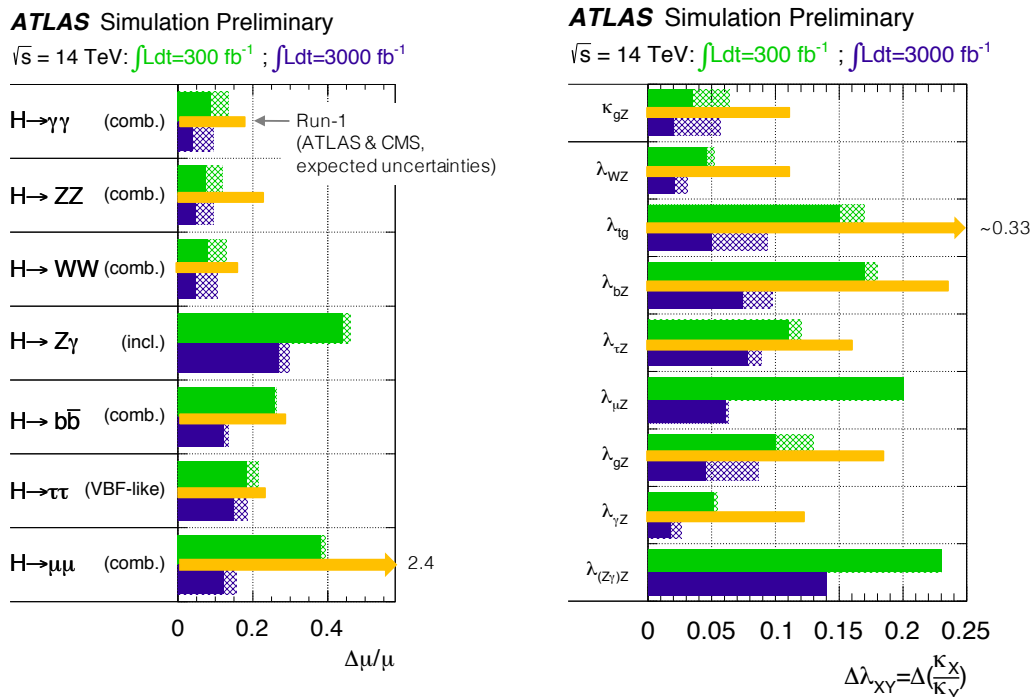
If one wants to succinctly highlight the main physics results of the LHC proton–proton programme during Run-1, one should emphasise the discovery of the Higgs boson, searches for additional new physics (all negative), multiple SM measurements, the observation of rare processes such as  $B_s \rightarrow \mu\mu$ , precision measurements of SM processes and parameters, and the study of CP asymmetries in the  $B_s$  sector. For Run-2 and Run-3, the focus lies on searches for new physics at the energy frontier, improved measurements of Higgs couplings in the main Higgs boson channels, consolidation and observation of the remaining Higgs decay and production modes, measurements of rare SM processes and more precision, improved measurements of rare  $B$  decays and CP asymmetries. Finally, the HL-LHC will serve for precision measurements of Higgs couplings, the search for and observation of very rare Higgs modes (among these di-Higgs production), the ultimate new physics search reach (on mass and forbidden decays such as FCNC), and ultimate SM and heavy flavour physics precision for rare processes (VBS, aT/QGC, etc.). Although any new physics found along the way would likely be a game changer in this planning process, these physics goals are “must do” topics for the HL-LHC.



**Fig. 54:** Sketch illustrating the integrated luminosity evolution during the various LHC phases [200]. LHC physics will hardly look the same again.

The substantial increase in luminosity will pose major technical challenges for the experiments. The average pileup will rise to  $\langle\mu\rangle = 140$  inelastic collisions per bunch crossing at (levelled)  $5 \cdot 10^{34} \text{ cm}^{-2} \text{s}^{-1}$ ,

<sup>25</sup>Recall that  $L \propto (\sigma_x \sigma_y)^{-1} = (\varepsilon_n \beta^* / \gamma)^{-1}$



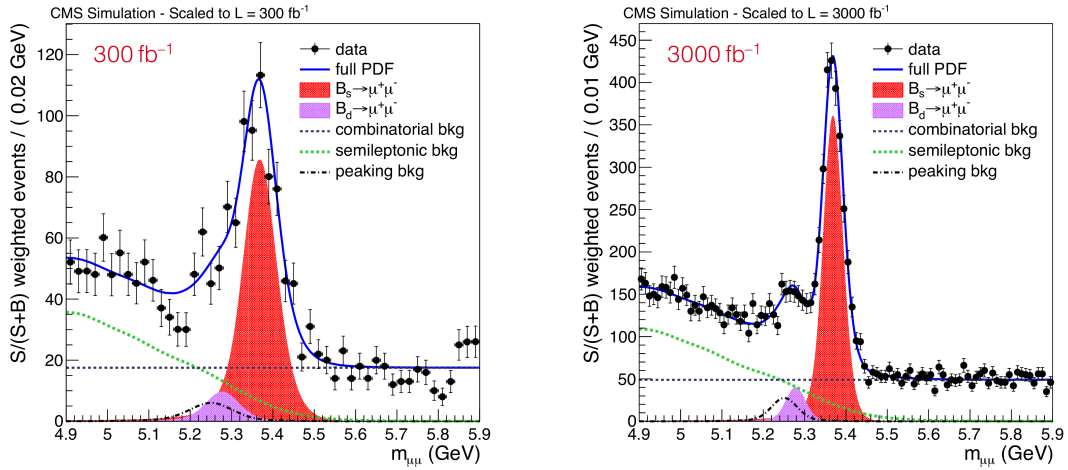
**Fig. 55:** Current (orange) and prospects for future precision (green for Run-3, blue for HL-LHC) on the measurements of the Higgs signal strengths (left panel) and the coupling modifier ratios (right). Hatched areas indicate the impact of theoretical uncertainties on expected cross-sections [201]. (The original figures have been modified.)

which will increase the background levels, the average event size and the time it takes to reconstruct the events (dominated by the track reconstruction). Faster detectors and readout electronics, as well as more sophisticated trigger systems will be required to efficiently identify physics signatures while keeping the transverse momentum thresholds at the current level. Finally, the detectors will need to withstand substantial radiation dose. Ambitious and costly upgrade programmes of the experiments address these challenges by improving the trigger and data acquisition systems, the front-end electronics, entirely replacing the inner tracking system (thereby increasing the tracker acceptance), and, in case of CMS, the endcap calorimeter, and more.

Among the large amount of prospective studies for the physics potential of the HL-LHC (and compared to the Run-3 integrated luminosity of  $300 \text{ fb}^{-1}$ ) I would like to mention here the prospects for Higgs coupling measurements, the constraint on the Higgs width from Higgs off-shell coupling measurements, and a precision measurement of  $B_s \rightarrow \mu\mu$ .

From a rather conservative extrapolation of the Run-1 Higgs coupling measurements ATLAS has derived the prospects shown in Fig. 55 assuming SM central values for the couplings. As a reminder, the coupling modifiers are defined by  $\kappa_i^2 = \sigma_i/\sigma_i^{\text{SM}}$  and  $\lambda_{ij} = \kappa_i/\kappa_j$ . The best precision of a few percent on the relative Higgs signal strengths is obtained for the diphoton, four-lepton and  $2\ell 2\nu$  decay channels. The decays to  $\tau\tau$  and  $b\bar{b}$  are challenging and will be limited by systematic uncertainties. The rare decays to  $Z\gamma$  and  $\mu\mu$  will have been observed and be statistically limited. The coupling modifier ratios (cf. right panel of Fig. 55) show a similar pattern. The important Higgs–top to Higgs–gluon coupling ratios are expected to be measured with a precision reaching 5% at the HL-LHC. Theory uncertainties are limiting the achievable precision in several cases. Some of the uncertainties cancel in the coupling modifier ratios.

Both CMS and ATLAS have constrained the Higgs off-shell coupling in Run-1 analyses and through



**Fig. 56:** Expected invariant mass distribution in the measurement of  $B_{(s)} \rightarrow \mu\mu$  for  $300 \text{ fb}^{-1}$  and  $3000 \text{ fb}^{-1}$  from a prospective study by CMS [206].

this obtained upper limits on the Higgs total width  $\Gamma_H$  [202, 203]. The method uses the independence of the off-shell cross section on  $\Gamma_H$  and relies on the assumption of identical on-shell and off-shell Higgs couplings.<sup>26</sup> One can then determine  $\Gamma_H$  ( $=4.1 \text{ MeV}$  in SM [204]) from the measurements of the off-shell and on-shell signal strengths  $\mu_{\text{off-shell}}$  and  $\mu_{\text{on-shell}}$  as follows:

$$\mu_{\text{off-shell}}(\hat{s}) = \frac{\sigma_{\text{off-shell}}^{gg \rightarrow H^* \rightarrow VV}(\hat{s})}{\sigma_{\text{off-shell,SM}}^{gg \rightarrow H^* \rightarrow VV}(\hat{s})} = \kappa_{g,\text{off-shell}}^2(\hat{s}) \cdot \kappa_{V,\text{off-shell}}^2(\hat{s}),$$

$$\mu_{\text{on-shell}} = \frac{\sigma_{\text{on-shell}}^{gg \rightarrow H \rightarrow VV^*}}{\sigma_{\text{on-shell,SM}}^{gg \rightarrow H \rightarrow VV^*}} = \frac{\kappa_{g,\text{off-shell}}^2(\hat{s}) \cdot \kappa_{V,\text{off-shell}}^2(\hat{s})}{\Gamma_H / \Gamma_{H,\text{SM}}}.$$

With the Run-1 datasets, limits of the order of 5 times  $\Gamma_{H,\text{SM}}$  were obtained by ATLAS and CMS. An ATLAS HL-LHC study [205] derived prospects for integrated luminosities of  $300 \text{ fb}^{-1}$  and  $3000 \text{ fb}^{-1}$  giving  $\mu_{\text{off-shell}} = 1_{-0.97}^{+0.80}$  and  $1_{-0.50}^{+0.43}$ , respectively. The latter precision allows to constrain  $\Gamma_H$  to remarkable  $4.1_{-2.1}^{+1.5} \text{ MeV}$ .

In the area of new physics searches, the emphasis will gradually move towards rare and difficult channels, such as low cross-section electroweak production and compressed scenarios in SUSY. Searches for WIMPs will require improvements in the data-driven determination of the backgrounds to take full benefit from the increased data sample.

Among the many other interesting prospects, one should also note the continuous gain in precision and reach for rare or suppressed processes in the flavour sector. The rise in luminosity during Run-2 will be slower for LHCb due to the luminosity levelling. The upgrade to 40 MHz trigger readout during the long shutdown 2 in 2019 will help increase the annual muonic  $B$  rate by a factor of ten. High-profile rare decay measurements performed by LHCb, ATLAS and CMS are  $B_{(s)} \rightarrow \mu\mu$  (and similar) as well as  $b \rightarrow s$  transitions such as  $B \rightarrow K^* \mu\mu$  and similar modes. Figure 56 shows the invariant mass distribution for  $B_{(s)} \rightarrow \mu\mu$  as expected from a prospective study by CMS [206]. The observation beyond  $5\sigma$  significance of the loop and CKM suppressed decay  $B \rightarrow \mu\mu$  is expected for the full HL-LHC integrated luminosity.

<sup>26</sup>The denominator of a relativistic Breit-Wigner resonance lineshape has the form  $(s - m^2)^2 + s^2 \Gamma^2 / m^2$ . For  $s \sim m^2$  (on-shell) the first term in the denominator vanishes so that the coupling depends reciprocally on the width  $\Gamma$ . In the off-shell regime  $s \gg m^2$  the first term dominates and the  $\Gamma$  dependence becomes negligible.

CP-violation measurements of the phase  $\phi_s$  will be performed by LHCb (dominant) and also by ATLAS and CMS, the unitarity triangle angle  $\gamma$  and other CKM parameters will be measured by LHCb. These important measurements will benefit from any increase in integrated luminosity. LHCb will also improve CP asymmetry measurements in the charm sector. Of high importance given the current results is to pursue measurements testing lepton universality in  $B$  decays (LHCb and Belle). Finally, further surprises and a better understanding of recently discovered heavy flavour spectroscopy states are expected by LHC, ATLAS and CMS.

## 10 Conclusions

The LHC Run-2 is a key period for particle physics. The first  $100 \text{ fb}^{-1}$  at  $13 \text{ TV}$  centre-of-mass energy are critical for new physics searches in all signatures. Further consolidation of the Higgs sector with the observation and measurement of  $H \rightarrow \tau\tau$ ,  $H \rightarrow bb$ , and associated  $t\bar{t}H$  production, as well as more precise coupling, fiducial and differential cross section measurements will be followed up with high priority by ATLAS and CMS. The luminosity of Run-2 will hugely increase the amount of interesting Standard Model and flavour physics measurements that can be performed.

Throughout Run-2 it is important to stay alert. New physics does not necessarily appear at high mass so that one needs to continue to search everywhere. High precision measurements are key for a better knowledge of the Standard Model. It is thereby extremely important to measure the detector performance in data as precisely as possible, and this may have priority over further improving the performance. Many results are dominated by theoretical uncertainties. The experiments need to produce measurements that allow to test theory, to improve PDFs, and that motivate theorists to improve calculations and event generators. We may cite William Thomson Kelvin, from a speech held to the British Association for the Advancement of Science in 2 Aug 1871: “Accurate and minute measurement seems to the non-scientific imagination, a less lofty and dignified work than looking for something new. But [many of] the grandest discoveries of science have been but the rewards of accurate measurement and patient long-continued labour in the minute sifting of numerical results.”

*I thank the organisers of the 2016 European School of High-Energy Physics for giving me the opportunity to lecture at this excellent school at a very pleasant location in Norway.*

## References

- [1] F. Englert and R. Brout, Phys. Rev. Lett. 13, 321 (1964); P.W. Higgs, Phys. Lett. 12, 132 (1964); P. W. Higgs, Phys. Rev. Lett. 13, 508 (1964); G.S. Guralnik, C.R. Hagen, and T.W.B. Kibble, Phys. Rev. Lett. 13, 585 (1964); P.W. Higgs, Phys. Rev. 145, 1156 (1966); T.W.B. Kibble, Phys. Rev. 155, 1554 (1967).
- [2] W. Bernreuther, Lect. Notes Phys. 591, 237 (2002) [hep-ph/0205279].
- [3] H1 and ZEUS Collaborations Eur. Phys. J. C 75, 580 (2015) [1506.06042].
- [4] ALEPH, DELPHI, L3, OPAL, SLD Collaborations, LEP Electroweak Working Group, SLD Electroweak and Heavy Flavour Groups, Phys. Rept. 427, 257 (2006) [hep-ex/0509008]; ALEPH, DELPHI, L3, OPAL Collaborations, LEP Electroweak Working Group, Phys. Rept. 532, 119 (2013) [1302.3415].
- [5] L. Evans and P. Bryant, JINST 3 (2008) S08001, ed. by L. Evans.
- [6] LHC luminosity calibration tool: <https://lpc.web.cern.ch/lumi2.html>.
- [7] ATLAS Collaboration, 1608.03953 (2016).
- [8] CMS Collaboration, CMS-PAS-LUM-13-001 (2013), CMS-PAS-LUM-15-001 (2016).

- [9] LHCb Collaboration, JINST 9 no. 12, 12005 (2014) [1410.0149].
- [10] S. van der Meer, CERN-ISR-PO-68-31, ISR-PO-68-31 (1968).
- [11] J. Stirling, Parton luminosity plots.
- [12] J. Collins, *Foundations of Perturbative QCD*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, 2011.
- [13] V.N. Gribov and L.N. Lipatov. Sov. J. Nucl. Phys. 15, 438 (1972); Y.L. Dokshitzer. Sov. Phys. JETP 46, 641 (1977); G. Altarelli and G. Parisi. Nucl. Phys. B 126, 298 (1977).
- [14] S. Agostinelli et al., Nucl. Instrum. Meth. A 506, 250 (2003), <http://cern.ch/geant4>.
- [15] G. Salam, Talk at LHCP 2016, Lund, Sweden, June 2016.
- [16] ATLAS Collaboration, ATL-PHYS-PUB-2015-013 (2015).
- [17] M. Cacciari, G.P. Salam, G. Soyez, JHEP 0804, 063 (2008) [0802.1189].
- [18] ATLAS Collaboration, JHEP 1309, 076 (2013) [1306.4945].
- [19] CMS Collaboration, JHEP 12, 017 (2014) [1410.4227].
- [20] CMS Collaboration, 1609.05331 (2016).
- [21] ATLAS Collaboration, Phys. Rev. Lett. 113, 141803 (2014) [1405.6241]; 1611.02428 (2016);
- [22] CMS Collaboration, Phys. Rev. Lett. 114, 051801 (2015) [1410.6315].
- [23] ATLAS Collaboration, JHEP 11, 172 (2015) [1509.05276].
- [24] CMS Collaboration, JHEP 01, 096 (2016) [1510.01131].
- [25] ATLAS Collaboration, Phys. Rev. D. 90, 112006 (2014) [1406.7844]; ATLAS-CONF-2014-007 (2014).
- [26] CMS Collaboration, JHEP 06, 090 (2014) [1403.7366].
- [27] ATLAS Collaboration, JHEP 1601, 064 (2016) [1510.03752].
- [28] CMS Collaboration, Phys. Rev. Lett. 112, 231802 (2014) [1401.2942].
- [29] ATLAS Collaboration, Phys. Lett. B 756, 228 (2016) [1511.05980].
- [30] CDF and D0 Collaborations, Phys. Rev. Lett. 112, 231803 (2014) [1402.5126].
- [31] ATLAS Collaboration, Phys. Lett. B 716, 1 (2012) [1207.7214].
- [32] CMS Collaboration, Phys. Lett. B 716, 30 (2012) [1207.7235].
- [33] The ALEPH, DELPHI, L3 and OPAL Collaborations, and LEP Working Group for Higgs Boson Searches, Phys. Lett. B 565, 61 (2003) [hep-ex/0306033].
- [34] CDF and D0 Collaborations, Phys. Rev. D 88, 052014 (2013) [1303.6346].
- [35] ALEPH, CDF, D0, DELPHI, L3, OPAL and SLD Collaborations, LEP Electroweak Working Group, Tevatron Electroweak Working Group, SLD Electroweak and Heavy Flavour Groups 1012.2367; J. Erler and P. Langacker (in: Review for Particle Data Group), Phys. Rev. D 86, 010001 (2012); M. Baak et al., Eur. Phys. J. C 72, 2003 (2012) [1107.0975].
- [36] LHC Higgs Cross Section Working Group, 1610.07922 (2016).
- [37] M. Carena, C. Grojean, and M. Kado, Review in PDG, Chin. Phys. C38, 090001 (2014).
- [38] ATLAS Collaboration, Phys. Rev. D 90, 112015 (2014) [1408.7084].
- [39] ATLAS Collaboration, Phys. Rev. D 91, 012006 (2015) [1408.5191].
- [40] ATLAS Collaboration, Phys. Rev. D 92, 012006 (2015) [1412.2641]; JHEP 08, 137 (2015) [1506.06641].
- [41] CMS Collaboration, Eur. Phys. J. C 74, 3076 (2014) [1407.0558].
- [42] CMS Collaboration, Phys. Rev. D 89, 092007 (2014) [1312.5353].
- [43] CMS Collaboration, JHEP 01, 096 (2014) [1312.1129].
- [44] ATLAS and CMS Collaborations, Phys. Rev. Lett. 114, 191803 (2015) [1503.07589].
- [45] ATLAS and CMS Collaborations, JHEP 1608, 045 (2016) [1606.02266].

- [46] S. Heinemeyer et al., 1307.1347 (2013).
- [47] ATLAS Collaboration, JHEP 04, 117 (2015) [1501.04943].
- [48] CMS Collaboration, JHEP 05, 104 (2014) [1401.5041].
- [49] ATLAS Collaboration, Phys. Lett. B 740, 222 (2015) [1409.3122]; Eur. Phys. J. C 75, 349 (2015) [1503.05066]; Phys. Lett. B 749, 519 (2015) [1506.05988].
- [50] CMS Collaboration, JHEP 05, 145 (2013) [1303.0763]; JHEP 09, 087 (2014) [1408.1682].
- [51] ATLAS Collaboration, JHEP 1601, 172 (2016) [1508.07869].
- [52] CMS Collaboration, Eur. Phys. J. C 74, 2980 (2014) [1404.1344]; 1610.09218 (2016).
- [53] ATLAS Collaboration, JHEP 1511, 206 (2015) [1509.00672].
- [54] CMS and LHCb Collaborations, Nature 522, 68 (2015) [1411.4413].
- [55] C. Bobeth et al., Phys. Rev. Lett. 112, 101801 (2014) [1311.0903].
- [56] ATLAS Collaboration, Eur. Phys. J. C 76, 513 (2016) [1604.04263].
- [57] ATLAS Collaboration, JHEP 08, 147 (2016) [1601.03297].
- [58] CMS Collaboration, Phys. Lett. B 757, 97 (2016) [1507.07527].
- [59] LHCb Collaboration, Phys. Rev. Lett. 114, 041801 (2015) [1411.3104].
- [60] LHCb Collaboration, LHCb-CONF-2016-001 (2016).
- [61] CKMfitter Group, updates at <http://ckmfitter.in2p3.fr> (2015).
- [62] LHCb Collaboration, Nature Physics 10, 1038 (2015) [1504.01568].
- [63] LHCb Collaboration, Eur.Phys.J. C 76, 412 (2016) [1604.03475].
- [64] LHCb Collaboration, Phys. Rev. Lett. 115, 031601 (2015) [1503.07089].
- [65] UA1 Collaboration, Europhys. Lett. 1, 327 (1986).
- [66] UA2 Collaboration, Phys. Lett. B 276, 354 (1992).
- [67] LHCb Collaboration, Phys.Rev.Lett. 116, 241601 (2016) [1603.04804].
- [68] R.N. Cahn, Rev. Mod. Phys. 68, 951 (1996).
- [69] BABAR Collaboration, Phys. Rev. Lett. 109, 101802 (2012) [1205.5442].
- [70] Belle Collaboration, Phys. Rev. D 92, 072014 (2015) [1507.03233].
- [71] LHCb Collaboration, Phys. Rev. Lett. 115, 111803 (2015) [1506.08614].
- [72] Belle Collaboration, 1603.06711 (2016).
- [73] Heavy Flavour Averaging Group,  $R_{D^{(*)}}$  combination (2016).
- [74] S. Fajfer, J.F. Kamenik, and I. Nisandzic, Phys. Rev. D 85, 094025 (2012) [1203.2654].
- [75] LHCb Collaboration, Phys. Rev. Lett. 113, 151601 (2014) [1406.6482].
- [76] LHCb Collaboration, Phys. Rev. Lett. 115, 072001 (2015) [1507.03414].
- [77] M.E. Peskin and T. Takeuchi, Phys. Rev. D 46, 381 (1992).
- [78] J.R. Ellis, G.L. Fogli and E. Lisi, Phys. Lett. B 343, 282 (1995).
- [79] M. Baak et al., Eur. Phys. J. C 72, 2003 (2012) [1107.0975].
- [80] M. Baak *et al.*, Eur. Phys. J. C 74, 3046 (2014) [1407.3792].
- [81] CDF and D0 Collaborations, Phys. Rev. D 88, 052018 (2013) [1307.7627].
- [82] CMS Collaboration, Phys. Rev. D 93, 072004 (2016) [1509.04044].
- [83] ATLAS Collaboration, Phys. Lett. B 761, 350 (2016) [1606.02179].
- [84] CDF and D0 Collaborations, 1407.2682 (2014).
- [85] CMS Collaboration, Eur. Phys. J. C 73, 2494 (2013) [1304.5783].
- [86] CMS Collaboration, CMS-PAS-TOP-15-014 (2015).
- [87] CMS Collaboration, 1603.02303 (2016).
- [88] ATLAS Collaboration, JHEP 09, 049 (2015) [1503.03709].



- [89] CMS Collaboration, Phys. Rev. D 84, 112002 (2011) [1110.2682].
- [90] LHCb Collaboration, JHEP 1511, 190 (2015) [1509.07645].
- [91] A. Bodek, Eur. Phys. J. C, 76, 1 (2016) [1512.08256].
- [92] CMS Collaboration, CMS-PAS-SMP-14-007 (2016).
- [93] ATLAS Collaboration, Eur. Phys. J. C 76, 1 (2016) [1512.02192].
- [94] C. Balazs, J.W. Qiu and C. Yuan, Phys. Lett. B 355, 548 (1995) [hep-ph/9505203].
- [95] S. Catani et al., Phys. Rev. Lett. 103, 082001 (2009) [0903.2120].
- [96] Gfitter Group (M. Baak et al.), Eur. Phys. J. C 72, 2205 (2012) [1209.2716].
- [97] J. Ellis et al., Phys.Lett. B 679, 369 (2009) [0906.0954].
- [98] B. Bertotti, L. Iess, P. Tortora, Nature 425, 374 (2003).
- [99] M. Mangano, Talk at SEARCH 2016, Oxford, UK, August 2016.
- [100] ATLAS Collaboration, Phys. Lett. B, 158 (2016) [1607.06605].
- [101] TOTEM Collaboration, Phys. Rev. Lett. 111, 012001 (2013).
- [102] ATLAS Collaboration, 1606.02625 (2016).
- [103] CMS Collaboration, Eur. Phys. J. C 76, 451 (2016) [1605.04436].
- [104] ATLAS Collaboration, Eur. Phys. J. C 72, 1909 (2012) [1110.3174].
- [105] ATLAS Collaboration, Phys.Lett. B 759, 601 (2016) [1603.09222].
- [106] CMS Collaboration, CMS-PAS-SMP-15-004 (2015).
- [107] LHCb Collaboration, LHCb-CONF-2016-002 (2016).
- [108] ATLAS Collaboration, Phys. Rev. Lett. 109, 012001 (2012) [1203.4051].
- [109] ATLAS Collaboration, Phys. Rev. Lett. 116, (2016) [1512.05314].
- [110] CMS Collaboration, CMS-PAS-SMP-16-006 (2016).
- [111] ATLAS Collaboration, Phys. Lett. B 762, 1 (2016) [1606.04017].
- [112] ALEPH, DELPHI, L3 and OPAL Collaborations, LEP Electroweak Working Group, Phys.Rept. 532, 119 (2013) [1302.3415].
- [113] D.R. Green, P. Meade, M.-A. Pleier, 1610.07572 (2016).
- [114] ATLAS Collaboration, Eur. Phys. J. C 74, 3109 (2014) [1406.5375].
- [115] ATLAS Collaboration, Phys. Lett. B 761, 136 (2016) [1606.02699].
- [116] J. Wenninger, CERN-ATS-2013-040 (2013).
- [117] ATLAS Collaboration, 1609.03920 (2016).
- [118] CMS Collaboration, 1610.00678 (2016).
- [119] ATLAS Collaboration, ATLAS-CONF-2016-065 (2015).
- [120] ATLAS Collaboration, 1609.01599 (2016).
- [121] CMS Collaboration, CMS-PAS-TOP-16-017 (2016).
- [122] ATLAS Collaboration, ATLAS-CONF-2016-079 (2016).
- [123] CMS Collaboration, CMS-PAS-HIG-16-033 (2016).
- [124] ATLAS Collaboration, ATLAS-CONF-2016-067 (2016).
- [125] CMS Collaboration, CMS-PAS-HIG-16-020 (2016).
- [126] ATLAS Collaboration, ATLAS-CONF-2016-081 (2016).
- [127] ATLAS Collaboration, ATLAS-CONF-2016-091 (2016).
- [128] ATLAS Collaboration, JHEP 01, 069 (2015) [1409.6212].
- [129] CMS Collaboration, Phys. Rev. D 89, 012003 (2014) [1310.3687].
- [130] ATLAS Collaboration, ATLAS-CONF-2016-041 (2016).
- [131] CMS Collaboration, CMS-PAS-HIG-16-022 (2016).

- [132] CMS Collaboration, CMS-PAS HIG-15-005 (2016).
- [133] CMS Collaboration, CMS-PAS HIG-16-004 (2016).
- [134] ATLAS Collaboration, ATLAS-CONF-2016-058 (2016); ATLAS-CONF-2016-067 (2016); ATLAS-CONF-2016-080 (2016).
- [135] ATLAS Collaboration, ATLAS-CONF-2016-06 (2016).
- [136] ATLAS Collaboration, ATLAS-CONF-2016-088 (2016).
- [137] CMS Collaboration, CMS-PAS-HIG-16-031 (2016).
- [138] ATLAS Collaboration, ATLAS-CONF-2016-085 (2016).
- [139] CMS Collaboration,
- [140] ATLAS Collaboration, ATLAS-CONF-2016-015 (2016).
- [141] ATLAS Collaboration, Phys. Rev. Lett. 114, 081802 (2015) [1406.5053].
- [142] ATLAS Collaboration, ATLAS-CONF-2016-004 (2016).
- [143] CMS Collaboration, CMS-PAS-HIG-16-032 (2016).
- [144] CMS Collaboration, CMS-PAS-HIG-16-029 (2016).
- [145] CMS Collaboration, CMS-PAS-HIG-16-024 (2016).
- [146] ATLAS Collaboration, Phys. Rev. D 94, 052002 (2016) [1606.04782].
- [147] CMS Collaboration, CMS-PAS-HIG-16-026 (2016).
- [148] CMS Collaboration, Phys. Lett. B 749, 337 (2015) [1502.07400].
- [149] ATLAS Collaboration, 1604.07730 (2016).
- [150] CMS Collaboration, CMS-PAS-HIG-16-005 (2016).
- [151] ATLAS Collaboration, ATLAS-CONF-2016-069 (2016).
- [152] ATLAS Collaboration, ATLAS-CONF-2016-070 (2016).
- [153] CMS Collaboration, 1611.03568 (2016).
- [154] ATLAS Collaboration, ATLAS-CONF-2016-030 (2016).
- [155] CMS Collaboration, Phys. Rev. Lett. 117, 031802 (2016) [1604.08907].
- [156] ATLAS Collaboration, 1603.08791 (2016); ATLAS-CONF-2016-014 (2016); ATLAS-CONF-2016-060 (2016).
- [157] CMS Collaboration, CMS-PAS-B2G-15-002 (2016).
- [158] ATLAS Collaboration, ATLAS-CONF-2016-045 (2016).
- [159] CMS Collaboration, CMS-PAS-EXO-16-031 (2016).
- [160] ATLAS Collaboration, ATLAS-CONF-2016-061 (2016).
- [161] CMS Collaboration, CMS-PAS-EXO-15-006 (2016).
- [162] ATLAS Collaboration, ATL-PHYS-PUB-2013-003 (2013); ATL-PHYS-PUB-2015-004 (2015)
- [163] H. Bachacou, Table shown at SEARCH2016 workshop, Oxford, UK (Sep 2016).
- [164] ATLAS Collaboration, JHEP 12, 55 (2015) [1506.00962].
- [165] CMS Collaboration, JHEP 08, 173 (2014) [1405.1994].
- [166] ATLAS Collaboration, ATLAS-CONF-2016-05 (2016).
- [167] CMS Collaboration, CMS-PAS-B2G-16-020 (2016); CMS-PAS-B2G-16-007 (2016).
- [168] ATLAS Collaboration, Phys. Rev. D 92, 032004 (2015) [1504.05511].
- [169] ATLAS Collaboration, Phys. Rev. Lett. 113, 171801 (2014) [1407.6583].
- [170] CMS Collaboration, Phys. Lett. B 750, 494 (2015) [1506.02301].
- [171] ATLAS Collaboration, Eur. Phys. J. C 76, 1 (2016) [1509.05051].
- [172] ATLAS Collaboration, 1606.03833 (2016).
- [173] CMS Collaboration, 1606.04093 (2016).

- [174] ATLAS Collaboration, ATLAS-CONF-2016-059 (2016).
- [175] CMS Collaboration, 1609.02507 (2016).
- [176] ATLAS Collaboration, JHEP 10, 134 (2015) [1508.06608].
- [177] CMS Collaboration, JHEP 10, 129 (2016) [1606.03577].
- [178] CMS Collaboration, CMS-PAS-SUS-16-014 (2016).
- [179] CMS Collaboration, CMS-PAS-SUS-16-015 (2016).
- [180] ATLAS Collaboration, ATLAS-CONF-2016-052 (2016).
- [181] J. Barnard, P. Cox, T. Gherghetta, A. Spray, JHEP 1603, 003 (2016) [1510.06405].
- [182] ATLAS Collaboration, Phys. Rev. D 88, 112006 (2013) [1310.3675].
- [183] ATLAS Collaboration, Phys. Rev. D 88, 112003 (2013) [1310.6584].
- [184] CMS Collaboration, Eur. Phys. J. C 75, 151 (2015) [1501.05603].
- [185] LHC Dark Matter Forum, 1507.00966 (2015).
- [186] ATLAS Collaboration, Phys.Rev. D 94, 032005 (2016) [1604.07773].
- [187] CMS Collaboration, CMS-PAS-EXO-16-037 (2016).
- [188] ATLAS Collaboration, JHEP 06, 059 (2016) [1604.01306].
- [189] CMS Collaboration, CMS-PAS-EXO-16-039 (2016).
- [190] ATLAS Collaboration, Phys. Lett. B 763, 251 (2016) [1608.02372].
- [191] CMS Collaboration, CMS-PAS-EXO-16-038 (2016); CMS-PAS-EXO-16-037 (2016); CMS-PAS-EXO-16-010 (2016).
- [192] ATLAS Collaboration, ATLAS-CONF-2016-086 (2016).
- [193] CMS Collaboration, CMS-PAS-B2G-15-007 (2016).
- [194] CMS Collaboration, CMS-PAS-EXO-16-028 (2016); CMS-PAS-EXO-16-005 (2016).
- [195] LUX Collaboration, Phys. Rev. Lett. 116, 161302 (2016) [1602.03489].
- [196] ATLAS Collaboration, Dark Matter searches summary plot.
- [197] The HI-LHC project: <http://hilumilhc.web.cern.ch>.
- [198] ATLAS Collaboration, CERN-LHCC-2015-020 (2015).
- [199] CMS Collaboration, CERN-LHCC-2015-019 (2015).
- [200] P. Ferreira da Silva at Moriond EW, 2016.
- [201] ATLAS Collaboration, ATL-PHYS-PUB-2014-016 (2014).
- [202] ATLAS Collaboration, Eur. Phys. J. C 75, 335 (2015) [1503.01060].
- [203] CMS Collaboration, JHEP 09, 051 (2016) [1605.02329].
- [204] LHC Higgs Cross Section Working Group, CERN-2013-004, 1307.1347 (2013).
- [205] ATLAS Collaboration, ATL-PHYS-PUB-2015-024 (2015).
- [206] CMS Collaboration, CMS-PAS-FTR-13-022 (2013).



# Practical Statistics for Particle Physicists

*L. Lista*

Istituto Nazionale di Fisica Nucleare, Sezione di Napoli, Italy

## Abstract

These three lectures provide an introduction to the main concepts of statistical data analysis useful for precision measurements and searches for new signals in High Energy Physics. The frequentist and Bayesian approaches to probability theory are introduced and, for both approaches, inference methods are presented. Hypothesis tests will be discussed, then significance and upper limit evaluation will be presented with an overview of the modern and most advanced techniques adopted for data analysis at the Large Hadron Collider.

## Keywords

Lectures; statistics; probability; frequentist; bayesian; statistical analysis.

## 1 Introduction

The main goal of an experimental particle physicist is to make precision measurements and possibly discover new natural phenomena. The starting ingredients to this task are particle collisions that are recorded in form of data delivered by detectors. Data provide measurements of the position of particle trajectories or energy releases in the detector, time of particles arrival, etc. Usually, a large number of collision events are collected by an experiment and each of such events may contain large amounts of data. Collision event data are all different from each other due to the intrinsic randomness of physics process. In Quantum Mechanics the probability (density) is proportional to the square of the process amplitude ( $P \propto |A|^2$ ). Detectors also introduce some degree of randomness in the data due to fluctuation of the response, like resolution effects, efficiency, etc. Theory provides prediction of the distributions of measured quantities in data. Those predictions depend on theory parameters, such as particles masses, particle couplings, cross section of observed processes, etc.

Given our data sample, we want to either measure the parameters that appear in the theory (e.g.: determine the top-quark mass to be:  $m_t = 173.44 \pm 0.49$  GeV [1]) or answer questions about the nature of data. For instance, as outcome of the search for the Higgs boson at the Large Hadron Collider, the presence of the boson predicted by Peter Higgs and François Englert was confirmed providing a quantitative measurement of how strong this evidence was. Modern experiments search for Dark Matter and they found no convincing evidence so far. Such searches can provide a range of parameters for theory models that predict Dark-Matter particle candidates that are allowed or excluded by the present experimental observation.

In order to achieve the methods that allow to perform the aforementioned measurements or searches for new signals, first of all a short introduction to probability theory will be given, in order to master the tools that describe the intrinsic randomness of our data. Then, methods will be introduced that allow to use probability theory on our data samples in order to address quantitatively our physics questions.

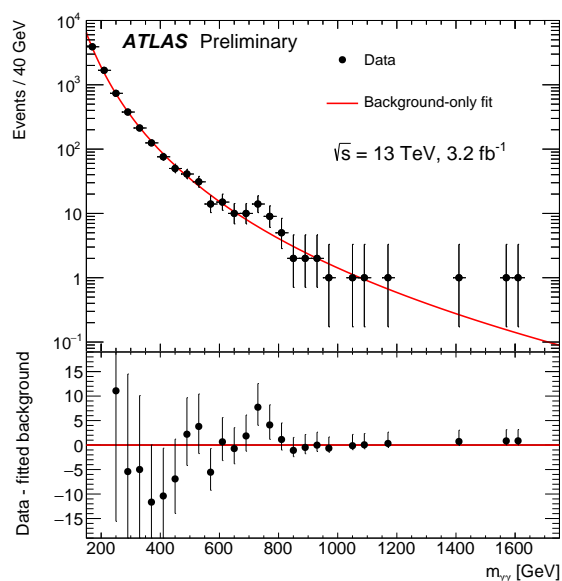
## 2 Probability theory

Probability can be defined in different ways, and the applicability of each definition depends on the kind of claim whose probability we are considering. One subjective approach expresses the *degree of belief/credibility* of a claim, which may vary from subject to subject. For repeatable experiments whose

outcome is uncertain, probability may be a measure of *how frequently* the claim is true. Repeatable experiments are a subset of the cases where the subjective approach may be applied.

Examples of probability that can be determined by means of repeatable experiments are the followings:

- *What is the probability to extract an ace in a deck of cards?*  
We can shuffle the deck and extract again the card.
- *What is the probability to win a lottery?*  
Though a specific lottery extraction can't be repeated, we can imagine to repeat the extraction process using the same device and extraction procedure.
- *What is the probability that a pion is incorrectly identified as a muon in detector capable of particle identification?*  
Most of the experiments have ways to obtain control samples where it's known that only pions are present (e.g.: test beams, or specific decay channels, etc.). One can count how many pions in a control sample are misidentified as muon.
- *What is the probability that a fluctuation in the background can produce a peak in the  $\gamma\gamma$  spectrum with a magnitude at least equal to what has been observed by ATLAS (Fig. 1, Ref. [2])?*  
At least in principle, the experiment can be repeated with the same running conditions. Anyway, this question is different with respect to another possible question: *what is the probability that the peak is due to a background fluctuation instead of a new signal?* This second question refers to a non-repeatable case.



**Fig. 1:** Invariant mass distribution of diphoton events selected by ATLAS. Figure from Ref. [2], where details about the analysis are described.

Examples of claims related to non-repeatable situations, instead, are the following:

- *What is the probability that tomorrow it will rain in Geneva?*  
The event is related to a specific date in the future. This specific event cannot be repeated.
- *What is the probability that your favorite team will win next championship?*  
Though every year there is a championship, a specific one can't be repeated.
- *What is the probability that dinosaurs went extinct because of an asteroid?*

This question is related to an event occurred in the past, but we don't know exactly what happened at that time.

- *What is the probability that Dark Matter is made of particles heavier than 1 TeV?*  
This question is related to an unknown property of our Universe.
- *What is the probability that climate changes are mainly due to human intervention?*  
This question is related to a present event whose cause is unknown.

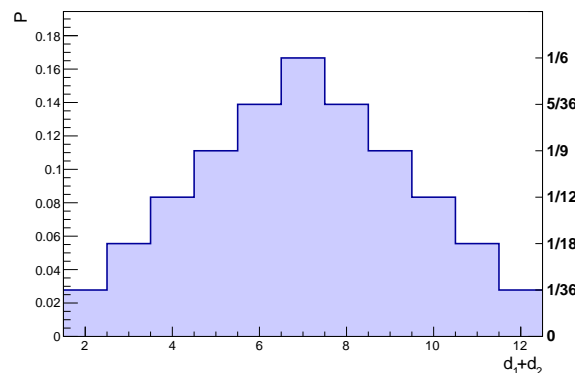
The first examples in the above list are related to events in the future, where it's rather natural to think in term of degree of belief about some prediction: we can wait and see if the prediction was true or not. But similar claims can also be related to past events, or, more in general, to cases where we just don't know whether the claims is true or not.

### 2.1 Classical probability

The simplest way to introduce probability is to consider the symmetry properties of a random device. Example could be a tossed coin (outcome may be *head* or *tail*) or a rolled dice (outcome may be a number from 1 to 6 for a cubic dice, but dices also exist with different shapes). According to the original definition due to Laplace [3], we can be "equally undecided" about an event outcome due to symmetry properties, and we can assign an equal *probability* to each of the outcomes. If we define an *event* from a statement about the possible outcome of one random extraction (e.g.: *a dice roll gives an odd number*), the *probability P* of that event (i.e.: that the statement is true) can be defined as:

$$P = \frac{\text{Number of favorable cases}}{\text{Total number of cases}} . \tag{1}$$

Probabilities related to composite cases can be computed using *combinatorial analysis* by reducing the composite event of interest into elementary equiprobable events. The set of all possible elementary events is called *sample space* (see also Sec. 2.4). For instance, in Fig. 2, the probability to obtain a given sum of two dices is reported. The computation can be done by simply counting the number of elementary cases.



**Fig. 2:** Probability distribution of the sum of two dices computed by counting all possible elementary outcomes.

For instance 2 can be obtained only as 1 + 1, while 3 can be obtained as 1 + 2 or 2 + 1, 4 as 1 + 3, 2 + 2 or 3 + 1, etc.

Textual statements about an event can be translated using set algebra considering that and/or/not correspond to intersection/union/complement in set algebra. For instance, the event "*sum of two dices is even and greater than four*" corresponds to the intersection of two sets:

$$\{(d_1, d_2) : \text{mod}(d_1 + d_2, 2) = 0\} \cap \{(d_1, d_2) : d_1 + d_2 > 4\} . \tag{2}$$

### 2.1.1 “Events” in statistics and in physics

It’s worth at this point to remarking the different meaning that the word *event* usually assumes in statistics and in physics. In statistics an event is a subset of the sample space. E.g.: “the sum of two dices is  $\geq 5$ ”. In particle physics usually an event is the result of a collision, as recorded by our experiment. In several concrete cases, an event in statistics may correspond to many possible collision events. E.g.: “ $p_T(\gamma) > 40$  GeV” may be an event in statistics, but it may correspond to many events from a data sample that have at least one photon with transverse momentum greater than 40 GeV.

## 2.2 Frequentist probability

The definition of *frequentis* probability relates probability to the fraction of times an event occurs, in the limit of very large number ( $N \rightarrow \infty$ ) of repeated trials:

$$P = \lim_{N \rightarrow \infty} \frac{\text{Number of favorable cases}}{N = \text{Number of trials}}. \quad (3)$$

This definition is exactly realizable only with an infinite number of trials, which conceptually may be unpleasant. Anyway, physicists may consider this definition pragmatically acceptable as approximately realizable in a large, but not infinite, number of cases. The definition in Eq. (3) is clearly only applicable to repeatable experiments.

## 2.3 Subjective (Bayesian) probability

Subjective probability expresses one’s *degree of belief* that a claim is true. A probability equal to 1 expresses certainty that the claim is true, 0 expresses certainty that the claim is false. Intermediate values from 0 to 1 quantify how strong the degree of belief that the claims is true is. This definition is applicable to all unknown events/claims, not only repeatable experiments, as it is the case for the frequentist approach. Each individual may have a different opinion/prejudice about one claim, so this definition is necessarily *subjective*. Anyway, quantitative rules exist about how subjective probability should be modified after learning about some observation/evidence. Those rules descend from the Bayes theorem (see Sec. 2.10), and this gives the name of *Bayesian* probability to subjective probability. Starting from a *prior probability*, following some observation, the probability can be modified into a *posterior probability*. The more information an individual receives, the more Bayesian probability is insensitive on prior probability, with the exception of pathological cases of prior probability. An example of such a case is a prior certainty that a claim is true (*dogma*) that is then falsified by the observation.

## 2.4 Komogorov axiomatic approach

An axiomatic definition of probability is due to Kolmogorov [4], which can be applied both to frequentist and Bayesian probabilities. The axioms assume that  $\Omega$  is a sample space,  $F$  is an event space made of subsets of  $\Omega$  ( $F \subseteq 2^\Omega$ ), and  $P$  is a *probability measure* that obeys the following three conditions:

1.  $P(E) \geq 0, \forall E \in F$
2.  $P(\Omega) = 1$  (normalization condition)
3.  $\forall (E_1, \dots, E_n) \in F^n : E_i \cup E_j = \emptyset, P\left(\bigcup_{i=1}^n E_i\right) = \sum_{i=1}^n P(E_i)$

The last condition states that the probability of the union of a set of disjoint events is equal to the sum of their individual probabilities.

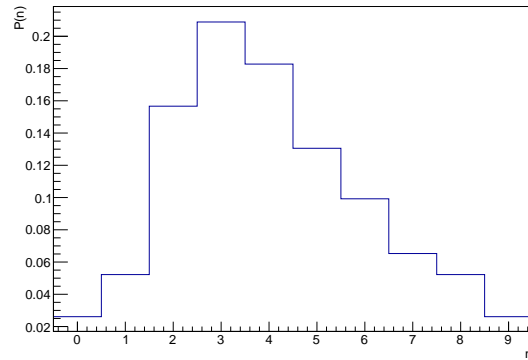
## 2.5 Probability distributions

Given a discrete *random variable*  $n$ , a probability can be assigned to each individual possible value of  $n$ :

$$P(n) = P(\{n\}). \quad (4)$$



Figure 3 shows an example of discrete probability distribution. In case of a continuous variable, the



**Fig. 3:** Example of probability distribution of a discrete random variable  $n$ .

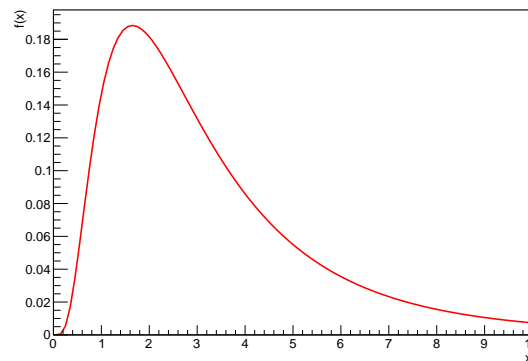
probability assigned to an individual value may be zero (e.g.:  $P(\{x\}) = 0$ ), and a *probability density function* (PDF) better quantifies the probability content of an interval with finite measure:

$$\frac{dP(x)}{dx} = f(x), \tag{5}$$

and:

$$P([x_1, x_2]) = \int_{x_1}^{x_2} f(x) dx. \tag{6}$$

Figure 4 shows an example of such a continuous distribution. Discrete and continuous distributions can



**Fig. 4:** Example of probability distribution of a continuous random variable  $x$ .

be combined using Dirac's delta functions. For instance, the following PDF:

$$\frac{dP(x)}{dx} = \frac{1}{2}\delta(x) + \frac{1}{2}f(x) \tag{7}$$

corresponds to a 50% probability to have  $x = 0$  ( $P(\{0\}) = 0.5$ ) and 50% probability to have a value  $x \neq 0$  distributed according to  $f(x)$ .

The *cumulative distribution* of a PDF  $f$  is defined as:

$$F(x) = \int_{-\infty}^x f(x) dx. \tag{8}$$

## 2.6 PDFs in more dimensions

In more dimensions, corresponding to  $n$  random variables, a PDF can be defined as:

$$\frac{d^n P(x_1, \dots, x_n)}{dx_1 \cdots dx_n} = f(x_1, \dots, x_n). \quad (9)$$

The probability associated to an event which corresponds to a subset  $E \subseteq \mathbb{R}^n$  is obtained by integrating the PDF over the set  $E$ , naturally extending Eq. (6):

$$P(E) = \int_E f(x_1, \dots, x_n) d^n x. \quad (10)$$

## 2.7 Mean, variance and covariance

For a PDF that models a random variable  $x$ , it's useful to define a number of quantities:

- The *mean* or *expected value* of  $x$  is defined as:

$$\mathbb{E}[x] = \langle x \rangle = \int x f(x) dx. \quad (11)$$

More in general, the mean or expected value of  $g(x)$  is:

$$\mathbb{E}[g(x)] = \langle g(x) \rangle = \int g(x) f(x) dx. \quad (12)$$

- The *variance* of  $x$  is defined as:

$$\mathbb{V}\text{ar}[x] = \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2. \quad (13)$$

The term  $\langle x^2 \rangle$  is called *root mean square*, or *r.m.s.*

- The *standard deviation* of  $x$  is the square root of the variance:

$$\sigma_x = \sqrt{\mathbb{V}\text{ar}[x]} = \sqrt{\langle (x - \langle x \rangle)^2 \rangle}. \quad (14)$$

Given two random variables  $x$  and  $y$ , the following quantities may be defined:

- The *covariance* of  $x$  and  $y$  is:

$$\mathbb{C}\text{ov}[x, y] = \langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle. \quad (15)$$

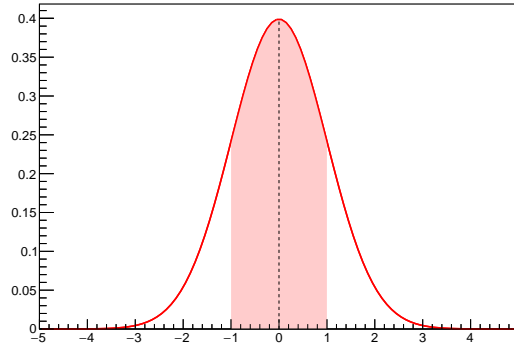
- The *correlation coefficient* is:

$$\rho_{xy} = \frac{\mathbb{C}\text{ov}[x, y]}{\sigma_x \sigma_y}. \quad (16)$$

Two variables with null covariance are said to be *uncorrelated*.

## 2.8 Commonly used distributions

Below a few examples of probability distributions are reported that are frequently used in physics and more in general in statistical applications.



**Fig. 5:** Example of Gaussian distribution with  $\mu = 0$  and  $\sigma = 1$ . The shaded interval  $[\mu - \sigma, \mu + \sigma]$  correspond to a probability of approximately 0.683.

**Table 1:** Probabilities for a Gaussian PDF corresponding to an interval  $[\mu - n\sigma, \mu + n\sigma]$ .

| $n$ | Prob.                    |
|-----|--------------------------|
| 1   | 0.683                    |
| 2   | 0.954                    |
| 3   | 0.997                    |
| 4   | $1 - 6.5 \times 10^{-5}$ |
| 5   | $1 - 5.7 \times 10^{-7}$ |

### 2.8.1 Gaussian distribution

A *Gaussian* or *normal* distribution is given by:

$$g(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \quad (17)$$

where  $\mu$  and  $\sigma$  are parameters equal to the average value and standard deviation of  $x$ , respectively. If  $\mu = 0$  and  $\sigma = 1$ , a Gaussian distribution is also called *standard normal distribution*. An example of Gaussian PDF is shown in Fig. 5. Probability values corresponding to intervals  $[\mu - n\sigma, \mu + n\sigma]$  for a Gaussian distribution are frequently used as reference, and are reported in Tab. 1. Many random variables in real experiments follow, at least approximately, a Gaussian distribution. This is mainly due to the *central limit theorem* that allows to approximate the sum of multiple random variables, regardless of their individual distributions, with a Gaussian distribution. Gaussian PDFs are frequently used to model detector resolution.

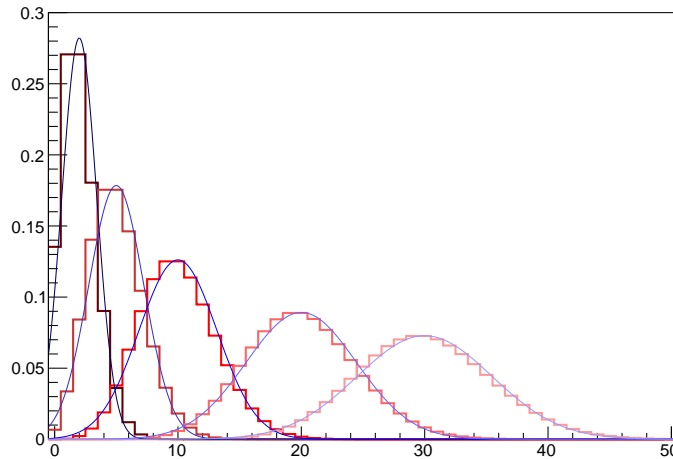
### 2.8.2 Poissonian distribution

A *Poissonian* distribution for an integer non-negative random variable  $n$  is:

$$P(n; \nu) = \frac{\nu^n}{n!} e^{-\nu}, \quad (18)$$

where  $\nu$  is a parameter equal to the average value of  $n$ . The variance of  $n$  is also equal to  $\nu$ .

Poissonian distributions model the number of occurrences of random event uniformly distributed in a measurement range whose rate is known. Examples are the number of rain drops falling in a given area and in a given time interval or the number of cosmic rays crossing a detector in a given time interval. Poissonian distributions may be approximated with a Gaussian distribution having  $\mu = \nu$  and  $\sigma = \sqrt{\nu}$  for sufficiently large values of  $\nu$ . Examples of Poissonian distributions are shown in Fig. 6 with superimposed Gaussian distributions as comparison.



**Fig. 6:** Example of Poisson distributions with different values of  $\nu$ . The continuous superimposed curves are Gaussian distributions with  $\mu = \nu$  and  $\sigma = \sqrt{\nu}$ .

### 2.8.3 Binomial distribution

A *binomial* distribution gives the probability to achieve  $n$  successful outcomes on a total of  $N$  independent trials whose individual probability of success is  $p$ . The binomial probability is given by:

$$P(n; N, p) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n}. \quad (19)$$

The average value of  $n$  for a binomial variable is:

$$\langle n \rangle = N p \quad (20)$$

and the variance is:

$$\text{Var}[n] = N p (1-p). \quad (21)$$

A typical example of binomial process in physics is the case of a detector with efficiency  $p$ , where  $n$  is the number of *detected* particles over a total number of particles  $N$  that *crossed* the detector.

## 2.9 Conditional probability

The probability of an event  $A$ , given the event  $B$  is defined as:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad (22)$$

and represents the probability that an event known to belong to set  $B$  also belongs to set  $A$ . It's worth noting that, given the sample space  $\Omega$  with  $P(\Omega) = 1$ :

$$P(A|\Omega) = \frac{P(A \cap \Omega)}{P(\Omega)}, \quad (23)$$

consistently with Eq. (22).

An event  $A$  is said to be independent on the event  $B$  if the probability of  $A$  given  $B$  is equal to the probability of  $A$ :

$$P(A|B) = P(A). \quad (24)$$

If an event  $A$  is independent on the event  $B$ , then  $P(A \cap B) = P(A)P(B)$ . Using Eq. (22), it's immediate to demonstrate that if  $A$  is independent on  $B$ , then  $B$  is independent on  $A$ .

The application of the concept of conditional probability to PDFs in more dimensions allows to introduce the concept of *independent variables*. Consider a two-variable PDF  $f(x, y)$  (but the result can be easily generalized to more than two variables), two *marginal distributions* can be defined as:

$$f_x(x) = \int f(x, y) dy, \quad (25)$$

$$f_y(y) = \int f(x, y) dx. \quad (26)$$

If we consider the sets:

$$A = \{x' : x < x' < x + \delta x\}, \quad (27)$$

$$B = \{y' : y < y' < y + \delta y\}, \quad (28)$$

where  $\delta x$  and  $\delta y$  are very small, if  $A$  and  $B$  are independent, we have:

$$P(A \cap B) = P(A)P(B), \quad (29)$$

which implies:

$$f(x, y) = f_x(x)f_y(y). \quad (30)$$

From Eq. (30), it's possible to define that  $x$  and  $y$  are independent variables if and only if their PDF can be factorized into the product of one-dimensional PDFs. Note that if two variables are uncorrelated they are not necessarily independent.

## 2.10 Bayes theorem

Considering two events  $A$  and  $B$ , using Eq. (22) twice, we can write:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad (31)$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \quad (32)$$

from which the following equation derives:

$$P(A|B)P(B) = P(B|A)P(A). \quad (33)$$

Eq. (33) can be written in the following form, that takes the name of *Bayes theorem*:

$$\boxed{P(A|B) = \frac{P(B|A)P(A)}{P(B)}}. \quad (34)$$

In Eq. (34),  $P(A)$  has the role of *prior* probability and  $P(A|B)$  has the role of *posterior* probability. Bayes theorem, that has its validity in any probability approach, including the frequentist one, can also be used to assign a posterior probability to a claim  $H$  that is necessarily not a random event, given a corresponding prior probability  $P(H)$  and the observation of an event  $E$  whose probability, if  $H$  is true, is given by  $P(E|H)$ :

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)}. \quad (35)$$

Eq. (35) is the basis of Bayesian approach to probability. It defines in a *rational* way a role to modify one's prior belief in a claim  $H$  given the observation of  $E$ .

The following problem is an example of application of Bayes theorem in a frequentist environment. Imagine you have a particle identification detector that identifies muons with high efficiency, say  $\varepsilon = 95\%$ . A small fraction of pions, say  $\delta = 5\%$ , are incorrectly identified as muons (*fakes*). Given a particle in a data sample that is identified as a muon, what is the probability that it is really a muon? The answer to this question can't be given unless we know more information about the composition of the sample, i.e.: what is the fraction of muons and pions in the data sample.

Using Bayes theorem, we can write:

$$P(\mu|+) = \frac{P(+|\mu)P(\mu)}{P(+)}, \quad (36)$$

where '+' denotes a positive muon identification,  $P(\mu|+) = \varepsilon$  is the probability to positively identify a muon,  $P(\mu)$  is the fraction of muons in our sample (*purity*) and  $P(+)$  is the probability to positively identify a particle randomly chosen from our sample.

It's possible to decompose  $P(+)$  as:

$$P(+)=P(+|\mu)P(\mu)+P(+|\pi)P(\pi), \quad (37)$$

where  $P(+|\pi) = \delta$  is the probability to positively identify a pion and  $P(\pi) = 1 - P(\mu)$  is the fraction of pions in our samples, that we suppose is only made of muons and pions. Eq. (37) is a particular case of the *law of total probability* which allows to decompose the probability of an event  $E_0$  as:

$$P(E_0)=\sum_{i=1}^n P(E_0|A_i)P(A_i), \quad (38)$$

where the sets  $A_i$  are all pairwise disjoint and constitute a partition of the sample space.

Using the decomposition from Eq. (37) in Eq. (36), one gets:

$$P(\mu|+)=\frac{\varepsilon P(\mu)}{\varepsilon P(\mu)+\delta P(\pi)}. \quad (39)$$

If we assume that our sample contains a fraction  $P(\mu) = 4\%$  of muons and  $P(\pi) = 96\%$  of pions, we have:

$$P(\mu|+)=\frac{0.95 \cdot 0.04}{0.95 \cdot 0.04 + 0.05 \cdot 0.96} \simeq 0.44. \quad (40)$$

In this case, even if the selection efficiency is very high, given the low sample purity, a particle positively identified as a muon has a probability less than 50% to be really a muon.

## 2.11 The likelihood function

The outcome of an experiment can be modeled as a set of random variables  $x_1, \dots, x_n$  whose distribution takes into account both intrinsic physics randomness (theory) and detector effects (like resolution, efficiency, etc.). Theory and detector effects can be described according to some parameters  $\theta_1, \dots, \theta_m$  whose values are, in most of the cases, unknown. The overall PDF, evaluated for our observations  $x_1, \dots, x_n$ , is called *likelihood function*:

$$L=f(x_1, \dots, x_n; \theta_1, \dots, \theta_m). \quad (41)$$

In case our sample consists of  $N$  *independent measurements*, typically each corresponding to a collision event, the likelihood function can be written as:

$$L=\prod_{i=1}^N f(x_1^i, \dots, x_n^i; \theta_1, \dots, \theta_m). \quad (42)$$

The likelihood function provides a useful implementation of Bayes rule (Eq. (35)) in the case of a measurement constituted by the observation of continuous random variables  $x_1, \dots, x_n$ . The posterior PDF of the unknown parameters  $\theta_1, \dots, \theta_m$  can be determined as:

$$P(\theta_1, \dots, \theta_m | x_1, \dots, x_n) = \frac{L(x_1, \dots, x_n; \theta_1, \dots, \theta_m) \pi(\theta_1, \dots, \theta_m)}{\int L(x_1, \dots, x_n; \theta_1, \dots, \theta_m) \pi(\theta_1, \dots, \theta_m) d\theta^m}, \quad (43)$$

where  $\pi(\theta_1, \dots, \theta_m)$  is the subjective prior probability and the denominator is a normalization factor obtained with a decomposition similar to Eq. (37). Equation (43) can be interpreted as follows: the observation of  $x_1, \dots, x_n$  modifies the prior knowledge of the unknown parameters  $\theta_1, \dots, \theta_m$ .

If  $\pi(\theta_1, \dots, \theta_m)$  is sufficiently smooth and  $L$  is sharply peaked around the true values of the parameters  $\theta_1, \dots, \theta_m$ , the resulting posterior will not be strongly dependent on the prior's choice.

Bayes theorem in the form of Eq. (43) can be applied sequentially for repeated independent observations. In fact, if we start with a prior  $P_0(\vec{\theta})$ , we can determine a posterior:

$$P_1(\vec{\theta}) \propto P_0(\vec{\theta}) \cdot L_1(\vec{x}_1; \vec{\theta}), \quad (44)$$

where  $L_1(\vec{x}_1; \vec{\theta})$  is the likelihood function corresponding to the observation  $\vec{x}_1$ . Subsequently, we can use  $P_1$  as new prior for a second observation  $\vec{x}_2$ , and we can determine a new posterior:

$$P_2(\vec{\theta}) \propto P_1(\vec{\theta}) \cdot L_2(\vec{x}_2; \vec{\theta}), \quad (45)$$

and so on:

$$P_3(\vec{\theta}) \propto P_2(\vec{\theta}) \cdot L_3(\vec{x}_3; \vec{\theta}). \quad (46)$$

For independent observations  $\vec{x}_1, \vec{x}_2, \vec{x}_3$ , the combined likelihood function can be written as the product of individual likelihood functions (Eq. (30)):

$$P_3(\vec{\theta}) \propto P_0(\vec{\theta}) \cdot L_1(\vec{x}_1; \vec{\theta}) \cdot L_2(\vec{x}_2; \vec{\theta}) \cdot L_3(\vec{x}_3; \vec{\theta}), \quad (47)$$

consistently with Eq. (46). This allows to use consistently the repeated application of Bayes rule as sequential improvement of knowledge from subsequent observations.

### 3 Inference

In Sec. 2 we presented how probability theory can model the fluctuation in data due to intrinsic randomness of observable data samples. Taking into account the distribution of data as a function of the values of unknown parameters, we can exploit the observed data in order to determine information about the parameters, in particular to measure their value (*central value*) within some *uncertainty*. This process is called *inference*.

#### 3.1 Bayesian inference

One example of inference is the use of Bayes theorem to determine the posterior PDF of an unknown parameter  $\theta$  given an observation  $x$ :

$$P(\theta|x) = \frac{L(x; \theta) \pi(\theta)}{\int L(x; \theta) \pi(\theta) d\theta}, \quad (48)$$

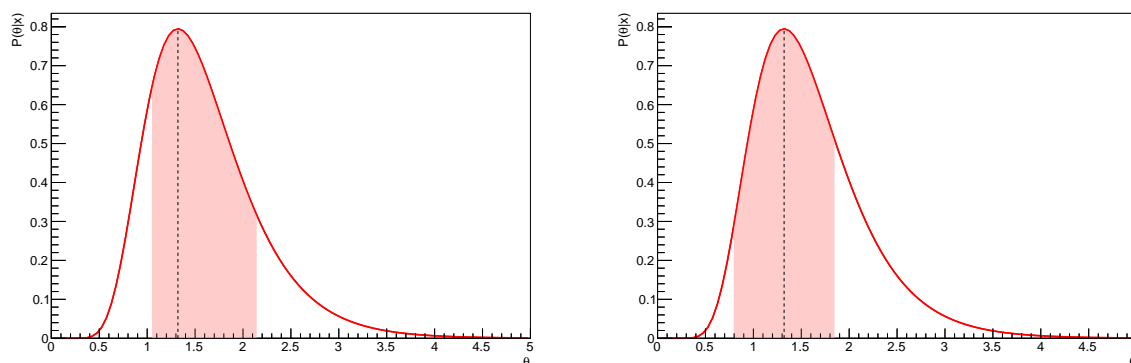
where  $\pi(\theta)$  is the prior PDF. The posterior  $P(\theta|x)$  contains all the information we can obtain from  $x$  about  $\theta$ . One example of possible outcome for  $P(\theta|x)$  is shown in Fig. 7 with two possible choices of uncertainty interval (left and right plots). The most probable value,  $\hat{\theta}$ , also called *mode*, shown as dashed line in both plots, can be taken as central value for the parameter  $\theta$ . It's worth noting that if  $\pi(\theta)$  is assumed to be a constant,  $\hat{\theta}$  corresponds to the maximum of the likelihood function (*maximum likelihood*

estimate, see Sec. 3.5). Different choices of 68.3% probability interval, or uncertainty interval, can be taken. A central interval  $[\theta_1, \theta_2]$ , represented in the left plot in Fig. 7 as shaded area, is obtained in order to have equal areas under the two extreme tails:

$$\int_{-\infty}^{\theta_1} P(\theta|x) d\theta = \frac{\alpha}{2}, \quad (49)$$

$$\int_{\theta_2}^{+\infty} P(\theta|x) d\theta = \frac{\alpha}{2}, \quad (50)$$

where  $\alpha = 1 - 68.3\%$ . Another example of a possible choice of 68.3% interval is shown in the right plot,



**Fig. 7:** Example of a posterior PDF of the parameter  $\theta$  with two possible choices of a 68.3% probability interval shown as shaded area: a central interval (left plot) and a symmetric interval (right plot). The dotted vertical line shows the most probable value (mode).

where a symmetric interval is taken, corresponding to:

$$\int_{\hat{\theta}-\delta}^{\hat{\theta}+\delta} P(\theta|x) d\theta = 1 - \alpha. \quad (51)$$

$$(52)$$

Two extreme choices of fully asymmetric probability intervals are shown in Fig. 8, leading to an upper (left) or lower (right) limit to the parameter  $\theta$ . For upper or lower limits, usually a 90% or 95% probability interval is chosen instead of the usual 68.3% used for central or symmetric intervals. The intervals in Fig. 8 are chosen such that:

$$\int_{-\infty}^{\theta^{\text{up}}} P(\theta|x) d\theta = 1 - \alpha \quad (\text{left plot}), \quad (53)$$

$$\int_{\theta^{\text{lo}}}^{+\infty} P(\theta|x) d\theta = 1 - \alpha \quad (\text{right plot}), \quad (54)$$

$$(55)$$

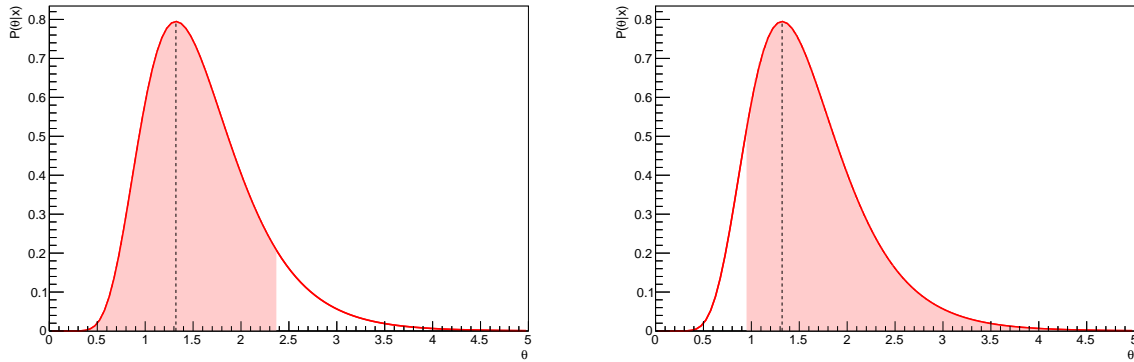
where in this case  $\alpha = 0.1$ .

### 3.1.1 Example of Bayesian inference: Poissonian counting

In a counting experiment, i.e.: the only information relevant to measure the yield of our signal is the number of events  $n$  that pass a given selection, a Poissonian can be used to model the distribution of  $n$  with an expected number of events  $s$ :

$$P(n; s) = \frac{s^n e^{-s}}{n!}. \quad (56)$$





**Fig. 8:** Extreme choices of 90% probability interval leading to an upper limit (left) and a lower limit (up) to the parameter  $\theta$ .

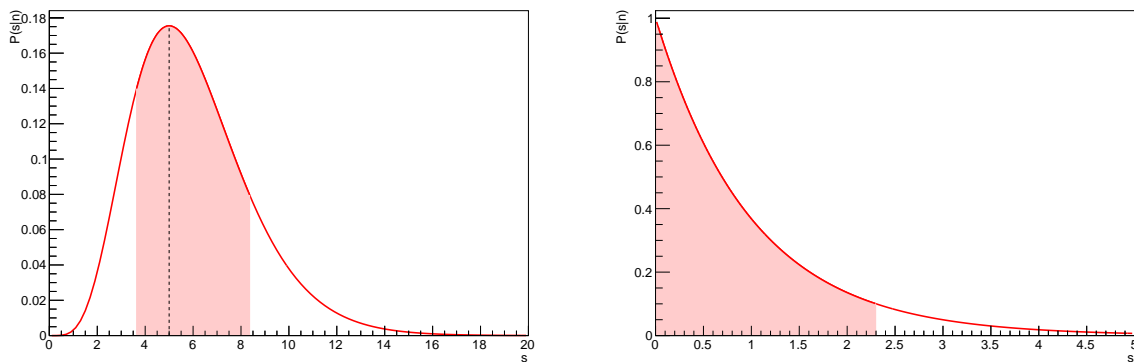
If a particular value of  $n$  is measured, the posterior PDF of  $s$  is (Eq. (48)):

$$P(s|n) = \frac{\frac{s^n e^{-s}}{n!} \pi(s)}{\int_0^\infty \frac{s'^n e^{-s'}}{n!} \pi(s') ds'}$$
(57)

where  $\pi(s)$  is the assumed prior for  $s$ . If we take  $\pi(s)$  to be uniform, performing the integration gives a denominator equal to one, hence:

$$P(s|n) = \frac{s^n e^{-s}}{n!}$$
(58)

Note that though Eqs. (56) and (58) lead to the same expression, the former is a probability for the discrete random variable  $n$ , the latter is a posterior PDF of the unknown parameter  $s$ . From Eq. (56), the mode  $\hat{s}$  is equal to  $n$ , but  $\langle s \rangle = n + 1$ , due to the asymmetric distribution of  $s$ , and  $\mathbb{V}\text{ar}[s] = n + 1$ , while the variance of  $n$  for a Poissonian distribution is  $\sqrt{s}$  (Sec. 2.8.2).



**Fig. 9:** Posterior PDF of a Poissonian parameter  $s$  for observed number of events  $n = 5$  (left) and for  $n = 0$  (right). In the left plot, a central 68.3% probability interval is chosen, while for the right plot a fully asymmetric 90% probability interval, leading to an upper limit, is chosen.

Figure 9 shows two cases of posterior PDF of  $s$ , for the cases  $n = 5$  (left) and for  $n = 0$  (right). In the case  $n = 5$ , a central value  $\hat{s} = 5$  can be taken as most probable value. In that plot, a central interval was chosen (Eq. (49, 50)). For the case  $n = 0$ , the most probable value of  $s$  is  $\hat{s} = 0$ . A fully asymmetric

interval corresponding to a probability  $1 - \alpha$  leads to an upper limit:

$$e^{-s^{\text{up}}} = \alpha, \quad (59)$$

which then leads to:

$$s < s^{\text{up}} = 2.303 \quad \text{for } \alpha = 0.1 \text{ (90\% probability)}, \quad (60)$$

$$s < s^{\text{up}} = 2.996 \quad \text{for } \alpha = 0.05 \text{ (95\% probability)}. \quad (61)$$

### 3.2 Error propagation with Bayesian inference

Error propagation is needed when applying a parameter transformation, say  $\eta = H(\theta)$ . A central value and uncertainty interval need to be determined for the transformed parameter  $\eta$ . With Bayesian inference, a posterior PDF of  $\theta$ ,  $f(\theta)$  is available, and the error propagation can be done transforming the posterior PDF of  $\theta$  into a PDF of  $\eta$ ,  $f'(\eta)$ : the central value and uncertainty interval for  $\eta$  can be computed from  $f'$ . In general, the PDF of the transformed variable  $\eta$ , given the PDF of  $\theta$ , is given by:

$$f'(\eta) = \int \delta(\eta - H(\theta))f(\theta) d\theta. \quad (62)$$

Transformations for cases with more than one variable proceed in a similar way. If we have two parameters  $\theta_1$  and  $\theta_2$ , and a transformed variable  $\eta = H(\theta_1, \theta_2)$ , then the PDF of  $\eta$ , similarly to Eq. (62), is given by:

$$f'(\eta) = \int \delta(\eta - H(\theta_1, \theta_2))f(\theta_1, \theta_2) d\theta_1 d\theta_2. \quad (63)$$

In case of a transformation from two parameters  $\theta_1$  and  $\theta_2$  in two other parameters  $\eta_1$  and  $\eta_2$ :  $\eta_1 = H_1(\theta_1, \theta_2)$ ,  $\eta_2 = H_2(\theta_1, \theta_2)$ , we have:

$$f'(\eta_1, \eta_2) = \int \delta(\eta_1 - H_1(\theta_1, \theta_2))\delta(\eta_2 - H_2(\theta_1, \theta_2))f(\theta_1, \theta_2) d\theta_1 d\theta_2. \quad (64)$$

### 3.3 Choice of the prior

One of the most questionable issue related to Bayesian inference is the subjectiveness of the result, being dependent on the choice of a prior. In particular, there is no unique choice of a prior that models one's ignorance about an unknown parameter. A choice of a uniform prior, such as it was done in Sec. 3.1.1, is also questionable: if the prior PDF is uniform in a chosen variable, it won't necessarily be uniform when applying a coordinate transformation to that variable. A typical example is the measurement of a particle's lifetime, which is the inverse of the particle's width. Given any choice of a regular prior for a parameter, there is always a transformation that makes the PDF uniform.

Harold Jeffreys provided a method [5] to chose a form of the prior that is invariant under parameter transformation. The choice uses the so-called Fishers information matrix, which, given a set of parameters  $\vec{\theta} = (\theta_1, \dots, \theta_m)$ , is defined as:

$$\mathcal{I}_{ij}(\vec{\theta}) = \left\langle \frac{\partial \ln L(\vec{x}; \vec{\theta})}{\partial \theta_i} \frac{\partial \ln L(\vec{x}; \vec{\theta})}{\partial \theta_j} \right\rangle. \quad (65)$$

Jeffrey's prior is then given by, up to a normalization factor:

$$\pi(\vec{\theta}) \propto \sqrt{\det \mathcal{I}(\vec{\theta})}. \quad (66)$$

It's possible to demonstrate that Eq. (66) is invariant under a parameter transformation  $\vec{\eta} = \vec{H}(\vec{\theta})$ .

### 3.4 Frequentist inference

Assigning a probability level to an unknown parameter makes no sense in the frequentist approach since unknown parameters are not random variables. A frequentist inference procedure should determine a central value and an uncertainty interval that depend on the observed measurements without introducing any subjective element. Such central value and interval extremes are random variables themselves. The function that returns the central value given an observed measurement is called *estimator*. The parameter value provided by an estimator is also called *best fit* value. Different estimator choices are possible, the most frequently adopted is the *maximum likelihood estimator* because of its statistical properties discussed in Sec. 3.7.

Repeating the experiment will result each time in a different data sample and, for each data sample, the estimator returns a different central value  $\hat{\theta}$ . An uncertainty interval  $[\hat{\theta} - \delta, \hat{\theta} + \delta]$  can be associated to the estimator value  $\hat{\theta}$ . In some cases, as for the Bayesian inference, an asymmetric interval choice is also possible with frequentist inference:  $[\hat{\theta} - \delta^-, \hat{\theta} + \delta^+]$ . Some of the intervals obtained with this method contain the fixed and unknown true value of  $\theta$ , corresponding to a fraction equal to 68.3% of the repeated experiments, in the limit of very large number of experiments. This property is called *coverage*.

The simplest example of frequentist inference assumes a Gaussian PDF (Eq. (17)) with a known  $\sigma$  and an unknown  $\mu$ . A single experiment provides a measurement  $x$ , and we can estimate  $\mu$  as  $\hat{\mu} = x$ . The distribution of  $\hat{\mu}$  is the original Gaussian because  $\hat{\mu}$  is just equal to  $x$ . A fraction of 68.3% of the experiments (in the limit of large number of repetitions) will provide an estimate  $\hat{\mu}$  within:  $\mu - \sigma < \hat{\mu} < \mu + \sigma$ . This means that we can quote:

$$\boxed{\mu = x \pm \sigma.} \quad (67)$$

### 3.5 Maximum likelihood estimates

The maximum likelihood method takes as best-fit values of the unknown parameter the values that maximize the likelihood function (defined Sec. 2.11). The maximization of the likelihood function can be performed analytically only in the simplest cases, while a numerical treatment is needed in most of the realistic cases. MINUIT [6] is historically the most widely used minimization software engine in High Energy Physics.

#### 3.5.1 Extended likelihood function

Given a sample of  $N$  measurements of the variables  $\vec{x} = (x_1, \dots, x_n)$ , the likelihood function expresses the probability density evaluated for our sample as a function of the unknown parameters  $\theta_1, \dots, \theta_m$ :

$$L(\vec{x}_1, \dots, \vec{x}_N) = \prod_{i=1}^N f(x_1^i, \dots, x_n^i; \theta_1, \dots, \theta_m). \quad (68)$$

The size  $N$  of the sample is in many cases also a random variable. In those cases, the *extended likelihood function* can be defined as:

$$L(\vec{x}_1, \dots, \vec{x}_N) = P(N; \theta_1, \dots, \theta_m) \prod_{i=1}^N f(x_1^i, \dots, x_n^i; \theta_1, \dots, \theta_m), \quad (69)$$

where  $P(N; \theta_1, \dots, \theta_m)$  is the distribution of  $N$ , and in practice is always a Poissonian whose expected rate parameter is a function of the unknown parameters  $\theta_1, \dots, \theta_m$ :

$$P(N; \theta_1, \dots, \theta_m) = \frac{\nu(\theta_1, \dots, \theta_m)^N e^{-\nu(\theta_1, \dots, \theta_m)}}{N!}. \quad (70)$$

In many cases, either with a standard or an extended likelihood function, it may be convenient to use  $-\ln L$  or  $-2 \ln L$  in the numerical treatment rather than  $L$ , because the product of the various

terms is transformed into the sum of the logarithms of those terms, which may have advantages in the computation.

For a Poissonian process that is given by the sum of a signal plus a background process, the extended likelihood function may be written as:

$$L(\vec{x}; s, b, \vec{\theta}) = \frac{(s+b)^N e^{-(s+b)}}{N!} \prod_{i=1}^N \left( f_s P_s(x_i; \vec{\theta}) + f_b P_b(x_i; \vec{\theta}) \right), \quad (71)$$

where  $s$  and  $b$  are the signal and background expected yields, respectively,  $f_s$  and  $f_b$  are the fraction of signal and background events, namely:

$$f_s = \frac{s}{s+b}, \quad (72)$$

$$f_b = \frac{b}{s+b}, \quad (73)$$

and  $P_s$  and  $P_b$  are the PDF of the variable  $x$  for signal and background, respectively. Replacing  $f_s$  and  $f_b$  into Eq. (71) gives:

$$L(\vec{x}; s, b, \vec{\theta}) = \frac{e^{-(s+b)}}{N!} \prod_{i=1}^N \left( s P_s(x_i; \vec{\theta}) + b P_b(x_i; \vec{\theta}) \right). \quad (74)$$

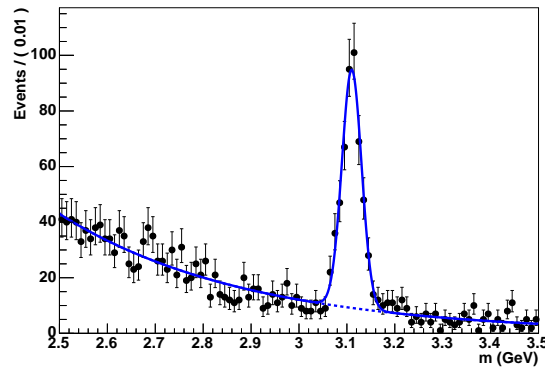
It may be more convenient to use the negative logarithm of Eq. (74), that should be minimize in order to determine the best-fit values of  $s$ ,  $b$  and  $\vec{\theta}$ :

$$-\ln L(\vec{x}; s, b, \vec{\theta}) = s + b - \sum_{i=1}^N \ln \left( s P_s(x_i; \vec{\theta}) + b P_b(x_i; \vec{\theta}) \right) + \ln N!. \quad (75)$$

The last term  $\ln N!$  is a constant with respect to the fit parameters, and can be omitted in the minimization. In many cases, instead of using  $s$  as parameter of interest, the *signal strength*  $\mu$  is introduced, defined by the following equation:

$$s = \mu s_0, \quad (76)$$

where  $s_0$  is the theory prediction for the signal yield.  $\mu = 1$  corresponds to the nominal value of the theory prediction for the signal yield.



**Fig. 10:** Example of an unbinned maximum likelihood fit. Data are fit using a Gaussian distribution for the signal and an exponential distribution for the background. This figure is taken from Ref. [7].

An example of unbinned maximum likelihood fit is given in Fig. 10, where the data are fit with a model inspired to Eq. (74), with  $P_s$  and  $P_b$  taken as a Gaussian and an exponential distribution, respectively. The observed variable has been called  $m$  in that case because the spectrum resembles an invariant mass peak, and the position of the peak at 3.1 GeV reminds a  $J/\psi$  particle. The two PDFs can be written as:

$$P_s(m) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(m-\mu)^2}{2\sigma^2}}, \quad (77)$$

$$P_b(m) = \lambda e^{-\lambda m}. \quad (78)$$

The parameters  $\mu$ ,  $\sigma$  and  $\lambda$  are fit together with the signal and background yields  $s$  and  $b$ . While  $s$  is our *parameter of interest*, because we will eventually determine a production cross section or branching fraction from its measurement, the other additional parameters, that are not directly related to our final measurement, are said *nuisance parameters*. In general, nuisance parameters are needed to model background yield, detector resolution and efficiency, various parameters modeling the signal and background shapes, etc. Nuisance parameters are also important to model *systematic uncertainties*, as will be discussed more in details in the following sections.

### 3.6 Estimate of Gaussian parameters

If we have  $n$  independent measurements  $\vec{x} = (x_1, \dots, x_n)$  all modeled (exactly or approximatively) with the same Gaussian PDF, we can write the negative of twice the logarithm of the likelihood function as follows:

$$-2 \ln L(\vec{x}; \mu) = \sum_{i=1}^n \frac{(x_i - \mu)^2}{\sigma^2} + n(\ln 2\pi + 2 \ln \sigma). \quad (79)$$

The first term,  $\sum_{i=1}^n \frac{(x_i - \mu)^2}{\sigma^2}$ , is an example of  $\chi^2$  variable (see Sec. 3.13).

An analytical minimization of  $-2 \ln L$  with respect to  $\mu$ , assuming  $\sigma^2$  is known, gives the *arithmetic mean* as maximum likelihood estimate of  $\mu$ :

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (80)$$

If  $\sigma^2$  is also unknown, the maximum likelihood estimate of  $\sigma^2$  is:

$$\widehat{\sigma^2} = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2. \quad (81)$$

The estimate in Eq. (81) can be demonstrated to have an unpleasant feature, called *bias*, that will be discussed in Sec. 3.7.2.

### 3.7 Estimator properties

This section illustrates the main properties of estimators. Maximum likelihood estimators are most frequently chosen because they have good performances for what concerns those properties.

#### 3.7.1 Consistency

For large number of measurements, the estimator  $\hat{\theta}$  should converge, in probability, to the true value of  $\theta$ ,  $\theta^{\text{true}}$ . Maximum likelihood estimators are consistent.

### 3.7.2 Bias

The bias of a parameter is the average value of its deviation from the true value:

$$\mathbb{b}[\hat{\theta}] = \langle \hat{\theta} - \theta^{\text{true}} \rangle = \langle \hat{\theta} \rangle - \theta^{\text{true}}. \quad (82)$$

An *unbiased estimator* has  $\mathbb{b}[\theta] = 0$ . Maximum likelihood estimators may have a bias, but the bias decreases with large number of measurements (if the model used in the fit is correct).

In the case of the estimate of a Gaussian's  $\sigma^2$ , the maximum likelihood estimate (Eq. (81)) underestimates the true variance. The bias can be corrected for by applying a multiplicative factor:

$$\widehat{\sigma}_{\text{unbias.}}^2 = \frac{n}{n-1} \widehat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{\mu})^2. \quad (83)$$

### 3.7.3 Efficiency

The variance of any consistent estimator is subject to a lower bound due to Cramér [8] and Rao [9]:

$$\text{Var}[\hat{\theta}] \geq \frac{\left(1 + \frac{\partial \mathbb{b}[\theta]}{\partial \theta}\right)^2}{\left\langle \left(\frac{\partial \ln L(\vec{x}; \theta)}{\partial \theta}\right)\right\rangle} = \mathbb{V}_{\text{CR}}[\hat{\theta}]. \quad (84)$$

For an unbiased estimator, the numerator in Eq. (84) is equal to one. The denominator in Eq. (84) is the Fisher information (Eq. (65)).

The *efficiency* of an estimator  $\hat{\theta}$  is the ratio of the Cramér–Rao bound and the estimator's variance:

$$\varepsilon(\hat{\theta}) = \frac{\mathbb{V}_{\text{CR}}[\hat{\theta}]}{\text{Var}[\hat{\theta}]} \quad (85)$$

The efficiency for maximum likelihood estimators tends to one for large number of measurements. In other words, maximum likelihood estimates have, asymptotically, the smallest variance of all possible consistent estimators.

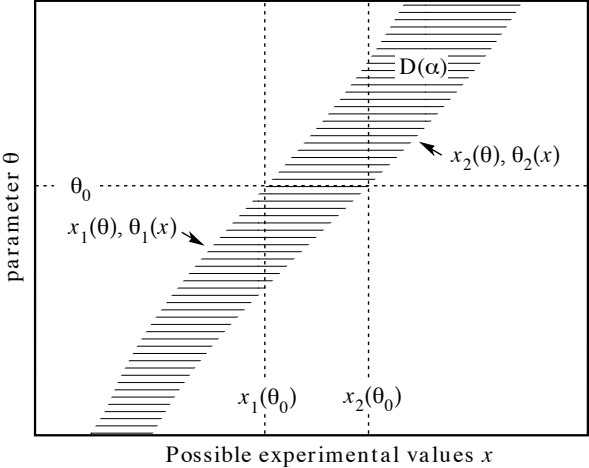
## 3.8 Neyman's confidence intervals

A procedure to determine frequentist *confidence intervals* is due to Neyman [10]. It proceeds as follows:

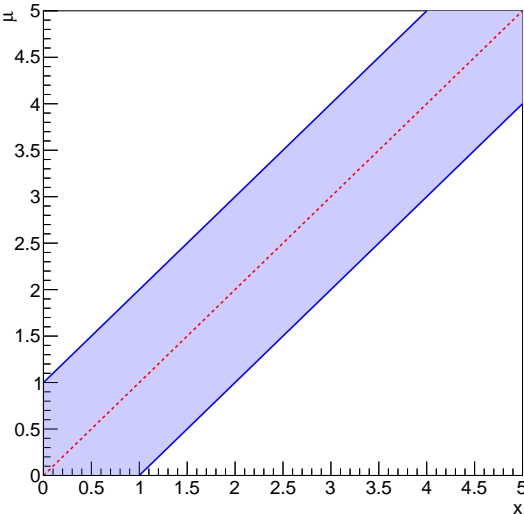
- Scan the allowed range of the unknown parameter of interest  $\theta$ .
- Given a value  $\theta_0$  of  $\theta$ , compute the interval  $[x_1(\theta_0), x_2(\theta_0)]$  that contains  $x$  with a probability  $1 - \alpha$  (*confidence level*, or CL) equal to 68.3% (or 90%, 95%). For this procedure, a choice of interval (*ordering rule*) is needed, as discussed in Sec. 3.1.
- For the observed value of  $x$ , invert the confidence belt: find the corresponding interval  $[\theta_1(x), \theta_2(x)]$ .

By construction, a fraction of the experiments equal to  $1 - \alpha$  will measure  $x$  such that the corresponding *confidence interval*  $[\theta_1(x), \theta_2(x)]$  contains (*covers*) the true value of  $\theta$ . It should be noted that the random variables are  $\theta_1(x)$  and  $\theta_2(x)$ , not  $\theta$ . An example of application of the Neyman's belt construction and inversion is shown in Fig. 11.

The simplest application of Neyman's belt construction can be done with a Gaussian distribution with known parameter  $\sigma = 1$ , as shown in Fig. 12. The belt inversion is trivial and gives the expected result: a central value  $\hat{\mu} = x$  and a confidence interval  $[\mu_1, \mu_2] = [x - \sigma, x + \sigma]$ . The result can be quoted as  $\mu = x \pm \sigma$ , similarly to what was determined with Eq. (67).



**Fig. 11:** Example Neyman’s belt construction and inversion. This figure is taken from Ref. [11].



**Fig. 12:** Example of Neyman’s belt construction for a Gaussian distribution with  $\sigma = 1$ ,  $1 - \alpha = 0.683$ .

### 3.9 Binomial intervals

The Neyman's belt construction may only guarantee approximate coverage in case of a discrete variable  $n$ . This because the interval for a discrete variable is a set of integer values,  $\{n_{\min}, \dots, n_{\max}\}$ , and cannot be "tuned" like in a continuous case. The choice of the discrete interval should be such to provide *at least* the desired coverage (i.e.: it may *overcover*). For a binomial distribution, the problem consists of finding the interval such that:

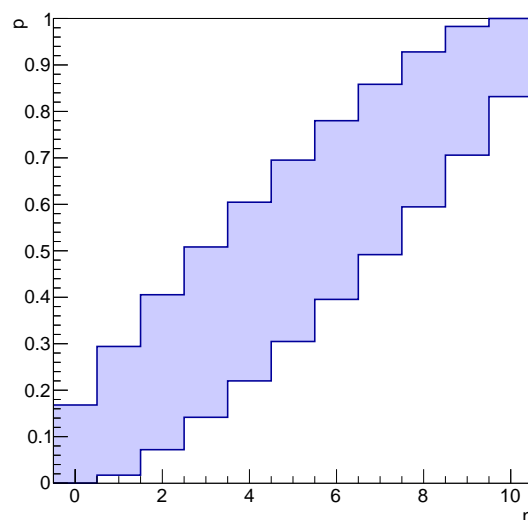
$$\sum_{n=n_{\min}}^{n_{\max}} \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n} \geq 1 - \alpha. \quad (86)$$

Clopper and Pearson [12] solved the belt inversion problem for central intervals. For an observed  $n = k$ , one has to find the lowest  $p^{\text{lo}}$  and highest  $p^{\text{up}}$  such that:

$$P(n \geq k | N, p^{\text{lo}}) = \frac{\alpha}{2}, \quad (87)$$

$$P(n \leq k | N, p^{\text{up}}) = \frac{\alpha}{2}. \quad (88)$$

An example of Neyman belt constructed using the Clopper–Pearson method is shown in Fig. 13. For



**Fig. 13:** Neyman belt construction for binomial intervals,  $N = 10$ ,  $1 - \alpha = 0.683$ .

instance for  $n = N$ , Eq. (87) becomes:

$$P(n \geq N | N, p^{\text{lo}}) = P(n = N | N, p^{\text{lo}}) = (p^{\text{lo}})^N = \frac{\alpha}{2}, \quad (89)$$

hence, for the specific case  $N = 10$ :

$$p^{\text{lo}} = \sqrt[10]{\frac{\alpha}{2}} = 0.83 \quad (1 - \alpha = 0.683), \quad 0.74 \quad (1 - \alpha = 0.90). \quad (90)$$

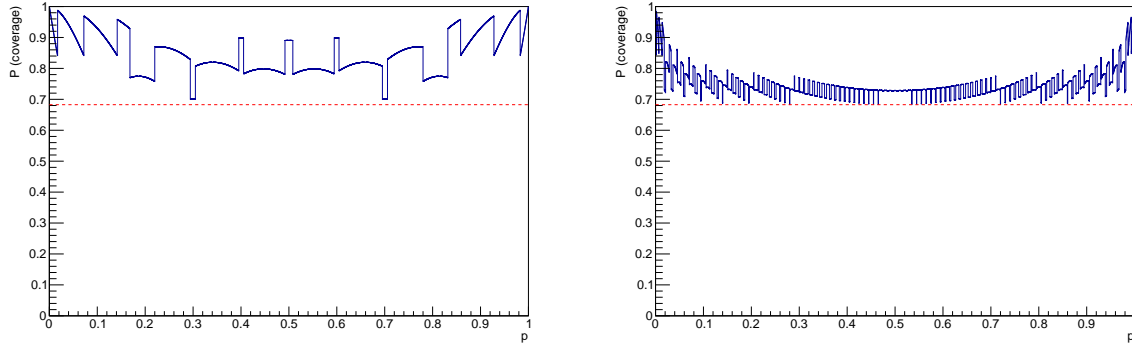
In fact, in Fig. 13, the bottom line of the belt reaches the value  $p = 0.83$  for  $n = 10$ . A frequently used approximation, inspired by Eq. (21) is:

$$\hat{p} = \frac{n}{N}, \quad \sigma_{\hat{p}} \simeq \sqrt{\frac{\hat{p}(1-\hat{p})}{N}}. \quad (91)$$



Eq. (91) gives  $\sigma_{\hat{p}} = 0$  for  $n = 0$  or  $N = n$ , which is clearly an underestimate of the uncertainty on  $\hat{p}$ . For this reason, Clopper–Pearson intervals should be preferred to the approximate formula in Eq. (91).

Clopper–Pearson intervals are often defined as “exact” in literature, though exact coverage is often impossible to achieve for discrete variables. Figure 14 shows the coverage of Clopper–Pearson intervals



**Fig. 14:** Coverage of Clopper–Pearson intervals for  $N = 10$  (left) and for  $N = 100$  (right).

as a function of  $p$  for  $N = 10$  and  $N = 100$  for  $1 - \alpha = 0.683$ . A “ripple” structure is present which, for large  $N$ , tends to get closer to the nominal 68.3% coverage.

### 3.10 Approximate error evaluation for maximum likelihood estimates

A parabolic approximation of  $-2 \ln L$  around the minimum is equivalent to a Gaussian approximation, which may be sufficiently accurate in many but not all cases. For a Gaussian model,  $-2 \ln L$  is given by:

$$-2 \ln L(\vec{x}; \mu, \sigma) = \sum_{i=1}^n \frac{(x_i - \mu)^2}{\sigma^2} + \text{const.} \quad (92)$$

An approximate estimate of the covariance matrix is obtained from the 2<sup>nd</sup> order partial derivatives with respect to the fit parameters at the minimum:

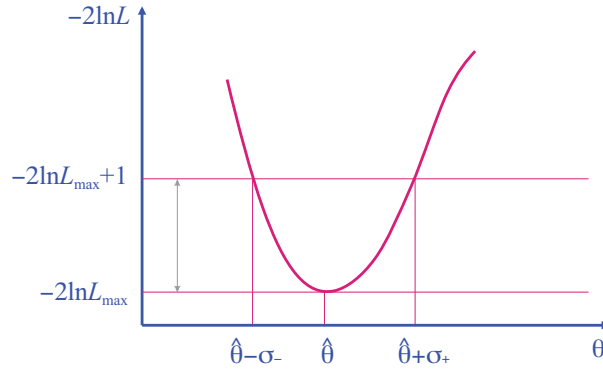
$$V_{ij}^{-1} = - \left. \frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j} \right|_{\theta_k = \hat{\theta}_k, \forall k} \quad (93)$$

Another approximation alternative to the parabolic one from Eq. (93) is the evaluation of the excursion range of  $-2 \ln L$  around the minimum, as visualized in Fig. 15. The uncertainty interval can be determined as the range around the minimum of  $-2 \ln L$  for which  $-2 \ln L$  increases by  $+1$  (or  $+n^2$  for a  $n\sigma$  interval). Errors can be asymmetric with this approach if the curve is asymmetric. For a Gaussian case the result is identical to the 2<sup>nd</sup> order derivative matrix (Eq. (93)).

### 3.11 Two-dimensional uncertainty contours

In more dimensions, i.e.: for the simultaneous determination of more unknown parameters from a fit, it’s still possible to determine multi-dimensional contours corresponding to  $1\sigma$  or  $2\sigma$  probability level. It should be noted that the scan of  $-2 \ln L$  in the multidimensional space, looking for an excursion of  $+1$  with respect to the value at the minimum, may give probability levels smaller than the corresponding values in one dimension. For a Gaussian case in one dimension, the probability associated to an interval  $[-n\sigma, +n\sigma]$  is given, integrating Eq. (17), by:

$$P_{1D}(n\sigma) = \sqrt{\frac{2}{\pi}} \int_0^n e^{-\frac{x^2}{2}} dx = \text{erf} \left( \frac{n}{\sqrt{2}} \right) \quad (94)$$



**Fig. 15:** Scan of  $-2 \ln L$  in order to determine asymmetric  $1\sigma$  errors. This figure is taken from Ref. [7].

For a two-dimensional Gaussian distribution, i.e.: the product of two independent Gaussian PDF, the probability associated to the contour with elliptic shape for which  $-2 \ln L$  increases by  $+(n\sigma)^2$  with respect to its minimum is:

$$P_{2D}(n\sigma) = \int_0^n e^{-\frac{r^2}{2}} r dr = 1 - e^{-\frac{n^2}{2}}. \quad (95)$$

Table. 2 reports numerical values for Eq. (94) and Eq. (95) for various  $n\sigma$  levels. In two dimensions, for

**Table 2:** Probabilities for 1D interval and 2D contours with different  $n\sigma$  levels..

| $n\sigma$     | $P_{1D}$ | $P_{2D}$ |
|---------------|----------|----------|
| $1\sigma$     | 0.6827   | 0.3934   |
| $2\sigma$     | 0.9545   | 0.8647   |
| $3\sigma$     | 0.9973   | 0.9889   |
| $1.515\sigma$ |          | 0.6827   |
| $2.486\sigma$ |          | 0.9545   |
| $3.439\sigma$ | 0.9973   |          |

instance, in order to recover a  $1\sigma$  probability level in one dimension (68.3%), a contour corresponding to an excursion of  $-2 \ln L$  from its minimum of  $+1.515^2$  should be considered, and for a  $2\sigma$  probability level in one dimension (95.5%), the excursion should be  $+2.486^2$ . Usually two-dimensional intervals corresponding to one or two sigma are reported, whose one-dimensional projection correspond to 68% or 95% probability content, respectively.

### 3.12 Error propagation

In case of frequentist estimates, error propagation can't be performed with a simple procedure as for the Bayesian case, where the full posterior PDF is available (Sec. 3.2).

Imagine we estimate from a fit the parameter set  $\vec{\theta} = (\theta_1, \dots, \theta_n) = \hat{\vec{\theta}}$  and we know their covariance matrix  $\Theta_{ij}$ , for instance using Eq. (93). We want to determine a new set of parameters that are functions of  $\vec{\theta}$ :  $\vec{\eta} = (\eta_1, \dots, \eta_m) = \vec{\eta}(\vec{\theta})$ . The best approach would be to rewrite the original likelihood function as a function of  $\vec{\eta}$  instead of  $\vec{\theta}$ , and perform the minimization and error estimate again for  $\vec{\eta}$ . In particular, the central value for  $\hat{\vec{\eta}}$  will be equal to the transformed of the central value  $\hat{\vec{\theta}}$ , but no obvious transformation rule exists for the uncertainty intervals.

Reparametrizing the likelihood function is not always feasible. One typical case is when central values and uncertainties for  $\vec{\theta}$  are given in a publication, but the full likelihood function is not available. For small uncertainties, a linear approximation may be sufficient to obtain the covariance matrix  $H_{ij}$  for

$\vec{\eta}$ . A Taylor expansion around the central value  $\hat{\theta}$  gives, using the error matrix  $\Theta_{ij}$ , at first order:

$$H_{ij} = \sum_{k,l} \frac{\partial \eta_i}{\partial \theta_k} \frac{\partial \eta_j}{\partial \theta_l} \Theta_{kl} \Big|_{\vec{\theta}=\hat{\theta}}. \quad (96)$$

The application of Eq. (96) gives well-known error propagation formulae reported below as examples, valid in case of null (or negligible) correlation:

$$\sigma_{x+y} = \sigma_{x-y} = \sqrt{\sigma_x^2 + \sigma_y^2}, \quad (97)$$

$$\frac{\sigma_{xy}}{xy} = \frac{\sigma_{x/y}}{x/y} = \sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}, \quad (98)$$

$$\sigma_{x^2} = 2x\sigma_x, \quad (99)$$

$$\sigma_{\ln x} = \frac{\sigma_x}{x}. \quad (100)$$

### 3.13 Likelihood function for binned samples

Sometimes data are available in form of a binned histogram. This may be convenient when a large number of entries is available, and computing an unbinned likelihood function (Eq. (42)) would be too much computationally expensive. In most of the cases, each bin content is independent on any other bin and all obey Poissonian distributions, assuming that bins contain event-counting information. The likelihood function can be written as product of Poisson PDFs corresponding to each bin whose number of entries is given by  $n_i$ . The expected number of entries in each bin depends on some unknown parameters:  $\mu_i = \mu_i(\theta_1, \dots, \theta_m)$ . The function to be minimized, in order to fit  $\theta_1, \dots, \theta_m$ , is the following:

$$-2 \ln L(\vec{n}; \vec{\theta}) = -2 \ln \prod_{i=1}^{n_{\text{bins}}} \text{Pois}(n_i; \mu_i(\theta_1, \dots, \theta_m)) \quad (101)$$

$$= -2 \sum_{i=1}^{n_{\text{bins}}} \ln \frac{e^{-\mu_i(\theta_1, \dots, \theta_m)} \mu_i(\theta_1, \dots, \theta_m)^{n_i}}{n_i!} \quad (102)$$

$$= 2 \sum_{i=1}^{n_{\text{bins}}} (\mu_i(\theta_1, \dots, \theta_m) - n_i \ln \mu_i(\theta_1, \dots, \theta_m) + \ln n_i!). \quad (103)$$

The expected number of entries in each bin,  $\mu_i$ , is often approximated by a continuous function  $\mu(x)$  evaluated at the center of the bin  $x = x_i$ . Alternatively,  $\mu_i$  can be given by the superposition of other histograms (*templates*), e.g.: the sum of histograms obtained from different simulated processes. The overall yields of the considered processes may be left as free parameters in the fit in order to constrain the normalization of simulated processes from data, rather than relying on simulation prediction, which may be affected by systematic uncertainties.

The distribution of the number of entries in each bin can be approximated, for sufficiently large number of entries, by a Gaussian with standard deviation equal to  $\sqrt{n_i}$ . Maximizing  $L$  is equivalent to minimize:

$$\chi^2 = \sum_{i=1}^{n_{\text{bins}}} \frac{(n_i - \mu(x_i; \theta_1, \dots, \theta_m))^2}{n_i} \quad (104)$$

Equation (104) defines the so-called Neyman's  $\chi^2$  variable. Sometimes, the denominator  $n_i$  is replaced by  $\mu_i = \mu(x_i; \theta_1, \dots, \theta_m)$  (Pearson's  $\chi^2$ ) in order to avoid cases with  $n_i$  equal to zero or very small.

Analytic solutions exist in a limited number of simple cases, e.g.: if  $\mu$  is a linear function. In most of the realistic cases, the  $\chi^2$  minimization is performed numerically, as for most of the unbinned

maximum likelihood fits. Binned fits are, in many cases, more convenient with respect to unbinned fits because the number of input variables decreases from the total number of entries to the number of bins. This leads usually to simpler and faster numerical implementations, in particular when unbinned fits become unpractical in cases of very large number of entries. Anyway, for limited number of entries, a fraction of the information is lost when moving from an unbinned to a binned sample and a possible loss of precision may occur.

The maximum value of the likelihood function obtained from an unbinned maximum likelihood fit doesn't in general provide information about the quality (*goodness*) of the fit. Instead, the minimum value of the  $\chi^2$  in a fit with a Gaussian underlying model is distributed according to a known PDF given by:

$$P(\chi^2; n) = \frac{2^{-n/2}}{\Gamma(n/2)} \chi^{n-2} e^{-\frac{\chi^2}{2}}, \quad (105)$$

where  $n$  is the *number of degrees of freedom*, equal to the number of bins minus the number of fit parameters. The cumulative distribution (Eq. (8)) of  $P(\chi^2; n)$  follows a uniform distribution between from 0 to 1, and it is an example of *p-value* (See Sec. 4). If the true PDF model deviates from the assumed distribution, the distribution of the *p-value* will be more peaked around zero instead of being uniformly distributed.

It's important to note that *p-values* are not the "probability of the fit hypothesis", because that would be a Bayesian probability, with a completely different meaning, and should be evaluated in a different way.

In case of a Poissonian distribution of the number of bin entries that may deviate from the Gaussian approximation, because of small number of entries, a better alternative to the Gaussian-inspired Neyman's or Pearson's  $\chi^2$  has been proposed by Baker and Cousins [13] using the following likelihood ratio as alternative to Eq. (103):

$$\chi_{\lambda}^2 = -2 \ln \prod_i \frac{L(n_i; \mu_i)}{L(n_i; n_i)} = -2 \ln \prod_i \frac{e^{-\mu_i} \mu_i^{n_i}}{n_i!} \frac{n_i!}{e^{-n_i} n_i^{n_i}} \quad (106)$$

$$= 2 \sum_i \left[ \mu_i(\theta_1, \dots, \theta_m) - n_i + n_i \ln \left( \frac{n_i}{\mu_i(\theta_1, \dots, \theta_m)} \right) \right]. \quad (107)$$

Equation (107) gives the same minimum value as the Poisson likelihood function, since a constant term has been added to the log-likelihood function in Eq. (103), but in addition it provides goodness-of-fit information, since it asymptotically obeys a  $\chi^2$  distribution with  $n - m$  degrees of freedom. This is due to Wilks' theorem, discussed in Sec. 5.6.

### 3.14 Combination of measurements

The simplest combination of two measurements can be performed when no correlation is present between them:

$$m = m_1 \pm \sigma_1, \quad (108)$$

$$m = m_2 \pm \sigma_2. \quad (109)$$

The following  $\chi^2$  can be built, assuming a Gaussian PDF model for the two measurements, similarly to Eq. (79):

$$\chi^2 = \frac{(m - m_1)^2}{\sigma_1^2} + \frac{(m - m_2)^2}{\sigma_2^2}. \quad (110)$$

The minimization of the  $\chi^2$  in Eq. (110) leads to the following equation:

$$0 = \frac{\partial \chi^2}{\partial m} = 2 \frac{(m - m_1)}{\sigma_1^2} + 2 \frac{(m - m_2)}{\sigma_2^2}, \quad (111)$$

which is solved by:

$$m = \hat{m} = \frac{\frac{m_1}{\sigma_1^2} + \frac{m_2}{\sigma_2^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}. \quad (112)$$

Eq. (112) can also be written in form of *weighted average*:

$$\hat{m} = \frac{w_1 m_1 + w_2 m_2}{w_1 + w_2}, \quad (113)$$

where the weights  $w_i$  are equal to  $\sigma_i^{-2}$ . The uncertainty on  $\hat{m}$  is given by:

$$\sigma_{\hat{m}}^2 = \frac{1}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}. \quad (114)$$

In case  $m_1$  and  $m_2$  are correlated measurements, the  $\chi^2$  changes from Eq. (110) to the following, including a non-null correlation coefficient  $\rho$ :

$$\chi^2 = \begin{pmatrix} m - m_1 & m - m_2 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}^{-1} \begin{pmatrix} m - m_1 \\ m - m_2 \end{pmatrix}. \quad (115)$$

In this case, the minimization of the  $\chi^2$  defined by Eq. (115) gives:

$$\hat{m} = \frac{m_1(\sigma_2^2 - \rho\sigma_1\sigma_2) + m_2(\sigma_1^2 - \rho\sigma_1\sigma_2)}{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2}, \quad (116)$$

with uncertainty given by:

$$\sigma_{\hat{m}}^2 = \frac{\sigma_1^2 \sigma_2^2 (1 - \rho)^2}{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2}. \quad (117)$$

This solution is also called best linear unbiased estimator (BLUE) [14] and can be generalized to more measurements. An example of application of the BLUE method is the world combination of the top-quark mass measurements at LHC and Tevatron [15].

It can be shown that, in case the uncertainties  $\sigma_1$  and  $\sigma_2$  are estimates that may depend on the assumed central value, a bias may arise, which can be mitigated by evaluating the uncertainties  $\sigma_1$  and  $\sigma_2$  at the central value obtained with the combination, then applying the BLUE combination, iteratively, until the procedure converges [16, 17].

Imagine we can write the two measurements as:

$$m = m_1 \pm \sigma_1' \pm \sigma_C, \quad (118)$$

$$m = m_2 \pm \sigma_2' \pm \sigma_C, \quad (119)$$

where  $\sigma_C^2 = \rho\sigma_1\sigma_2$ . This is the case where the two measurements are affected by a statistical uncertainty, which is uncorrelated between the two measurements, and a fully correlated systematic uncertainty. In those case, Eq. (116) becomes:

$$\hat{m} = \frac{\frac{m_1}{\sigma_1'^2} + \frac{m_2}{\sigma_2'^2}}{\frac{1}{\sigma_1'^2} + \frac{1}{\sigma_2'^2}}, \quad (120)$$

i.e.: it assumes again the form of a weighted average with weights  $w_i = \sigma_i'^{-2}$  computed on the uncorrelated uncertainty contributions. The uncertainty on  $\hat{m}$  is given by:

$$\sigma_{\hat{m}}^2 = \frac{1}{\frac{1}{\sigma_1'^2} + \frac{1}{\sigma_2'^2}} + \sigma_C^2, \quad (121)$$

which is the sum in quadrature of the uncertainty of the weighted average (Eq. (114)) and the common uncertainty  $\sigma_C$  [18].

In a more general case, we may have  $n$  measurements  $m_1, \dots, m_n$  with a  $n \times n$  covariance matrix  $C_{ij}$ . The expected values for  $m_1, \dots, m_n$  are  $M_1, \dots, M_n$  and may depend on some unknown parameters  $\vec{\theta} = (\theta_1, \dots, \theta_m)$ . For this case, the  $\chi^2$  to be minimized is:

$$\chi^2 = \sum_{i,j=1}^n (m_i - M_i(\vec{\theta})) C_{ij}^{-1} (m_j - M_j(\vec{\theta})) \quad (122)$$

$$= \begin{pmatrix} m_1 - M_1(\vec{\theta}) & \dots & m_n - M_n(\vec{\theta}) \end{pmatrix} \begin{pmatrix} C_{11} & \dots & C_{1n} \\ \vdots & \ddots & \vdots \\ C_{n1} & \dots & C_{nn} \end{pmatrix}^{-1} \begin{pmatrix} m_1 - M_1(\vec{\theta}) \\ \dots \\ m_n - M_n(\vec{\theta}) \end{pmatrix}. \quad (123)$$

An example of application of such a combination of measurement is given by fit of the Standard Model parameters using the electroweak precision measurements at colliders [19, 20].

#### 4 Hypothesis tests

Hypothesis testing addresses the question whether some observed data sample is more compatible with one theory model or another alternative one.

The terminology used in statistics may sometimes be not very natural for physics applications, but it has become popular among physicists as well as long as more statistical methods are becoming part of common practice. In a test, usually two hypotheses are considered:

- $H_0$ , the *null hypothesis*.  
Example 1: “a sample contains only background”.  
Example 2: “a particle is a pion”.
- $H_1$ , the *alternative hypothesis*.  
Example 1: “a sample contains background + signal”.  
Example 2: “a particle is a muon”.

A *test statistic* is a variable computed from our data sample that discriminates between the two hypotheses  $H_0$  and  $H_1$ . Usually it is a ‘summary’ of the information available in the data sample.

In physics it’s common to perform an event selection based on a discriminating variable  $x$ . For instance, we can take as signal sample all events whose value of  $x$  is above a threshold,  $x > x_{\text{cut}}$ .  $x$  is an example of *test statistic* used to discriminate between the two hypotheses,  $H_1 = \text{“signal”}$  and  $H_2 = \text{“background”}$ .

The following quantities are useful to give quantitative information about a test:

- $\alpha$ , the *significance level*: probability to reject  $H_0$  if  $H_0$  is assumed to be true (type I error, or false negative). In physics  $\alpha$  is equal to one minus the selection efficiency.
- $\beta$ , the *misidentification probability*, i.e.: probability to reject  $H_1$  if  $H_1$  is assumed to be true (type II error, or false negative).  $1 - \beta$  is also called *power of the test*.
- a  $p$ -value is the probability, assuming  $H_0$  to be true, of getting a value of the test statistic as result of our test at least as extreme as the observed test statistic.

In case of multiple discriminating variables, a selection of a signal against a background may be implemented in different ways. E.g.: applying a selection on each individual variable, or on a combination of those variables, or selecting an area of the multivariate space which is enriched in signal events.

#### 4.1 The Neyman–Pearson lemma

The Neyman–Pearson lemma [21] ensures that, for a fixed significance level ( $\alpha$ ) or equivalently a signal efficiency ( $1 - \alpha$ ), the selection that gives the lowest possible misidentification probability ( $\beta$ ) is based on a likelihood ratio:

$$\lambda(x) = \frac{L(x|H_1)}{L(x|H_0)} > k_\alpha, \quad (124)$$

where  $L(x|H_0)$  and  $L(x|H_1)$  are the values of the likelihood functions for the two considered hypotheses.  $k_\alpha$  is a constant whose value depends on the fixed significance level  $\alpha$ .

The likelihood function can't always be determined exactly. In cases where it's not possible to determine the exact likelihood function, other discriminators can be used as test statistics. Neural Networks, Boosted Decision Trees and other machine-learning algorithms are examples of discriminators that may closely approximate the performances of the exact likelihood ratio, approaching the Neyman–Pearson optimal performances [22].

In general, algorithms that provide a test statistic for samples with multiple variables are referred to as *multivariate discriminators*. Simple mathematical algorithms exist, as well as complex implementations based on extensive CPU computations. In general, the algorithms are ‘trained’ using input samples whose nature is known (*training samples*), i.e.: where either  $H_0$  or  $H_1$  is known to be true. This is typically done using data samples simulated with computer algorithms (Monte Carlo) or, when possible, with control samples obtained from data. Among the most common problems that arise with training of multivariate algorithms, the size of training samples is necessarily finite, hence the true distributions for the considered hypotheses can't be determined exactly from the training sample distribution. Moreover, the distribution assumed in the simulation of the input samples may not reproduce exactly the true distribution of real data, for instance because of systematic errors that affect our simulation.

#### 4.2 Projective likelihood ratio

In case of independent variables, the likelihood functions appearing in the numerator and denominator of Eq. (124) can be factorized as product of one-dimensional PDF (Eq. (30)). Even in the cases when variables are not independent, this can be taken as an approximate evaluation of the Neyman–Pearson likelihood ratio, so we can write:

$$\lambda(x) = \frac{L(x_1, \dots, x_n|H_1)}{L(x_1, \dots, x_n|H_0)} \simeq \frac{\prod_{i=1}^n f_i(x_i|H_1)}{\prod_{i=1}^n f_i(x_i|H_0)}. \quad (125)$$

The approximation may be improved if a proper rotation is first applied to the input variables in order to eliminate their correlation. This approach is called *principal component analysis*.

#### 4.3 Fisher discriminant

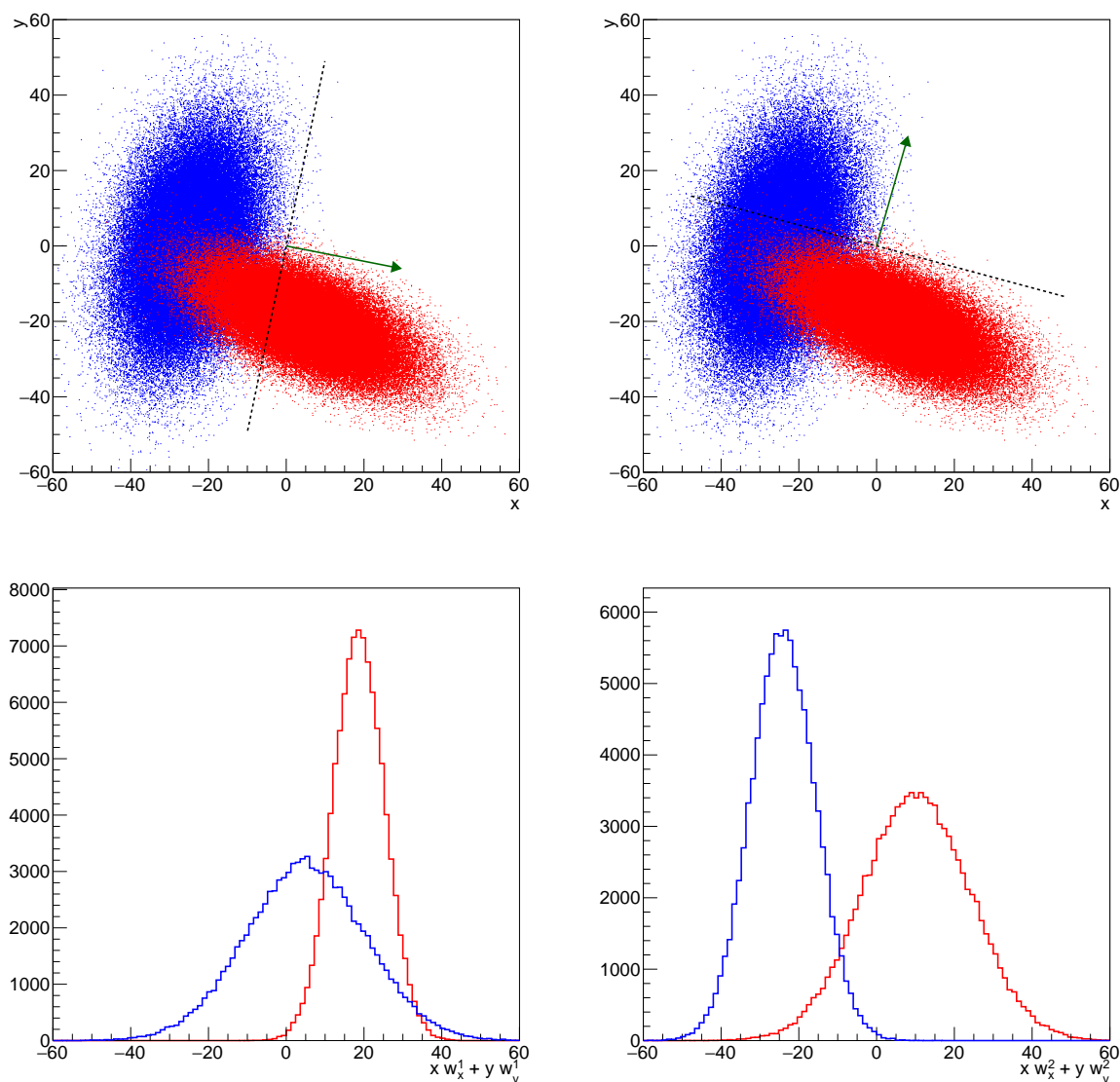
Fisher [23] introduced a discriminator based on a linear combination of input variables that maximizes the distance of the means of the two classes while minimizing the variance, projected along a direction  $\mathbf{w}$ :

$$J(\mathbf{w}) = \frac{|\mu_0 - \mu_1|^2}{\sigma_0^2 + \sigma_1^2} = \frac{\mathbf{w}^T \cdot (\mathbf{m}_0 - \mathbf{m}_1)}{\mathbf{w}^T (\boldsymbol{\Sigma}_0 + \boldsymbol{\Sigma}_1) \mathbf{w}}. \quad (126)$$

The selection is achieved by requiring  $J(\mathbf{w}) > J_{\text{cut}}$ , which determines an hyperplane perpendicular to  $\mathbf{w}$ . Examples of two different projections for a two-dimensional case is shown in Fig. 16. The problem of maximising  $J(\mathbf{w})$  over all possible directions  $\mathbf{w}$  can be solved analytically using linear algebra.

#### 4.4 Artificial Neural Network

Artificial Neural Networks (ANN) are computer implementations of simplified models of how neuron cells work. The schematic structure of an ANN is shown in Fig. 17. Each node in the network receives



**Fig. 16:** Examples of Fisher projections. Two samples are distributed according to the red and blue distributions in two dimensions and two possible projection direction  $\mathbf{w}$  are shown as dotted line, the green arrows are perpendicular to them (top plots). The corresponding one-dimensional projections along the chosen direction show different overlap between the red and blue distribution (bottom plots), depending on the choice of the projection.

inputs from either the input variables (input layer) or from the previous layer, and provides an output either of the entire network (output layer) or which is used as input to next layer. Within a node, inputs are combined linearly with proper weights that are different for each of the nodes. Each output is then transformed using a sigmoid function  $\varphi$ :

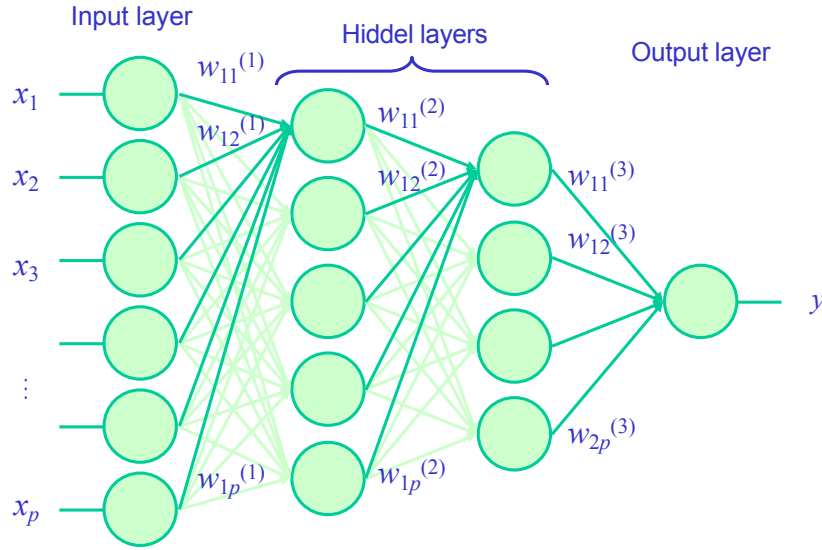
$$y^{(n)}(\vec{x}) = \varphi \left( \sum_{j=1}^p w_{kj}^{(n)} x_j \right), \quad (127)$$

where  $\varphi$  is typically:

$$\varphi(\nu) = \frac{1}{1 + e^{-\lambda\nu}}, \quad (128)$$

so that the output values are bound between 0 and 1.





**Fig. 17:** Structure of an Artificial Neural Network.

In order to find the optimal set of network weights  $w_{ij}^{(n)}$ , a minimization is performed on the *loss function* defined as the following sum over a training sample of size  $N$ :

$$L(w) = \sum_{i=1}^N (y_i^{\text{true}} - y(\vec{x}_i))^2, \quad (129)$$

$y_i^{\text{true}}$  being usually equal to 1 for signal ( $H_1$ ) and 0 for background ( $H_0$ ). Iteratively, weights are modified (*back propagation*) for each training event (or each group of training events) using the *stochastic gradient descent* technique:

$$w_{ij} \rightarrow w_{ij} - \eta \frac{\partial L(w)}{\partial w_{ij}}. \quad (130)$$

The parameter  $\eta$  controls the learning rate of the network. Variations of the training implementation exist.

Though it can be proven [24] that, under some regularity conditions, neural networks with a single hidden layer can approximate any analytical function with a sufficiently high number of neurons, in practice this limit is hard to achieve. Networks with several hidden layers can better manage complex variables combinations, e.g.: exploiting invariant mass distributions features using only four-vectors as input [25]. Those complex implementation that were almost intractable in the past can now be better approached thanks to the availability of improved training algorithms and more easily available CPU power.

#### 4.5 Boosted Decision Trees

A *decision tree* is a sequence of simple cuts that are sequentially applied on events in a data sample. Each cut splits the sample into nodes that may be further split by the application of subsequent cuts. Nodes where signal or background is largely dominant are classified as leaves. Alternatively, the splitting may stop if too few events per node remain, or if the total number of nodes too high. Each branch on the tree represents one sequence of cuts. Cuts can be optimized in order to achieve the best split level. One possible implementation is to maximize for each node the gain of Gini index after a splitting:

$$G = P(1 - P), \quad (131)$$

where  $P$  is the purity of the node (i.e.: the fraction of signal events).  $G$  is equal to zero for nodes containing only signal or background events. Alternative metrics can be used (e.g.: the *cross entropy*, equal to:  $-(P \ln P + (1 - P) \ln(1 - P))$ ) in place of the Gini index.

An optimized single decision tree does not usually provide optimal performances or stability, hence multiple decision trees are usually combined. Each tree is added iteratively after weights are applied to test events. *Boosting* is achieved by iteratively reweighting the events in the training sample according to the classifier output in the previous iteration. The *boosted decision tree* (BDT) algorithm usually proceeds as follows:

- Events are reweighted using the previous iteration’s classifier result.
- A new tree is build and optimized using the reweighted events as training sample.
- A score is given to each tree.
- The final BDT classifier result is a weighted average over all trees:

$$y(\vec{x}) = \sum_{k=1}^{N_{\text{trees}}} w_k C^{(k)}(\vec{x}). \quad (132)$$

One of the most popular algorithm is the *adaptive boosting* [26]: misclassified events only are reweighted according to the fraction of classification error of the previous tree:

$$\frac{1-f}{f}, f = \frac{N_{\text{misclassified}}}{N_{\text{tot}}}. \quad (133)$$

The weights applied to each tree are also related to the misclassification fraction:

$$y(\vec{x}) = \sum_{k=1}^{N_{\text{trees}}} \ln \left( \frac{1-f^{(k)}}{f^{(k)}} \right) C^{(k)}(\vec{x}). \quad (134)$$

This algorithm enhances the weight of events misclassified on the previous iteration in order to improve the performance on those events. Further variations and more algorithms are available.

#### 4.6 Overtraining

Algorithms may learn too much from the training sample, exploiting features that are only due to random fluctuations. It may be important to check for overtraining comparing the discriminator’s distributions for the training sample and for an independent *test sample*: compatible distributions will be an indication that no overtraining occurred.

### 5 Discoveries and upper limits

The process towards a discovery, from the point of view of data analysis, proceeds starting with a test of our data sample against two hypotheses concerning the theoretical underlying model:

- $H_0$ : the data are described by a model that contains background only;
- $H_1$ : the data are described by a model that contains a new signal plus background.

The discrimination between the two hypotheses can be based on a test statistic  $\lambda$  whose distribution is known under the two considered hypotheses. We may assume that  $\lambda$  tends to have (conventionally) large values if  $H_1$  is true and small values if  $H_0$  is true. This convention is consistent with using as test statistic the likelihood ratio  $\lambda = L(x|H_1)/L(x|H_0)$ , as in the Neyman–Pearson lemma (Eq. (124)). Under the frequentist approach, it’s possible to compute a  $p$ -value equal to the probability that  $\lambda$  is greater or equal

to than the value  $\lambda^{\text{obs}}$  observed in data. Such  $p$ -value is usually converted into an equivalent probability computed as the area under the rightmost tail of a standard normal distribution:

$$p = \int_Z^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 1 - \Phi(Z), \quad (135)$$

where  $\Phi$  is the cumulative (Eq. (8)) of a standard normal distribution.  $Z$  in Eq. (135) is called *significance level*. In literature conventionally a signal with a significance of at least 3 ( $3\sigma$  level) is claimed as *evidence*. It corresponds to a  $p$ -value of  $1.35 \times 10^{-3}$  or less. If the significance exceeds 5 ( $5\sigma$  level), i.e.: the  $p$ -value is below  $2.9 \times 10^{-7}$ , one is allowed to claim the *observation* of the new signal. It's worth noting that the probability that background produces a large test statistic is not equal to the probability of the null hypothesis (background only), which has only a Bayesian sense.

Finding a large significance level, anyway, is only part of the discovery process in the context of the scientific method. Below a sentence is reported from a recent statement of the American Statistical Association:

*The  $p$ -value was never intended to be a substitute for scientific reasoning. Well-reasoned statistical arguments contain much more than the value of a single number and whether that number exceeds an arbitrary threshold. The ASA statement is intended to steer research into a 'post  $p < 0.05$  era' [27].*

This was also remarked by the physicists community, for instance by Cowan *et al.*:

*It should be emphasized that in an actual scientific context, rejecting the background-only hypothesis in a statistical sense is only part of discovering a new phenomenon. One's **degree of belief** that a new process is present will depend in general on other factors as well, such as the plausibility of the new signal hypothesis and the degree to which it can describe the data [28].*

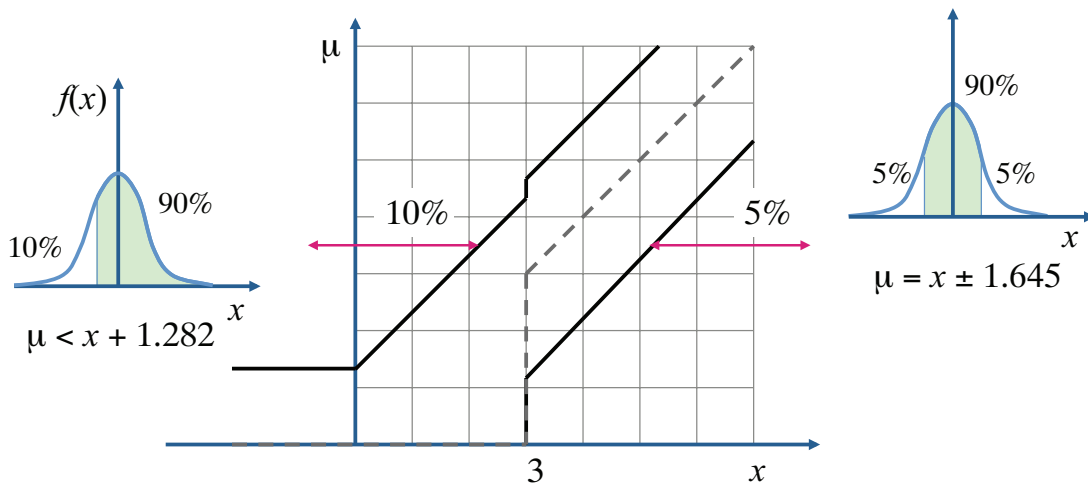
## 5.1 Upper limits

Upper limits measure the amount of excluded region in the theory's parameter space resulting from our negative results of a search for a new signal. Upper limits are obtained by building a fully asymmetric Neyman confidence belt (Sec. 3.8) based on the considered test statistic (Fig. 8). The belt can be inverted in order to find the allowed fully asymmetric interval for the signal yield  $s$ . The upper limit  $s^{\text{up}}$  is the upper extreme of the asymmetric confidence interval  $[0, s^{\text{up}}]$ . In case the considered test statistic is a discrete variable (e.g.: the number of selected events  $n$  in a counting experiments), the coverage may not be exact, as already discussed in Sec 3.9.

The procedure described above to determine upper limits, anyway, may incur the so-called *flip-flopping* issue [29]: given an observed result of our test statistic, when should we quote a central value or upper limit? A choice that is sometimes popular in scientific literature is to quote a (90% or 95% CL) upper limit if the significance is below  $3\sigma$  or to quote a central value if the significance is at least  $3\sigma$ . The choice to “flip” from an upper limit to a central value can be demonstrated to produce an incorrect coverage. This can be visually seen in Fig. 18, where a Gaussian belt at 90% CL is constructed, similarly to Fig. 12 (where instead a 68.3% CL was used). In Fig. 18, for  $x < 3$ , anyway, the belt is modified because those values correspond to a significance below the  $3\sigma$  level, where an upper limit (a fully asymmetric confidence interval) was chosen. As shown with the red arrows in the figures, there are intervals in  $x$  that, in this way, have a coverage reduced to 85% instead of the required 90%.

## 5.2 Feldman–Cousins intervals

A solution to the flip-flopping problem was developed by Feldman and Cousins [29]. They proposed to select confidence interval based on a likelihood-ratio criterion. Given a value  $\theta = \theta_0$  of the parameter of



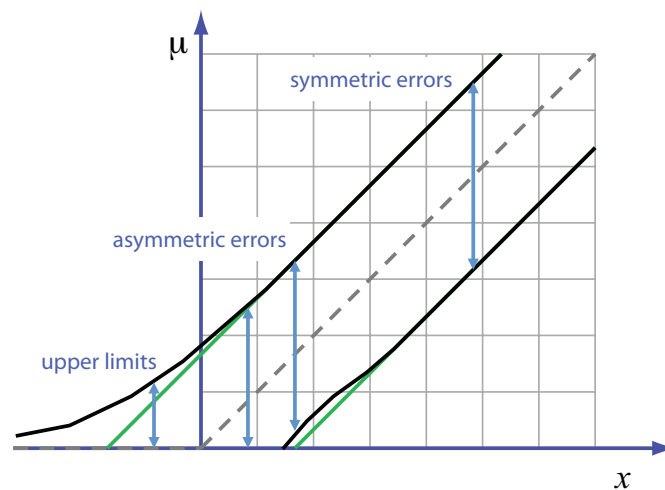
**Fig. 18:** Illustration of flip-flopping with a Gaussian confidence belt. This figure is taken from Ref. [7].

interest, the Feldman–Cousins confidence interval is defined as:

$$R_\mu = \left\{ x : \frac{L(x; \theta_0)}{L(x; \hat{\theta})} > k_\alpha \right\}, \quad (136)$$

where  $\hat{\theta}$  is the best-fit value for  $\theta$  and the constant  $k_\alpha$  should be set in order to ensure the desired confidence level  $1 - \alpha$ .

An example of the confidence belt computed with the Feldman–Cousins approach is shown in Fig. 19 for the Gaussian case illustrated in Fig. 18. With the Feldman–Cousins approach, the confidence



**Fig. 19:** Feldman–Cousins belt for a Gaussian case. This figure is taken from Ref. [7].

interval smoothly changes from a fully asymmetric one, which leads to an upper limit, for low values of  $x$ , to an asymmetric interval for higher values of  $x$  interval, then finally a symmetric interval (to a very good approximation) is obtained for large values of  $x$ , recovering the usual result as in Fig. 18.

Even for the simplest Gaussian case, the computation of Feldman–Cousins intervals requires numerical treatment and for complex cases their computation may be very CPU intensive.

### 5.3 Upper limits for event counting experiments

The simplest search for a new signal consists of counting the number of events passing a specified selection. The number of selected events  $n$  is distributed according to a Poissonian distribution where the expected value, in case of presence of signal plus background ( $H_1$ ) is  $s + b$ , and for background only ( $H_0$ ) is  $b$ . Assume we count  $n$  events, we then want to compare the two hypotheses  $H_1$  and  $H_0$ . As simplest case, we can assume that  $b$  is known with negligible uncertainty. If not, uncertainty on its estimate must be taken into account.

The likelihood function for this case can be written as:

$$L(n; s) = \frac{(s + b)^n}{n!} e^{-(s+b)}. \tag{137}$$

$H_0$  corresponds to the case  $s = 0$ .

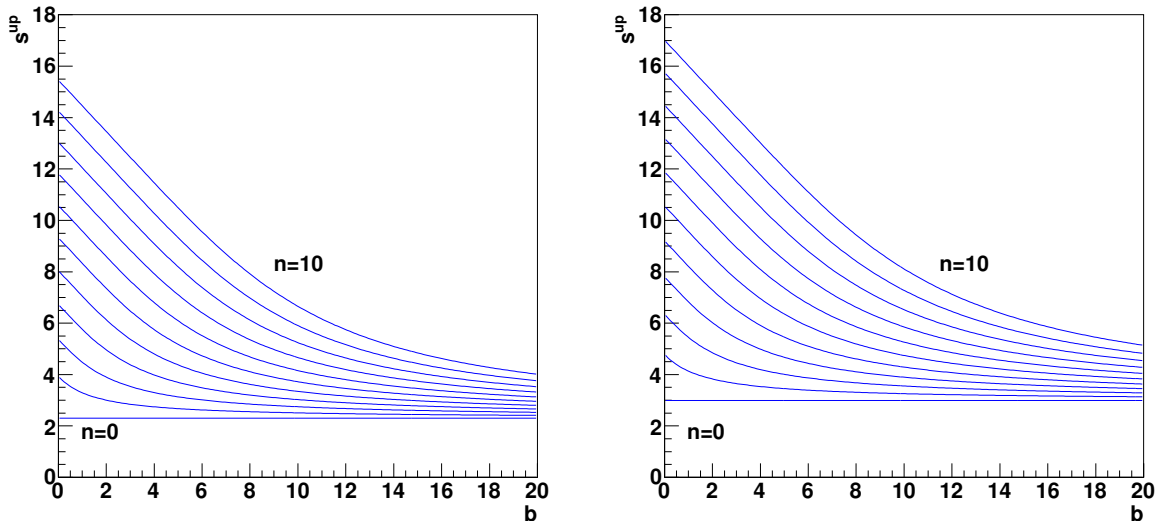
Using the Bayesian approach, an upper limit  $s^{\text{up}}$  on  $s$  can be determined by requiring that the posterior probability corresponding to the interval  $[0, s^{\text{up}}]$  is equal to the confidence level  $1 - \alpha$ :

$$1 - \alpha = \int_0^{s^{\text{up}}} P(s|n) ds = \frac{\int_0^{s^{\text{up}}} L(n; a)\pi(s) ds}{\int_0^{+\infty} L(n; a)\pi(s) ds}. \tag{138}$$

The choice of a uniform prior,  $\pi(s) = 1$ , simplifies the computation and Eq. (138) reduces to [30]:

$$\alpha = e^{-s^{\text{up}}} \frac{\sum_{m=0}^n \frac{(s^{\text{up}} + b)^m}{m!}}{\sum_{m=0}^n \frac{b^m}{m!}}. \tag{139}$$

Upper limits obtained with Eq. (139) are shown in Fig. 20. In the case  $b = 0$ , the results obtained in



**Fig. 20:** Bayesian upper limits at 90% (left) and 95% (right) CL for the expected signal yield  $s$  for a counting experiment with different level of expected background  $b$ . This figure is taken from Ref. [7].

Eq. (60) and (61) are again recovered.

Frequentist upper limits for a counting experiment can be easily computed in case of negligible background ( $b = 0$ ). If zero events are observed ( $n = 0$ ), the likelihood function simplifies to:

$$L(n = 0; s) = \text{Pois}(0; s) = e^{-s}. \quad (140)$$

The inversion of the fully asymmetric Neyman belt reduces to:

$$P(n \leq 0; s^{\text{UP}}) = P(n = 0; s^{\text{UP}}) = \alpha \implies s^{\text{UP}} = -\ln \alpha, \quad (141)$$

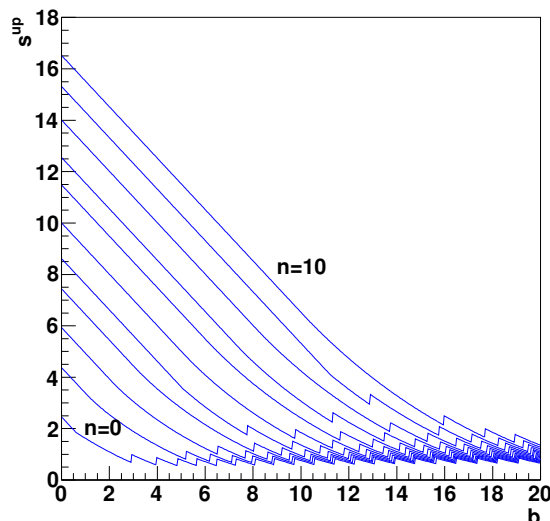
which lead to results that are numerically identical to the Bayesian computation:

$$s < s^{\text{UP}} = 2.303 \quad \text{for } \alpha = 0.1 \text{ (90\% CL)}, \quad (142)$$

$$s < s^{\text{UP}} = 2.996 \quad \text{for } \alpha = 0.05 \text{ (95\% CL)}. \quad (143)$$

In spite of the numerical coincidence, the interpretation of frequentist and Bayesian upper limits remain very different.

Upper limits from Eq. (142) and (143) anyway suffer from the flip-flopping problem and the coverage is spoiled when deciding to switch from an upper limit to a central value depending on the observed significance level. Feldman–Cousins intervals cure the flip-flopping issue and ensure the correct coverage (or may overcover for discrete variables). Upper limits at 90% computed with the Feldman–Cousins approach for a counting experiment are reported in Fig. 21. The “ripple” structure is due to the discrete nature of Poissonian counting. It’s evident from the figure that, even for  $n = 0$ , the upper



**Fig. 21:** Feldman–Cousins upper limits at 90% CL for a counting experiment. This figure is taken from Ref. [7].

limit decrease as  $b$  increases (apart from ripple effects). This means that if two experiment are designed for an expected background of –say– 0.1 and 0.01, the “worse” experiment (i.e.: the one which expects 0.1 events) achieves the best upper limit in case no event is observed ( $n = 0$ ), which is the most likely outcome if no signal is present. This feature was noted in the 2001 edition of the PDG [31]

*The intervals constructed according to the unified procedure [Feldman–Cousins] for a Poisson variable  $n$  consisting of signal and background have the property that for  $n = 0$  observed events, the upper limit decreases for increasing expected background. This is counter-intuitive, since it is known that if  $n = 0$  for the experiment in question, then no background was observed, and therefore one may argue that the expected background should not be relevant. The extent to which one should regard this feature as a drawback is a subject of some controversy.*

This counter-intuitive feature of frequentist upper limits is one of the reasons that led to the use in High-Energy Physics of a modified approach, whose main feature is that it also prevents rejecting cases where the experiment has little sensitivity due to statistical fluctuation, as will be described in next Section.

#### 5.4 The modified frequentist approach

A *modified frequentist approach* [32] was proposed for the first time for the combination of the results of searches for the Higgs boson by the four LEP experiments, ALEPH, DELPHI, L3 and OPAL [33]. Given a test statistic  $\lambda(x)$  that depends on some observation  $x$ , its distribution should be determined under the two hypotheses  $H_1$  (signal plus background) and  $H_0$  (background only). The following  $p$ -values can be used, where we assume that the test statistic  $\lambda$  tends to have small values for  $H_1$  and larger values for  $H_0$ :

$$p_{s+b} = P(\lambda(x|H_1) \geq \lambda^{\text{obs}}), \quad (144)$$

$$p_b = P(\lambda(x|H_0) \leq \lambda^{\text{obs}}). \quad (145)$$

$p_{s+b}$  and  $p_b$  can be interpreted as follows:

- $p_{s+b}$  is the probability to obtain a result which is less compatible with the signal than the observed result, assuming the signal hypothesis;
- $p_b$  is the probability to obtain a result less compatible with the background-only hypothesis than the observed one, assuming background only.

Instead of requiring, as for a frequentist upper limit,  $p_{s+b} \leq \alpha$ , the modified approach introduces a new quantity,  $\text{CL}_s$ , defined as:

$$\boxed{\text{CL}_s = \frac{p_{s+b}}{1 - p_b}}, \quad (146)$$

and the upper limit is set by requiring  $\text{CL}_s \leq \alpha$ . For this reason, the modified frequentist approach is also called “ $\text{CL}_s$  method”.

In practice, in most of the realistic cases,  $p_b$  and  $p_{s+b}$  are computed from simulated pseudoexperiments (*toy Monte Carlo*) by approximating the probabilities defined in Eq. (144, 145) with the fraction of the total number of pseudoexperiments satisfying their respective condition:

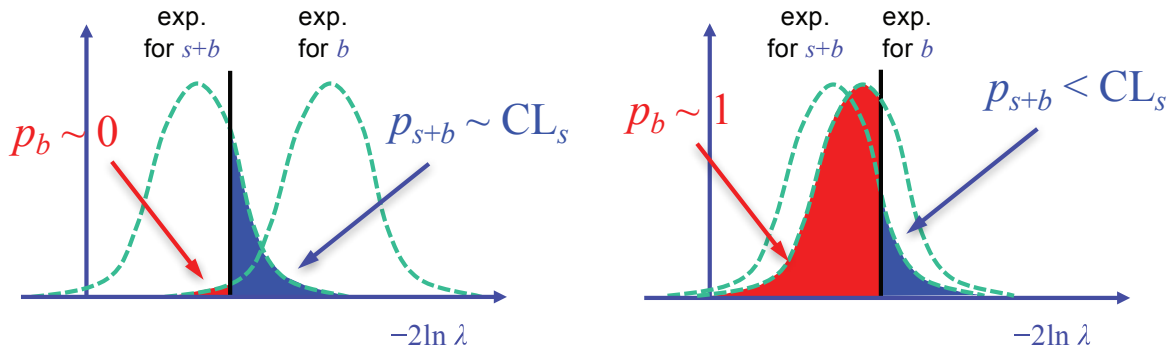
$$\text{CL}_s = \frac{p_{s+b}}{1 - p_b} = \frac{N(\lambda_{s+b} \geq \lambda^{\text{obs}})}{N(\lambda_b \geq \lambda^{\text{obs}})}. \quad (147)$$

Since  $1 - p_b \leq 1$ , then  $\text{CL}_s \geq p_{s+b}$ , hence upper limits computed with the  $\text{CL}_s$  method are always *conservative*.

In case the distributions of the test statistic  $\lambda$  (or equivalently  $-2 \ln \lambda$ ) for the two hypotheses  $H_0$  and  $H_1$  are well separated (Fig. 22, left), if  $H_1$  is true, then  $p_b$  will have a very high chance to be very small, hence  $1 - p_b \simeq 1$  and  $\text{CL}_s \simeq p_{s+b}$ . In this case  $\text{CL}_s$  and the purely frequentist upper limits coincide. If the two distributions instead largely overlap (Fig. 22, right), indicating that the experiment has poor sensitivity on the signal, in case  $p_b$  is large, because of a statistical fluctuation, then  $1 - p_b$  becomes small. This prevents  $\text{CL}_s$  to become too small, i.e.: it prevents to reject cases where the experiment has little sensitivity.

If we apply the  $\text{CL}_s$  method to the previous counting experiment, using the observed number of events  $n^{\text{obs}}$  as test statistic, then  $\text{CL}_s$  can be written, considering that  $n$  tends to be large in case of  $H_1$ , for this case, as:

$$\text{CL}_s = \frac{P(n \leq n^{\text{obs}} | s + b)}{P(n \leq n^{\text{obs}} | b)}. \quad (148)$$



**Fig. 22:** Illustration of the application of the  $CL_s$  method in cases of well separated distributions of the test statistic  $-2 \ln \lambda$  for the  $s + b$  and  $b$  hypotheses (left) and in case of largely overlapping distributions (right).

Expliciting the Poisson distribution, the computation gives the same result as for the Bayesian case with a uniform prior (Eq. (139)). In many cases, the  $CL_s$  upper limits give results that are very close, numerically, to Bayesian computations performed assuming a uniform prior. Of course, this does not allow to interpret  $CL_s$  upper limits as Bayesian upper limits. Concerning the interpretation of  $CL_s$ , it's worth reporting from Ref [32] the following statements:

*A specific modification of a purely classical statistical analysis is used to avoid excluding or discovering signals which the search is in fact not sensitive to.*

*The use of  $CL_s$  is a conscious decision not to insist on the frequentist concept of full coverage (to guarantee that the confidence interval doesn't include the true value of the parameter in a fixed fraction of experiments).*

*Confidence intervals obtained in this manner do not have the same interpretation as traditional frequentist confidence intervals nor as Bayesian credible intervals.*

## 5.5 Treatment of nuisance parameters

Nuisance parameters have been introduced in Sec. 3.5.1. Usually, signal extraction procedures (either parameter fits or upper limits determinations) determine, together with parameters of interest, also nuisance parameters that model effects not strictly related to our final measurement, like background yield and shape, detector resolution, etc. Nuisance parameters are also used to model sources of systematic uncertainties. Often, the true value of a nuisance parameter is not known, but we may have some estimate from sources that are external to our problem. In those cases, we can refer to *nominal values* of the nuisance parameter and their uncertainty. Nominal values of nuisance parameters are random variables distributed according to some PDF that depend on their true value.

A Gaussian distribution is the simplest assumption for nominal values of nuisance parameters. Anyway, this may give negative values corresponding to the leftmost tail, which are not suitable for non-negative quantities like cross sections. For instance, we may have an estimate of some background yield  $b$  given by:

$$b = \beta \sigma_b L_{\text{int}}, \quad (149)$$

where  $L_{\text{int}}$  is the estimate of the integrated luminosity (assumed to be known with negligible uncertainty),  $\sigma_b$  is the background cross section evaluated from theory, and  $\beta$  is a nuisance parameter, whose nominal value is equal to one, representing the uncertainty on the cross-section evaluation. If the uncertainty on  $\beta$



is large, one may have a negative value of  $\beta$  with non-negligible probability, hence an unphysical negative value of the background yield  $b$ . A safer assumption in such cases is to take a log normal distribution for the uncertain non-negative quantities:

$$b = e^\beta \sigma_b L_{\text{int}}, \quad (150)$$

where  $\beta$  is again distributed according to a normal distribution with nominal value equal to zero, in this case.

Under the Bayesian approach, nuisance parameters don't require a special treatment. If we have a parameter of interest  $\mu$  and a nuisance parameter  $\theta$ , a Bayesian posterior will be obtained as (Eq. (48)):

$$P(\mu, \theta | \vec{x}) = \frac{L(\vec{x}; \mu, \theta) \pi(\mu, \theta)}{\int L(\vec{x}; \mu', \theta') \pi(\mu', \theta') d\mu' d\theta'}. \quad (151)$$

$P(\mu | \vec{x})$  can be obtained as marginal PDF of  $\mu$  by integrating  $P(\mu, \theta | \vec{x})$  over  $\theta$ :

$$P(\mu | \vec{x}) = \int P(\mu, \theta | \vec{x}) d\theta = \frac{\int L(\vec{x}; \mu, \theta) \pi(\mu, \theta) d\theta}{\int \int L(\vec{x}; \mu', \theta') \pi(\mu', \theta') d\mu' d\theta'}. \quad (152)$$

In the frequentist approach, one possibility is to introduce in the likelihood function a model for a data sample that can constrain the nuisance parameter. Ideally, we may have a control sample  $\vec{y}$ , complementary to the main data sample  $\vec{x}$ , that only depends on the nuisance parameter, and we can write a global likelihood function as:

$$L(\vec{x}, \vec{y}; \mu, \theta) = L_x(\vec{x}; \mu, \theta) L_y(\vec{y}; \theta). \quad (153)$$

Using control regions to constrain nuisance parameters is usually a good method to reduce systematic uncertainties. Anyway, it may not always be feasible and in many cases we may just have information about the nominal value  $\theta^{\text{nom}}$  of  $\theta$  and its distribution obtained from a complementary measurement:

$$L(\vec{x}, \vec{y}; \mu, \theta) = L_x(\vec{x}; \mu, \theta) L_\theta(\theta^{\text{nom}}; \theta). \quad (154)$$

$L_\theta$  may be a Gaussian or log normal distribution in the easiest cases.

In order to achieve a likelihood function that does not depend on nuisance parameters, for many measurements at LEP or Tevatron a method proposed by Cousins and Highland was adopted [34] which consists of integrating the likelihood function over the nuisance parameters, similarly to what is done in the Bayesian approach (Eq. (152)). For this reason, this method was also called hybrid. Anyway the Cousins–Highland does not guarantee to provide exact coverage, and was often used as pragmatic solution in the frequentist context.

## 5.6 Profile likelihood

Most of the recent searches at LHC use the so-called *profile likelihood* approach for the treatment of nuisance parameters [28]. The approach is based on the test statistic built as the following likelihood ratio:

$$\lambda(\mu) = \frac{L(\vec{x}; \mu, \hat{\theta}(\mu))}{L(\vec{x}; \hat{\mu}, \hat{\theta})}, \quad (155)$$

where in the denominator both  $\mu$  and  $\theta$  are fit simultaneously as  $\hat{\mu}$  and  $\hat{\theta}$ , respectively, and in the numerator  $\mu$  is fixed, and  $\hat{\theta}(\mu)$  is the best fit of  $\theta$  for the fixed value of  $\mu$ . The motivation for the choice of Eq. (155) as the test statistic comes from Wilks' theorem that allows to approximate asymptotically  $-2 \ln \lambda(\mu)$  as a  $\chi^2$  [35].

In general, Wilks' theorem applies if we have two hypotheses  $H_0$  and  $H_1$  that are *nested*, i.e.: they can be expressed as sets of nuisance parameters  $\vec{\theta} \in \Theta_0$  and  $\vec{\theta} \in \Theta_1$ , respectively, such that  $\Theta_0 \subseteq \Theta_1$ . Given the likelihood function:

$$L = \prod_{i=1}^N L(\vec{x}_i, \vec{\theta}), \quad (156)$$

if  $H_0$  and  $H_1$  are nested, then the following quantity, for  $N \rightarrow \infty$ , is distributed as a  $\chi^2$  with a number of degrees of freedom equal to the difference of the  $\Theta_1$  and  $\Theta_0$  dimensionalities:

$$\chi_r^2 = -2 \ln \frac{\sup_{\vec{\theta} \in \Theta_0} \prod_{i=1}^N L(\vec{x}_i; \vec{\theta})}{\sup_{\vec{\theta} \in \Theta_1} \prod_{i=1}^N L(\vec{x}_i; \vec{\theta})}. \quad (157)$$

In case of a search for a new signal where the parameter of interest is  $\mu$ ,  $H_0$  corresponds to  $\mu = 0$  and  $H_1$  to any  $\mu \geq 0$ , Eq. (157) gives:

$$\chi_r^2(\mu) = -2 \ln \frac{\sup_{\vec{\theta}} \prod_{i=1}^N L(\vec{x}_i; \mu, \vec{\theta})}{\sup_{\mu', \vec{\theta}} \prod_{i=1}^N L(\vec{x}_i; \mu', \vec{\theta})}. \quad (158)$$

Considering that the supremum is equivalent to the best fit value, the profile likelihood defined in Eq. (155) is obtained.

As a concrete example of application of the profile likelihood, consider a signal with a Gaussian distribution over a background distributed according to an exponential distribution. A pseudoexperiment that was randomly-extracted according to such a model is shown in Fig. 23, where a signal yield  $s = 40$  was assumed on top of a background yield  $b = 100$ , exponentially distributed in the range of the random variable  $m$  from 100 to 150 GeV. The signal was assumed centered at 125 GeV with a standard deviation of 6 GeV, reminding the Higgs boson invariant mass spectrum. The signal yields  $s$  is fit from data. All parameters in the model are fixed, except the background yield, which is assumed to be known with some level of uncertainty modeled with a log normal distribution whose corresponding nuisance parameter is called  $\beta$ . The likelihood function for the model, which only depends on two parameters,  $s$  and  $\beta$ , is, in case of a single measurement  $m$ :

$$L(m; s, \beta) = L_0(m; s, b_0 = be^\beta) L_\beta(\beta; \sigma_\beta), \quad (159)$$

where:

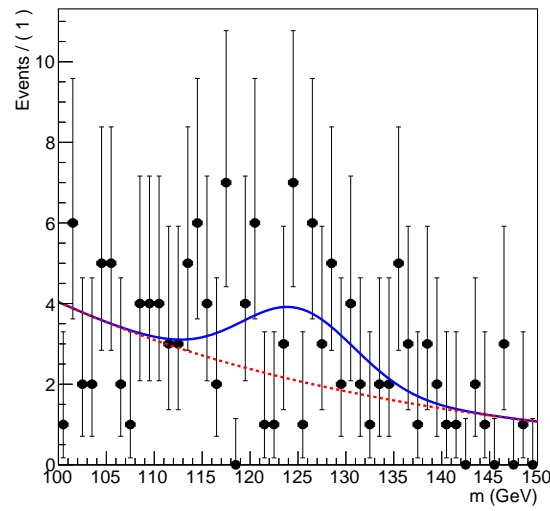
$$L_0(m; s, b_0) = \frac{e^{-(s+b_0)}}{n!} \left( s \frac{1}{\sqrt{2\pi}\sigma} e^{-(m-\mu)^2/2\sigma^2} + b_0 \lambda e^{-\lambda m} \right), \quad (160)$$

$$L_\beta(\beta; \sigma_\beta) = \frac{1}{\sqrt{2\pi}\sigma_\beta} e^{-\beta^2/2\sigma_\beta^2}. \quad (161)$$

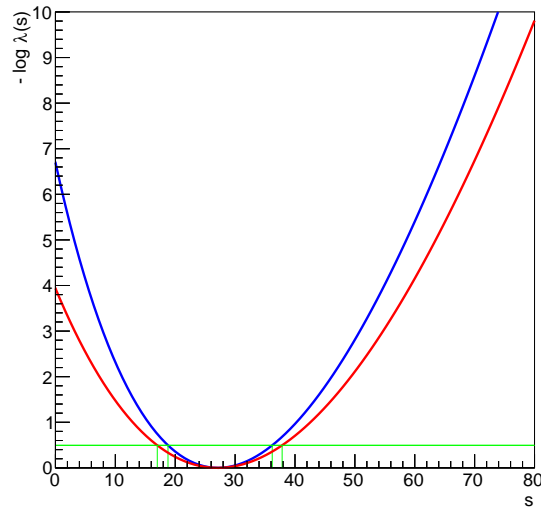
If we measure a set values  $\vec{m} = (m_1, \dots, m_N)$ , the likelihood function is:

$$L(\vec{m}; s, \beta) = \prod_{i=1}^N L(m_i; s, \beta). \quad (162)$$

The scan of  $-\ln \lambda(s)$  is shown in Fig. 24, where the profile likelihood was evaluated assuming  $\sigma_\beta = 0$



**Fig. 23:** Example of pseudoexperiment generated with a Gaussian-distributed signal over an exponential background. The assumed distribution for the background is the red dashed line, while the distribution for signal plus background is the blue solid line.



**Fig. 24:** Scan of the negative logarithm of the profile likelihood as a function of the signal yield  $s$ . The blue curve shows the profile likelihood curve defined assuming the background yield to be a constant (i.e.:  $b$  known without any uncertainty), the red curve shows the same curve defined with  $\sigma_\beta = 0.3$ . The green line at  $-\ln \lambda(s) = 0.5$  determines the uncertainty interval on  $s$ .

(no uncertainty on  $b$ , blue curve) or  $\sigma_b = 0.3$  (red curve). The minimum value of  $-\ln \lambda(s)$  is equal to zero, since at the minimum numerator and denominator in Eq. (155) are identical. Introducing the uncertainty on  $\beta$  (red curve) makes the curve broader. This causes an increase of the uncertainty on the estimate of  $s$ , whose uncertainty interval is obtained by intersecting the curve of the negative logarithm of the profile likelihood with an horizontal line at  $-\ln \lambda(s) = 0.5$  (green line in Fig. 24<sup>1</sup>).

In order to evaluate the significance of the observed signal, Wilks' theorem can be used. If we assume  $\mu = 0$  (null hypothesis), the quantity  $q_0 = -2 \ln \lambda(0)$  can be approximated with a  $\chi^2$  having one degree of freedom. Hence, the significance can be approximately evaluated as:

$$Z \simeq \sqrt{q_0}. \quad (163)$$

$q_0$  is twice the intercept of the curve in Fig. 24 with the vertical axis, and gives an approximate significance of  $Z \simeq \sqrt{2 \times 6.66} = 3.66$ , in case of no uncertainty on  $b$ , and  $Z \simeq \sqrt{2 \times 3.93} = 2.81$ , in case the uncertainty on  $b$  is considered. In this example, the effect of background yield uncertainty reduces the significance bringing it below the evidence level ( $3\sigma$ ). Those numerical values can be verified by running many pseudo experiments (toy Monte Carlo) assuming  $\mu = 0$  and computing the corresponding  $p$ -value. In complex cases, the computation of  $p$ -values using toy Monte Carlo may become unpractical, and Wilks' approximation provides a very convenient, and often rather precise, alternative calculation.

### 5.7 Variations on test statistics

A number of test statistics is proposed in Ref. [28] that better serve various purposes. Below the main ones are reported:

– **Test statistic for discovery:**

$$q_0 = \begin{cases} -2 \ln \lambda(0), & \hat{\mu} \geq 0, \\ 0, & \hat{\mu} < 0. \end{cases} \quad (164)$$

In case of a negative estimate of  $\mu$  ( $\hat{\mu} < 0$ ), the test statistic is set to zero in order to consider only positive  $\hat{\mu}$  as evidence against the background-only hypothesis. Within an asymptotic approximation, the significance is given by:  $Z \simeq \sqrt{q_0}$ .

– **Test statistic for upper limit:**

$$q_\mu = \begin{cases} -2 \ln \lambda(\mu), & \hat{\mu} \leq \mu, \\ 0, & \hat{\mu} > \mu. \end{cases} \quad (165)$$

If the  $\hat{\mu}$  estimate is larger than the assumed value for  $\mu$ , an upward fluctuation occurred. In those cases,  $\mu$  is not excluded by setting the test statistic to zero.

– **Test statistic for Higgs boson search:**

$$\tilde{q}_\mu = \begin{cases} -2 \ln \frac{L(\bar{x}|\mu, \hat{\theta}(\mu))}{L(\bar{x}|0, \hat{\theta}(0))}, & \hat{\mu} < 0, \\ -2 \ln \frac{L(\bar{x}|\mu, \hat{\theta}(\mu))}{L(\bar{x}|\mu, \hat{\theta}(\mu))}, & 0 \leq \hat{\mu} < \mu, \\ 0, & \hat{\mu} \geq \mu. \end{cases} \quad (166)$$

This test statistics both protects against unphysical cases with  $\mu < 0$  and, as the test statistic for upper limits, protects upper limits in cases of an upward  $\hat{\mu}$  fluctuation.

<sup>1</sup> The plot in Fig. 24 was generated with the library ROOSTATS in ROOT [36], which by default, uses  $-\ln \lambda$  instead of  $-2 \ln \lambda$ .

A number of measurements performed at LEP and Tevatron used a test statistic based on the ratio of the likelihood function evaluated under the signal plus background hypothesis and under the background only hypothesis, inspired by the Neyman–Pearson lemma:

$$q = -2 \ln \frac{L(\vec{x}|s+b)}{L(\vec{x}|b)}. \quad (167)$$

In many LEP and Tevatron analyses, nuisance parameters were treated using the hybrid Cousins–Hyghland approach. Alternatively, one could use a formalism similar to the profile likelihood, setting  $\mu = 0$  in the denominator and  $\mu = 1$  in the numerator, and minimizing the likelihood functions with respect to the nuisance parameters:

$$q = -2 \ln \frac{L(\vec{x}|\mu=1, \hat{\theta}(1))}{L(\vec{x}|\mu=0, \hat{\theta}(0))}. \quad (168)$$

For all the mentioned test statistics, asymptotic approximations exist and are reported in Ref. [28]. Those are based either on Wilks’ theorem or on Wald’s approximations [37]. If a value  $\mu$  is tested, and the data are supposed to be distributed according to another value of the signal strength  $\mu'$ , the following approximation holds, asymptotically:

$$-2 \ln \lambda(\mu) = \frac{(\mu - \hat{\mu})^2}{\sigma^2} + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right), \quad (169)$$

where  $\hat{\mu}$  is distributed according to a Gaussian with average  $\mu'$  and standard deviation  $\sigma$ . The covariance matrix for the nuisance parameters is given, in the asymptotic approximation, by:

$$V_{ij}^{-1} = \left\langle \frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j} \right\rangle \Bigg|_{\mu=\mu'}, \quad (170)$$

where  $\mu'$  is assumed as value for the signal strength.

In some cases, asymptotic approximations (Eq. (169)) can be written in terms of an *Asimov dataset* [38]:

*We define the Asimov data set such that when one uses it to evaluate the estimators for all parameters, one obtains the true parameter values* [28].

In practice, an Asimov dataset is a single “representative” dataset obtained by replacing all observable (random) variables with their expected value. In particular, all yields in the data sample (e.g.: in a binned case) are replaced with their expected values, that may be non integer values. The median significance for different cases of test statistics can be computed in this way without need of producing extensive sets of toy Monte Carlo. The implementation of those asymptotic formulate is available in the ROOSTATS library, released as part an optional component ROOT [36].

## 5.8 The look-elsewhere effect

When searching for a signal peak on top of a background that is smoothly distributed over a wide range, one can either know the position of the peak or not. One example in which the peak position is known is the search for a rare decay of a known particle, like  $B_s \rightarrow \mu^+ \mu^-$ . A case when the position was not know was the search for the Higgs boson, whose mass is not predicted by theory. In a case like the decay of a particle of known mass, it’s easy to compute the peak significance: from the distribution of

the test statistic  $f(q)$  computed assuming  $\mu = 0$  (background only), given the observed value of the test statistic  $q^{\text{obs}}$ , a  $p$ -value can be determined and then translated into a significance level:

$$p = \int_{q^{\text{obs}}}^{+\infty} f(q|\mu = 0) dq, \quad Z = \Phi^{-1}(1 - p). \quad (171)$$

In case, instead, the search is performed without knowing the position of the peak, Eq. (171) gives only a *local*  $p$ -value, which means it reflects the probability that a background fluctuation *at a given mass value*  $m$  gives a value of the test statistic greater than the observed one:

$$p(m) = \int_{q^{\text{obs}}(m)}^{+\infty} f(q|\mu = 0) dq. \quad (172)$$

The *global*  $p$ -value, instead, should quantify the probability that a background fluctuation *at any mass value* gives a value of the test statistic greater than the observed one.

The chance that an overfluctuation occurs for *at least* one mass value increases with the size of the search range, and the magnitude of the effect depends on the resolution.

One possibility to evaluate a global  $p$ -value is to let also  $m$  fluctuate in the test statistic:

$$\hat{q} = -2 \ln \frac{L(\mu = 0)}{L(\hat{\mu}; \hat{m})}. \quad (173)$$

Note that in the numerator  $L$  doesn't depend on  $m$  for  $\mu = 0$ . This is a case where Wilks' theorem doesn't apply, and no simple asymptotic approximations exist. The global  $p$ -value can be computed, in principle, as follows:

$$p^{\text{glob}} = \int_{\hat{q}^{\text{obs}}}^{+\infty} f(\hat{q}|\mu = 0) d\hat{q}_0. \quad (174)$$

The effect in practice can be evaluated with brute-force toy Monte Carlo:

- Produce a large number of pseudoexperiments simulating background-only samples.
- Find the maximum  $\hat{q}$  of the test statistic  $q$  in the entire search range.
- Determine the distribution of  $\hat{q}$ .
- Compute the global  $p$ -value as probability to have a value of  $\hat{q}$  greater than the observed one.

This procedure usually requires very large toy Monte Carlo samples in order to treat a discovery case: a  $p$ -value close to  $3 \times 10^{-7}$  ( $5\sigma$  level) requires a sample significantly larger than  $\sim 10^7$  entries in order to determine the  $p$ -value with small uncertainty.

An asymptotic approximation for the global  $p$ -value is given by the following inequation [39]<sup>2</sup>:

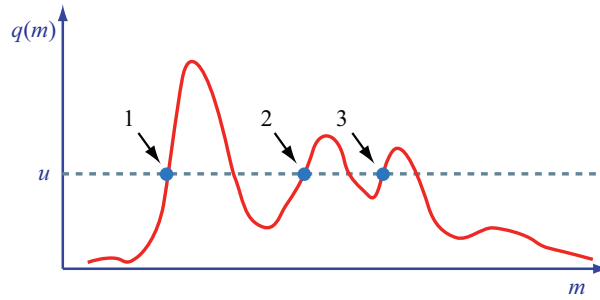
$$p^{\text{glob}} = P(\hat{q} > u) \leq \langle N_u \rangle + P(\chi^2 > u), \quad (175)$$

where  $P(\chi^2 > u)$  is a standard  $\chi^2$  probability and  $\langle N_u \rangle$  is the average number of *upcrossings* of the test statistic, i.e.: the average number of times that the curve  $q(m)$  crosses a given horizontal line at a level  $u$  with a positive derivative, as illustrated in Fig. 25.

The number of upcrossings may be very small for some values of  $u$ , but an approximate scaling law exists and allows to perform the computation at a more convenient level  $u_0$ :

$$\langle N_u \rangle = \langle N_{u_0} \rangle e^{-(u-u_0)/2}. \quad (176)$$

<sup>2</sup> In case of a test statistic for discovery  $q_0$  (Eq. (164)), the term  $P(\chi^2 > u)$  in Eq. (175) achieves an extra factor 1/2, which is usually not be present for other test statistics.



**Fig. 25:** Illustration of the evaluation of the number of upcrossing of a test statistic curve  $q(m)$ . The upcrossings are noted as 1, 2, and 3, hence the corresponding  $N_u$  is equal to 3.

So,  $\langle N_{u_0} \rangle$  can be more conveniently evaluated using a reasonable number of toy Monte Carlo generations, then it can be extrapolated following the exponential scaling law. Numerical comparisons of this approach with the full toy Monte Carlo show that good agreement is achieved for sufficiently large number of observations.

In case more parameters are estimated from data, e.g.: when searching for a new resonance whose mass and width are both unknown, the look-elsewhere effect can be addressed with an extension of the approach described above, as detailed in Ref. [40].

## References

- [1] CMS collaboration, “Measurement of the top quark mass using proton-proton data at  $\sqrt{s} = 7$  and 8 TeV,” *Phys. Rev. D*, vol. 93, p. 072004, 2016.
- [2] ATLAS Collaboration, “Search for resonances decaying to photon pairs in 3.2 fb<sup>-1</sup> of  $pp$  collisions at  $\sqrt{s} = 13$  tev with the ATLAS detector,” 2015.
- [3] P. Laplace, *Essai philosophique sur les probabilités*. Paris: Courcier Imprimeur, 3<sup>rd</sup> ed., 1816.
- [4] A. Kolmogorov, *Foundations of the Theory of Probability*. New York: Chelsea Publishing Company, 1956.
- [5] H. Jeffreys, “An invariant form for the prior probability in estimation problems,” *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 186, pp. 453–461, 1946.
- [6] F. James and M. Roos, “MINUIT: Function minimization and error analysis.” Cern Computer Centre Program Library, Geneve Long Write-up No. D506, 1989.
- [7] L. Lista, “Statistical methods for data analysis in particle physics,” *Lect. Notes Phys.*, vol. 909, pp. 1–172, 2016.
- [8] H. Cramér, *Mathematical Methods of Statistics*. Princeton Univ. Press., 1946.
- [9] C. R. Rao, “Information and the accuracy attainable in the estimation of statistical parameters,” *Bulletin of the Calcutta Mathematical Society*, vol. 37, pp. 81–89, 1945.
- [10] J. Neyman, “Outline of a theory of statistical estimation based on the clasiscal theory of probability,” *Philosophical Transactions of the Royal Society of London, Series A, Mathematical and Physical Sciences*, vol. 236, pp. 333–380, August 1937.
- [11] K. A. Olive *et al.*, “The review of particle physics,” *Chin. Phys. C*, vol. 38, p. 010009, 2014 and 2015 update.
- [12] C. C.J. and E. Pearson, “The use of confidence or fiducial limits illustrated in the case of the binomial,” *Biometrika*, vol. 26, pp. 404–413, 1934.
- [13] S. Baker and R. Cousins, “Clarification of the use of chi-square and likelihood functions in fit to histograms,” *Nucl. Instr. Meth.*, vol. A221, pp. 437–442, 1984.

- [14] L. Lions, D. Gibaut, and P. Clifford, “How to combine correlated estimates of a single physical quantity,” *Nucl. Instr. Meth.*, vol. A270, pp. 110–117, 1988.
- [15] ATLAS, CMS, CDF and D0 collaborations, “First combination of Tevatron and LHC measurements of the top-quark mass,” 2014.
- [16] L. Lyons, A. J. Martin, and D. H. Saxon, “On the determination of the B lifetime by combining the results of different experiments,” *Phys. Rev.*, vol. D41, pp. 982–985, 1990.
- [17] L. Lista, “The bias of the unbiased estimator: a study of the iterative application of the BLUE method,” *Nucl. Instr. Meth.*, vol. A764, pp. 82–93, 2014. and corr. *ibid.* A773, pp. 87–96, 2015.
- [18] A. Valassi and R. Chierici, “Information and treatment of unknown correlations in the combination of measurements using the BLUE method,” *Eur. Phys. J.*, vol. C74, p. 2717, 2014.
- [19] The ALEPH, DELPHI, L3, OPAL Collaborations, the LEP Electroweak Working Group, “Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP,” *Phys. Rept.*, vol. 532, p. 119, 2013.
- [20] M. Baak, J. Cúth, J. Haller, A. Hoecker, R. Kogler, K. Mönig, M. Schott, and J. Stelzer, “The global electroweak fit at NNLO and prospects for the LHC and ILC,” *Eur. Phys. J.*, vol. C74, p. 3046, 2014.
- [21] J. Neyman and E. Pearson, “On the problem of the most efficient tests of statistical hypotheses,” *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 231, p. 289–337, 1933.
- [22] P. R. Byron, Y. Hai-Jun, Z. Ji, L. Yong, n. S. Io, and M. Gordon, “Boosted decision trees as an alternative to artificial neural networks for particle identification,” *Nucl. Instr. Meth.*, vol. A543, pp. 577–584, 2005.
- [23] R. A. Fisher, “The use of multiple measurements in taxonomic problems,” *Annals of Eugenics*, vol. 7, pp. 179–188, 1936.
- [24] H. N. Mhaskar, “Neural networks for optimal approximation of smooth and analytic functions,” *Neural Computation*, vol. 8, pp. 164–177, 1996.
- [25] P. Baldi, P. Sadowski, and D. Whiteson, “Searching for Exotic Particles in High-Energy Physics with Deep Learning,” *Nature Commun.*, vol. 5, p. 4308, 2014.
- [26] Y. Freund and R. Schapire, “A decision-theoretic generalization of on-line learning and an application to boosting,” 1994.
- [27] R. L. Wasserstein and N. A. Lazar, “The asa’s statement on p-values: context, process, and purpose,” *The American Statistician*, vol. 70, pp. 129–133, 2016.
- [28] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics,” *Eur. Phys. J.*, vol. C71, p. 1554, 2011.
- [29] G. Feldman and R. Cousins, “Unified approach to the classical statistical analysis of small signals,” *Phys. Rev.*, vol. D57, pp. 3873–3889, 1998.
- [30] O. Helene, “Upper limit of peak area,” *Nucl. Instr. and Meth.*, vol. A212, p. 319, 1983.
- [31] G. D.E. *et al.*, “The review of particle physics,” *The European Physical Journal C*, vol. 15, p. 1, 2000, 2001.
- [32] A. Read, “Modified frequentist analysis of search results (the  $CL_s$  method),” in *1st Workshop on Confidence Limits*, (CERN), 2000.
- [33] G. Abbiendi *et al.*, “Search for the standard model Higgs boson at LEP,” *Physics Letters*, vol. B565, pp. 61–75, 2003.
- [34] R. Cousins and V. Highland, “Incorporating systematic uncertainties into an upper limit,” *Nucl. Instr. Meth.*, vol. A320, pp. 331–335, 1992.
- [35] S. Wilks, “The large-sample distribution of the likelihood ratio for testing composite hypotheses,” *Ann. Math. Stat.*, vol. 9, pp. 60–62, 1938.
- [36] R. Brun and F. Rademakers, “ROOT - an object oriented data analysis framework,” *Proceedings*



*AIHENP'96 Workshop, Lausanne, Sep. 1996, Nucl. Inst. Meth.*, vol. A389, pp. 81–86, 1997. See also <http://root.cern.ch/>.

- [37] A. Wald, “Tests of statistical hypotheses concerning several parameters when the number of observations is large,” *Trans. Am. Math. Soc.*, vol. 54, pp. 426–482, November 1943.
- [38] I. Asimov, “Franchise”, in I. Asimov, “*The Complete Stories*”, vol. 1. Broadway Books, 1990.
- [39] E. Gross and O. Vitells, “Trial factors for the look elsewhere effect in high energy physics,” *Eur. Phys. J.*, vol. C70, p. 525, 2010.
- [40] O. Vitells and E. Gross, “Estimating the significance of a signal in a multi-dimensional search,” *Astropart. Phys.*, vol. 35, pp. 230–234, 2011.



# Cosmology and Dark Matter

*I. Tkachev*

Institute for Nuclear Research of the Russian Academy of Sciences, Moscow, Russia

## Abstract

This lecture course covers cosmology from the particle physicist perspective. Therefore, the emphasis will be on the evidence for the new physics in cosmological and astrophysical data together with minimal theoretical frameworks needed to understand and appreciate the evidence. I review the case for non-baryonic dark matter and describe popular models which incorporate it. In parallel, the story of dark energy will be developed, which includes accelerated expansion of the Universe today, the Universe origin in the Big Bang, and support for the Inflationary theory in CMBR data.

## Keywords

Cosmology; dark matter; lectures; dark energy; inflation.

## 1 Introduction

The deeper we dig into microphysics, the deeper connections between smallest and largest scales are revealed. The world which surrounds us on Earth in everyday life can be understood in frameworks of electrodynamics with atoms and molecules in hands. But we need nuclear physics to describe the Sun and other stars. Moving to even bigger scales, explaining galaxies and the Universe as a whole, we need good understanding of particle physics.

This relation works the other way around as well. In particular, cosmology tells us that the Standard Model of particle physics is incomplete. Namely, that the galaxies and galaxy clusters are made mainly of the Dark Matter, which overweights usual baryonic matter, and for which there is no room in the Standard Model of particle physics. It tells also that at the scale of the Universe the Dark Energy overrules, which can be simple or complicated substance, but which is not matter. We conclude that there ought to be a new physics and new particles outside of the Standard Model frameworks. At present we learn their properties from cosmology only.

The plan of this lectures is as follows. In Section 2, I review the basics of cosmology: Friedman equations, Hubble expansion and cosmography. Evidence for the dark energy is also presented in this section. In Section 3 the Hot Big Bang theory is outlined: relevant thermodynamal facts and relations are presented and Cosmic Microwave Background Radiation (CMBR) is introduced. In Section 4 the study of CMBR anisotropies is described as a tool of precision cosmology. In Section 5 basics of inflationary theory are given, while some important technical details of this theory are moved into Appendices. In Section 6 the evidence for dark matter is described together with dark matter models and dark matter searches results.

## 2 Basics of Cosmology

There are many excellent recent books and reviews on cosmology in the market, which readers may consult for missing details. I would especially recommend the balanced, contemporary and comprehensive book by Gorbunov and Rubakov [1].

### 2.1 Expansion of the Universe

Cosmological dynamics is provided by General Relativity - the Einstein field equations

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G T_{\mu\nu} , \quad (1)$$

where  $T_{\mu\nu}$  is a stress energy tensor describing the distribution of mass in space,  $G$  is Newton's gravitational constant, and the curvature  $R_{\mu\nu}$  is certain function of the metric  $g_{\mu\nu}$  and its first and second derivatives. Immediate consequence of Einstein equations is energy momentum conservation

$$T_{\mu}{}^{\nu}{}_{;\nu} = 0. \quad (2)$$

These equations take simple form in important physical situations with special symmetries. At large scales the Universe as a whole is homogeneous and isotropic and these symmetries form basis for the construction of cosmological models. The most general space-time metric describing such universe is the Robertson-Walker metric

$$ds^2 = dt^2 - a^2(t) d\mathbf{l}^2, \quad (3)$$

where  $a(t)$  is the dimensionless scale factor by which all distances vary as a function of cosmic time  $t$ . The scale factor contains all the dynamics of the Universe as a whole, while the vector product  $d\mathbf{l}^2$  describes the geometry of the space,

$$d\mathbf{l}^2 = \frac{dr^2}{1 - k r^2} + r^2 (d\theta^2 + \sin^2 \theta d\phi^2),$$

which can be either Euclidian, or positively or negatively curved. For the spatial 3-dimensional curvature we find, explicitly

$${}^{(3)}R = \frac{6k}{a^2(t)} \begin{cases} k = -1 & \text{Open} \\ k = 0 & \text{Flat} \\ k = +1 & \text{Closed} \end{cases} \quad (4)$$

E.g., the space with  $k = +1$  can be thought of as a 3-dimensional sphere with a curvature being inversely proportional to the square of its radius. In this Section we will model the matter content of the Universe as a perfect fluid with energy density  $\rho$  and pressure  $p$ , for which the stress-energy tensor in the rest frame of the fluid is

$$T_{\mu}{}^{\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & -p & 0 & 0 \\ 0 & 0 & -p & 0 \\ 0 & 0 & 0 & -p \end{pmatrix} \quad (5)$$

With these assumptions the Einstein equations simplify to the Friedmann equations, which form the dynamical basis of cosmology

$$\frac{\dot{a}^2}{a^2} = \frac{8\pi G}{3} \rho - \frac{k}{a^2}, \quad (6)$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3} (\rho + 3p). \quad (7)$$

As Alexander Friedmann have shown in 1922, a universe described by such equations cannot be static, it inevitably expands or collapses. Solution of Eqs. (6) and (7) can be found in the following way, which also highlights the physics behind these equations. Differentiating the first Friedmann equation and combining result with the second one we obtain

$$\frac{d\rho}{dt} + 3 \frac{\dot{a}}{a} (\rho + p) = 0, \quad (8)$$

which is nothing but the energy-momentum conservation  $T^{\mu\nu}{}_{;\nu} = 0$  written for the homogeneous isotropic medium. On the other hand the result Eq. (8) also corresponds to the First Law of thermodynamics

$$dE + p dV = T dS, \quad (9)$$

with  $dS = 0$ , where  $S$  is entropy,  $E = \rho V$ , and  $V \propto a^3$ . This is expected since in derivation of Friedmann equations the energy-momentum tensor of an ideal fluid was assumed. It turns out that this is valid approximation most of the time<sup>1</sup> and therefore the expansion of the Universe is adiabatic,  $S = \text{const}$ .

Let  $s$  be entropy density,  $S = sV$ . We know from thermodynamic (and I'll derive this in Section 3.2) that in thermal equilibrium the entropy density is given by

$$s = \frac{2\pi^2}{45} g_* T^3, \quad (10)$$

where  $g_*$  is the number of effective relativistic degrees of freedom

$$g_* = \sum_{i=\text{bosons}} g_i + \frac{7}{8} \sum_{j=\text{fermions}} g_j, \quad (11)$$

and  $g_i$  is the number of spin states of a given particle, e.g. for photons, electrons and positrons  $g_\gamma = g_{e^-} = g_{e^+} = 2$ . In this expression only particles with  $m \ll T$  are counted, i.e.  $g_*$  is a function of temperature. It is displayed in Fig. 2 by the solid line for the Standard Model and by the dashed line for the minimal supersymmetric extension of the SM. Therefore,  $S = \text{const}$  is equivalent to

$$g_*^{1/3} T a = \text{const}. \quad (12)$$

This is very useful relation. In particular, it gives  $T \propto a^{-1}$  (neglecting the change in the number of degrees of freedom), and we arrive to the concept of the Hot Universe right away. Indeed, currently the Universe is expanding, in the past it was smaller and therefore hotter. In the era of precision cosmology the change in  $g_*(T)$  should be counted too, if we go beyond simple estimates. We do not know all particle content of complete theory describing Nature, but LHC will fix the actual shape of  $g_*(T)$  in the region of highest  $T$  of Fig. 2 and, in fact, a sharp MSSM like rise in the number of degrees of freedom is ruled out already.

Only relativistic particles contribute to entropy, but everything existing contributes to the energy density,  $\rho$ . Even vacuum. The equation of state,  $w$ , of a substance contributing to  $\rho$  is defined as  $w \equiv p/\rho$ . If  $w = \text{const}$ , the energy-momentum conservation, Eq. (8), gives  $\rho = a^{-3(1+w)} \rho_0$ . Using this result, we find from the first Friedmann equation, Eq. (6), the scale factor as a function of time,  $a = (t/t_0)^{2/3(1+w)}$ .

During the first half of of the last century cosmologists were assuming that the Universe is filled with a “dust”,  $p = 0$ , while the “dust” particles were represented by galaxies made of usual matter. Nowadays we know that the Universe is multicomponent. Its energy density was dominated in turn by radiation ( $p = \rho/3$ ), by dark matter ( $p = 0$ , as for the “dust” of galaxies) and finally by dark energy (equation of state consistent with  $p = -\rho$ ). In the Table 1 we list: substances known to contribute into energy balance in the Universe, their defining equations of state, the corresponding scaling of energy density with expansion, and corresponding solution for  $a(t)$  if  $k = 0$ .

**Table 1:** Substances contributing into the energy balance in the Universe

| Substance | Equation of state | $\rho(a)$              | $a(t)$              |
|-----------|-------------------|------------------------|---------------------|
| Radiation | $w = 1/3$         | $\rho = a^{-4} \rho_0$ | $a = (t/t_0)^{1/2}$ |
| Matter    | $w = 0$           | $\rho = a^{-3} \rho_0$ | $a = (t/t_0)^{2/3}$ |
| Vacuum    | $w = -1$          | $\rho = \text{const}$  | $a = \exp(H_0 t)$   |

To parameterize the Fiedmann equations and their solution  $a(t)$ , cosmologists introduce *cosmological parameters*. One of such parameters we have already encountered, this is  $k$  entering Frideman

<sup>1</sup>It is violated though during special moments, in particular at the initial matter creation after inflation, see below, or at strongly first order phase transitions, if those existed in the Universe past.

equations explicitly. Despite parametrizing spatial geometry of the Universe it was used to predict the Universe future. Namely, it immediately follows from Eq. (6) that for  $k = -1$  or  $k = 0$  the Universe will expand forever. For  $k = +1$  the Universe should recollapse at the point when r.h.s. of Eq. (6) turns to zero. This should happen in radiation or matter dominated Universe, but never happens in the universe dominated by the dark energy with  $\rho_0 > 0$ . Since nowadays we know that our Universe is dominated by such dark energy, we already know that it will expand forever.

Inherently related to the parameter  $k$  is *critical density*. This is the density at which the Universe is spatially flat,  $k = 0$

$$\rho_c \equiv \frac{3}{8\pi G} \left( \frac{\dot{a}}{a} \right)^2. \quad (13)$$

Critical density can be expressed using a second, directly observable and very important cosmological parameter, the Hubble constant,  $H \equiv \dot{a}/a$ . To quantify relative contribution,  $\rho_i$ , of each of the components in the total energy budget of the Universe,  $\rho$ , the following notations are introduced,  $\Omega_i \equiv \rho_i/\rho_c$  and  $\sum_i \rho_i = \rho$ . The current knowledge of the numerical values of some of these parameters at  $t = t_0$  is summarized in the Table 2 below.

**Table 2:** Cosmological parameters

| Symbol & Definition                    | Description          | Present value, from Ref. [2]                                |
|--|----------------------|---|
| $t$                                    | Age of the Universe  | $t_0 = 13.81 \pm 0.03$ Gyr                                  |
| $H = \dot{a}/a$                        | Hubble parameter     | $H_0 = 67.27 \pm 0.66$ km s <sup>-1</sup> Mpc <sup>-1</sup> |
| $\Omega = \rho/\rho_c$                 | Spatial curvature    | $1 - \Omega = 0.000 \pm 0.005$                              |
| $\Omega_\gamma = \rho_\gamma/\rho_c$   | Fraction of photons  | $\Omega_\gamma = 2.48 \cdot 10^{-5} h^{-2}$                 |
| $\Omega_b = \rho_b/\rho_c$             | Baryonic fraction    | $\Omega_b h^2 = 0.02225 \pm 0.00016$                        |
| $\Omega_m = \rho_m/\rho_c$             | Matter fraction      | $\Omega_m = 0.316 \pm 0.009$                                |
| $\Omega_\Lambda = \rho_\Lambda/\rho_c$ | Dark Energy fraction | $\Omega_\Lambda = 0.684 \pm 0.009$                          |

Accuracy of numerical values presented in this table should not be over-appreciated since those were derived with some model assumptions, e.g. that the dark matter is cold and the equation of state for dark energy is  $w = -1$ . Relaxing such assumptions changes presented values somewhat.

## 2.2 The Hubble law

The velocity with which distance  $r = a(t)r_0$  between two arbitrary galaxies increases in expanding Universe follows trivially from the definition of the Hubble parameter,  $H \equiv \dot{a}/a$ , namely  $v = \dot{r} = H r$ . This relation, known as the Hubble law, makes the basis for direct observational determination of the Hubble parameter and was discovered by Edwin Hubble in 1929. It is also used to set units for measuring  $H$ . For convenience, in many cosmological relations, dimensionless ‘‘small  $h$ ’’ is introduced as  $H = 100 h$  km s<sup>-1</sup>Mpc<sup>-1</sup>, then  $h \sim O(1)$ . Latest value of the Hubble constant obtained from direct mapping of recession velocities versus distance corresponds to  $H_0 = (73.00 \pm 1.75)$  km s<sup>-1</sup> Mpc<sup>-1</sup> [3]. Note that this value is  $3.4\sigma$  higher than the value presented in Table 2 which was derived indirectly from other cosmological data. This may suggest some unaccounted systematic uncertainties or may indicate new physics [4].

Looking at cosmologically distant objects we see just the light they emit. One can wonder then, how observables are derived from this limited information? E.g. how distance and velocity can be measured separately to determine  $H$ ? The answer is simple. Velocity is measured by the frequency shift of known signal, similarly to what police is doing when checking for speeding cars using Doppler radars. Distance can be derived measuring dimming of objects with calibrated intrinsic luminosity: objects which are further away are less bright. Now we shall explain this in more details.

### 2.2.1 Redshift

Photon motion in any metric is described by basic equation,  $ds^2 = 0$ . In the Robertson-Walker metric this becomes  $ds^2 = dt^2 - a^2 d\chi^2 = 0$ , where  $\chi$  is *comoving* distance along particle trajectory. Define conformal time  $\eta$  as

$$d\eta = dt/a. \quad (14)$$

Then  $ds^2 = a^2(d\eta^2 - d\chi^2) = 0$ . Remarkably, solution for a photon world line in conformal coordinates is the same as in Minkowskian space-time  $\chi = \pm\eta + \text{const}$ . Therefore, the conformal time lapse between two events of light emission at one point will be the same as for their detection at another point, regardless of distance traveled,  $d\eta|_{\text{emission}} = d\eta|_{\text{detection}}$ . Therefore, the proper time lapses for emissions and detections are related as

$$\frac{dt}{a}|_{\text{detection}} = \frac{dt}{a}|_{\text{emission}}.$$

Let  $dt$  corresponds to the period of some monochromatic signal, then for its frequency we will have  $\omega a_d = \omega_0 a_e$ . As a result, we can say that the wavelength of a signal stretches together with the expansion of the Universe. An this interpretation is often used. However, it is incorrect. For example, the space does not stretches inside galaxies, but  $\omega a_d = \omega_0 a_e$  will be always true regardless of how many times the light signal passed through galaxies between detection and emission. The signal frequency changes between the point of emission,  $\omega_0$ , and in the point of detection,  $\omega$ , because clocks run differently in those points. In a similar way the gravitational redshift or blueshift can be also derived and interpreted. At largest cosmological scales we always have redshift since Universe is expanding.

In measurements, the redshift is quantified as  $z \equiv (\omega_0 - \omega)/\omega$  and measured as a frequency shift of emission or absorption lines of various chemical elements. With this definition we get

$$1 + z = \frac{a_d}{a_e}. \quad (15)$$

In other words, redshift can be used also to label cosmological epoch. For convenience, we can always normalise the scale factor today as  $a_d = 1$ . Recall relation  $a \propto T^{-1}$  (or, more precisely, Eq. 12). This gives temperature of the Universe as a function of redshift,  $T(z) = T_0(1 + z)$ . Also, using our previous result  $\rho = a^{-3(1+w)} \rho_0$  and Friedmann equation (6) we can rewrite Hubble parameter as a function of  $z$  and cosmological parameters  $H_0$  and  $\Omega_i$  in the following often used form

$$H^2(z) = H_0^2 \sum_i \Omega_i (1 + z)^{3(1+w_i)}. \quad (16)$$

Differentiating Eq. (15) we find  $dz = -Hd\chi$  which is a local form of the Hubble law, partially expressed through observables already. Instead of  $v$  we have directly measurable redshift  $z$ . This is understandable since in the non-relativistic limit cosmological redshift reduces to the Doppler effect - at the end, galaxies are receding from us with the Universe expansion. Finally, integral of the local form of the Hubble law,  $dz = -Hd\chi$ , gives

$$\chi(z) = \int_0^z \frac{dz'}{H(z')}. \quad (17)$$

Now we have to find the way to measure distances to remote objects to complete construction of the Hubble law generalisation, which would be expressed through observables only and which would be valid to any redshift.

### 2.2.2 Luminosity distance

Consider two objects with identical luminosities (Standard Candles) placed at different distances from us. The radiation flux,  $F$ , scales with distance as  $F^{-1/2}$ . Therefore, measuring fluxes from a standard

candles, we can determine the ratio of distances to them. Moreover, the distance to an object with known luminosity can be defined as Luminosity Distance,  $D_L$ , and measured via measuring flux from it

$$D_L^2 \equiv \frac{L}{4\pi F}.$$

Consider now this idea in cosmology. Again, let us write metric in conformal time, but now we should keep track of changing area with distance at fixed solid angle,  $ds^2 = a^2 (d\eta^2 - d\chi^2 - \chi^2 d\Omega)$ . Surface area at the point of detection is  $4\pi \chi^2$  (we can always normalise scale factor today as  $a = 1$ ). Further, energy and arrival rates of registered photons are redshifted. This reduces the flux by  $(1+z)^2$ , where  $z$  is redshift of emitter. We get for observed bolometric flux

$$F = \frac{L}{4\pi \chi^2 (1+z)^2},$$

and luminosity distance is

$$D_L \equiv \sqrt{\frac{L}{4\pi F}} = (1+z) \chi.$$

Therefore, measuring flux, we can determine comoving distance to a standard candle. We may not know the value of intrinsic luminosity  $L$ , but this is not necessary. It is important only that  $L$  should not vary from an object to object. Then we can compare ratio of fluxes at different redshifts and from there derive cosmology. For historical reasons astronomers measure flux in magnitudes, which are defined as  $\mu \propto 5 \log_{10} F$ . Ratio of fluxes will be difference in magnitudes. Now, we want to see how different are magnitudes of a standard candles at a given redshift in different cosmologies, say in cosmology which predicts  $\chi(z)$  and in a ‘‘base’’ cosmology which predicts  $\chi(z)_{\text{base}}$ . We find

$$\Delta\mu = \mu - \mu_{\text{base}} = 5 \log_{10} \left[ \frac{\chi(z)}{\chi(z)_{\text{base}}} \right] \quad (18)$$

Here  $\chi(z)$  and  $\chi(z)_{\text{base}}$  are given by Eq. (17) with its own sets of cosmological parameters each.

### 2.2.3 Dark Energy

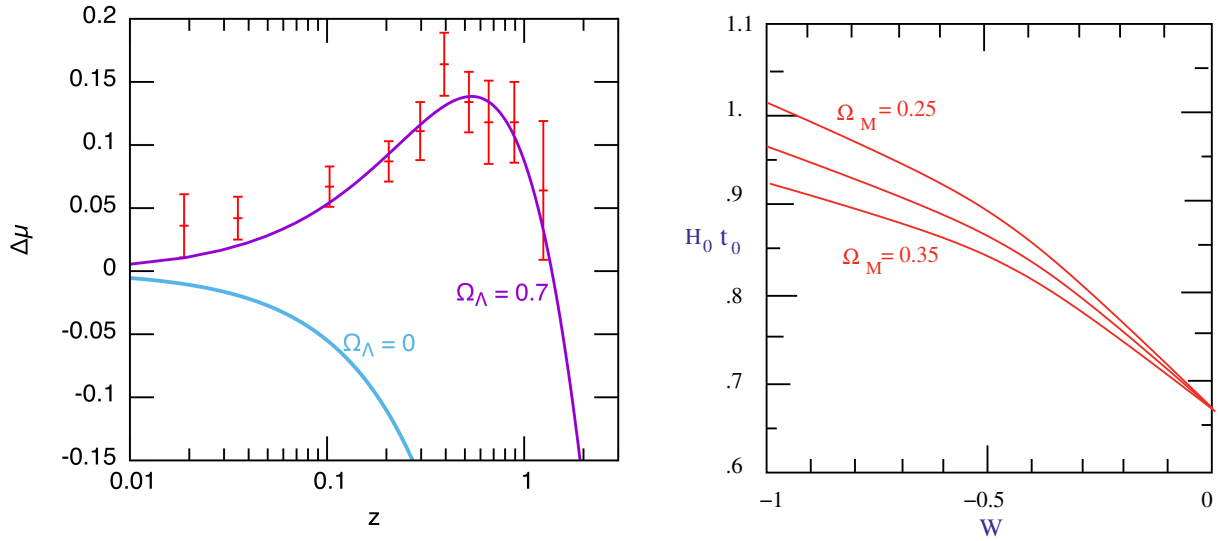
Supernova of type Ia have been shown to be good standard candles. In measurements of luminosity distance to them the dark energy has been discovered. Below I illustrate this result in Fig. 1, left panel, using modern data and relation Eq. (18). Blue curve corresponds to the Universe without dark energy,  $\Omega_\Lambda = 0$ , while violet curve corresponds to the best fit over dark energy which gives  $\Omega_\Lambda = 0.7$ . As a base cosmology which was subtracted I used the Universe with  $\ddot{a} = 0$ , but the subtraction is not important here and is needed for the visualisation purposes only, to enhance separation of curves on a graph.

In fact, prior to the dark energy discovery in Supernova data, scientists already suspected for a while that it exists. Several hints existed, I illustrate the one derived from attempts to determine the Universe age. During matter dominated expansion  $a \propto t^{2/3}$ . Therefore, for a matter dominated Universe without dark energy we would have  $H_0 t_0 = 2/3$ . However direct measurements of the Hubble constant gave at the time  $H_0 = 70 \pm 7 \text{ km sec}^{-1} \text{ Mpc}^{-1}$ , while the age was estimated (using ages of the oldest stars) as  $t_0 = 13 \pm 1.5 \text{ Gyr}$ . Therefore, measurements producing  $H_0 t_0 = 0.93 \pm 0.15$  were in contradiction with prediction for the matter dominated Universe.

Let us see what happens if we add dark energy to a matter. The age of the multicomponent universe can be found integrating  $dt = ad\chi$  and using  $dz = -Hd\chi$ , which for the universe age at redshift  $z$  gives

$$t(z) = \int_z^\infty \frac{dz'}{(1+z')H(z')}.$$





**Fig. 1:** Evidence for the dark energy from Supernova data, left panel, and from the Universe age, right panel.

In particular, if equation of state of dark energy corresponds to a vacuum,  $w = -1$ , and universe is spatially flat,  $\Omega_\Lambda + \Omega_M = 1$ , this gives

$$H_0 t_0 = \frac{2}{3\sqrt{\Omega_\Lambda}} \ln \left( \frac{1 + \sqrt{\Omega_\Lambda}}{\sqrt{\Omega_M}} \right).$$

Such a universe with  $\Omega_\Lambda \simeq 0.7$  is a good fit to observations as opposed to a matter dominated universe, see Fig. 1, right panel and compare to a modern data in Table 2 which give  $H_0 t_0 \simeq 0.95$ .

### 3 Hot Big Bang

So far we have considered cosmography of the late Universe and found that the Universe should be filled with matter and dark energy. However, the Universe should contain radiation also. Today its contribution is negligible, but in the early Universe it was dominant fraction. Indeed, energy density of radiation is fastest growing fraction when we look back in time,  $a \rightarrow 0$ , see Table 1. And the Universe was hotter as well in this limit, see Eq. (12). To reach such conclusions we have to assume also that the Universe was in thermal equilibrium in the past. But this is inevitable too since in a denser medium relaxation time is shorter. Universe was indeed in thermal equilibrium in the past, as we will shortly see. The concept of the Hot Universe is so natural and so inevitable, that it is hard to imagine nowadays that it was not widely accepted until relict radiation has been observed.

#### 3.1 Cosmic Microwave Background Radiation

The Universe is filled with radiation which is left-over from the Big Bang. The name for this first light is Cosmic Microwave Background Radiation (CMBR). Measurements of tiny fluctuations (anisotropy) in CMBR temperature give a wealth of cosmological information and became a most powerful probe of cosmology.

This radiation was predicted by Georgi Gamov in 1946, who estimated its temperature to be  $\sim 5 K^\circ$ . Gamov was trying to understand the origin of chemical elements and their abundances. Most abundant, after hydrogen, is helium, with its shear being  $\sim 25\%$ . One possibility which Gamov considered was nucleo-synthesis of He out of H in stars. Dividing the total integrated luminosity of the stars by the energy released in one reaction, he estimated the number of produced He nuclei. This number was

too small in comparison with observations. Gamov assumed then that the oven where the light elements were cooked-up was the hot early Universe itself. He calculated abundances of elements successfully and found that the redshifted relic of thermal radiation left over from this hot early epoch should correspond to  $\sim 5 K^\circ$  at present. In one stroke G. Gamov founded two pillars (out of four) on which modern cosmology rests: CMBR and Big Bang Nucleosynthesis (BBN). Hot Big Bang theory was born.

Cosmic microwave background has been accidentally discovered by Penzias and Wilson [5] at Bell Labs in 1965 as the excess antenna temperature which, within the limits of their observations, was isotropic, unpolarized, and free from seasonal variations. A possible explanation for the observed excess noise temperature was immediately given by Dicke, Peebles, Roll, and Wilkinson and was published in a companion letter in the same issue [6]. Actually, they were preparing dedicated CMBR search experiment, but were one month late. Penzias and Wilson measured the excess temperature as  $\sim 3.5 \pm 1 K^\circ$ . It is interesting to note that the first (unrecognized) direct measurements of the CMB radiation was done by T. Shmaonov at Pulkovo in 1955, also as an excess noise while calibrating the RATAN antenna [7]. He published the temperature as  $(3.7 \pm 3.7) K^\circ$ . And even prior to this, in 1940, Andrew McKellar [8] had observed the population of excited rotational states of CN molecules in interstellar absorption lines, concluding that it was consistent with being in thermal equilibrium with a temperature of  $\approx 2.7 K^\circ$ . Its significance was unappreciated and the result essentially forgotten. Finally, before the discovery, in 1964 Doroshkevich and Novikov in an unnoticed paper emphasized [9] the detectability of a microwave blackbody as a basic test of Gamov's Hot Big Bang model.

The spectrum of CMBR is a perfect blackbody, with a temperature [10]

$$T_0 = 2.7255 \pm 0.0006 K^\circ, \quad (19)$$

as measured by modern instruments. This corresponds to 410 photons per cubic centimeter or to the flux of 10 trillion photons per second per squared centimeter.

### 3.2 Thermodynamics of the Universe

There is no explanation to CMBR but the hot Big Bang. And since CMB is the radiation with black body spectrum, we know that the Universe was once in the thermal equilibrium. It immediately follows from Eq. (12) that in the past the Universe was hotter since it was smaller. We can and should use thermodynamics describing the early Universe.

For particles in thermal equilibrium the phase-space distribution functions are:

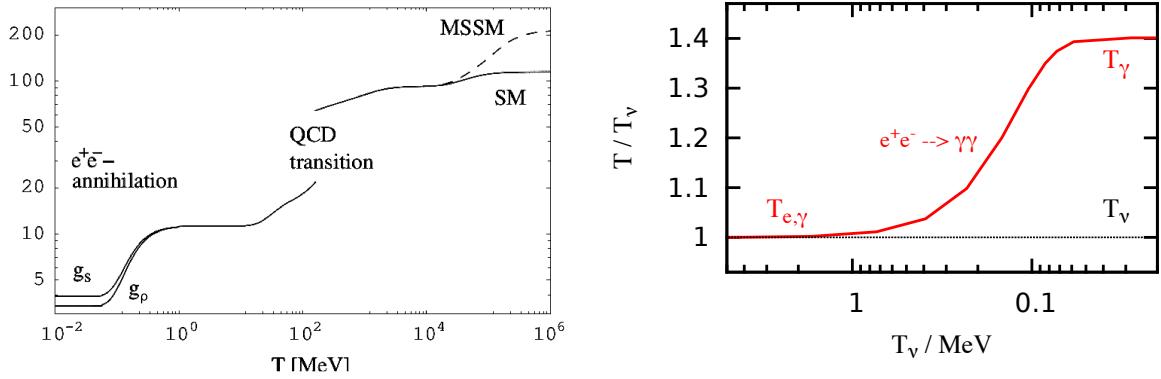
$$f_i(k) = \frac{g_i}{(2\pi)^3} \frac{1}{e^{(k_0 - \mu_i)/T} \pm 1}, \quad (20)$$

where  $k_0$  is particle energy,  $k_0 = \sqrt{\vec{k}^2 + m_i^2}$ ,  $\mu$  is chemical potential and the plus (minus) sign corresponds to fermions (bosons). Index  $i$  refers to different particles species and  $g_i$  is the number of their spin states, e.g. for photons, electrons and positrons  $g_\gamma = g_{e^-} = g_{e^+} = 2$  correspondingly, for neutrino and antineutrino  $g_\nu = g_{\bar{\nu}} = 1$ . All thermodynamical relations which we will need are derived using this function. In particular, number density of  $i$ -th particle species and their contribution into energy-momentum tensor are, correspondingly

$$n_i = \int d^3k f_i(k), \quad (21)$$

$$T_{\mu\nu}(i) = \int d^3k \frac{k_\mu k_\nu}{k_0} f_i(k). \quad (22)$$

Equation (22) gives energy density as  $\rho = T_{00}$ , while pressure is expressed through the trace over spatial part of the energy-momentum tensor,  $p = -T_j^j/3$ . To find overall energy density and pressure



**Fig. 2:** Left panel. Number of relativistic degrees of freedom,  $g_*$ , as a function of temperature. Solid line - the Standard Model case, dashed line - hypothetical behaviour in the Minimal Supersymmetric Standard Model (MSSM). Important events in the Universe evolutions are also indicated. Right panel. Increase of photon temperature over neutrino temperature during  $e^+e^-$  annihilation.

entering Friedmann equations one has to sum over all particle species, i. Entropy density is calculated as  $s = (\rho + p - \mu n)/T$ . Let us consider now important limits of these expressions.

1. *Relativistic particles.* First of all, for relativistic particles, regardless of particular form of  $f(k)$ , we have  $p = \rho/3$ . In other words, this relation is valid even out of thermal equilibrium and simply follows from definitions since  $k_0 = |\vec{k}|$  for  $m = 0$ . Further, for relativistic plasma without chemical potentials, integrals in Eqs. (21) and (22) can be calculated analytically and are slightly different for bosons and fermions. Summing out over all particles we find

$$n = g'_* \frac{\zeta(3)}{\pi^2} T^3 \quad (23)$$

$$\rho = g_* \frac{\pi^2}{30} T^4 \quad (24)$$

$$s = g_* \frac{2\pi^2}{45} T^3 \quad (25)$$

where  $\zeta(3) \approx 1.2$  is Rieman zeta function and

$$g'_* = \sum_{\text{bosons}} g_i + \frac{3}{4} \sum_{\text{fermions}} g_i \cdot$$

$$g_* = \sum_{\text{bosons}} g_i + \frac{7}{8} \sum_{\text{fermions}} g_i \cdot$$

In these expressions particles with  $m \ll T$  should be counted only, i. e.  $g_*$  and  $g'_*$  are functions of the temperature. Temperature dependence of  $g_*$  is shown in Fig. 2. Why it splits on  $g_s$  and  $g_\rho$  at  $T \lesssim \text{MeV}$  will be explained later on.

2. *Non-relativistic particles.* For non-relativistic particles all densities are exponentially suppressed in thermal equilibrium and, again in the case without chemical potentials, we find

$$n_i = g_i \left( \frac{m_i T}{2\pi} \right)^{3/2} e^{-m_i/T} \quad (26)$$

$$\rho_i = m_i n_i + \frac{3}{2} T n_i$$

$$p_i = T n_i$$

Here expression (26) for  $n_i$  is most important. In particular, it makes the basis for *Saha equation*, which will be used repeatedly throughout the lectures. This equation gives surviving amount of particles when they go out of equilibrium and will be used to discuss Big Bang nucleosynthesis, hydrogen recombination and abundance of thermally produced dark matter.

### 3.2.1 Cosmological density of neutrino

In the expanding Universe particle concentrations,  $n$ , are in equilibrium as long as reaction rate is sufficiently high  $\sigma n v > H$ , where  $\sigma$  is corresponding crosssection. After that distributions do not change in a comoving volume, i.e. "freeze-out". Weak interaction rate for neutrino  $\sigma_W n \sim G_F^2 T^2 \cdot T^3$  matches expansion rate,  $H \sim T^2/M_{\text{Pl}}$ , when  $G_F^2 M_{\text{Pl}} T^3 \approx 1$ . We conclude that neutrino are in thermal equilibrium at temperatures  $T \gg 1$  MeV and decouple from the rest of plasma at lower temperatures. Therefore, Standard Model neutrinos, which have small masses, decouple when they are still relativistic. The number density of neutrino at this time is given by Eq. (23) with  $g'_* = 2$ . Below this temperature, neutrinos are no longer in thermal equilibrium with the rest of the plasma, and their temperature simply decreases as  $T \propto 1/a$ . However, the cosmological background of photons is heated up by the  $e^+e^-$  annihilations shortly after neutrino "freeze-out". Let us find a relation between  $T_\nu$  and  $T_\gamma$ , which will also give the relation between  $n_\nu$  and  $n_\gamma$ .

Recall that entropy in the comoving volume conserves,  $g_* T^3 = \text{const}$ . Before annihilation  $g_* = g_\gamma + g_e \cdot (7/8) = 2 + 4 \cdot (7/8) = 11/2$ . After annihilation  $g_* = g_\gamma = 2$ . (Neutrinos are decoupled already and do not participate in these relations.) Since before annihilation  $T_\nu = T_\gamma$  the condition  $g_* T^3 = \text{const}$  gives for the neutrino temperature after positron annihilation

$$T_\nu = (4/11)^{1/3} T_\gamma. \quad (27)$$

This can be compared to the result of numerical integration of corresponding Boltzmann equations which is shown in Fig. 2, right panel. Present day photon temperature  $T_\gamma = 2.7255$  K, therefore, present day neutrino temperature is  $T_\nu = 1.9454$  K. For the number density of one flavour of left-handed neutrino and antineutrino we find  $n_\nu = 3n_\gamma/11 = 115 \text{ cm}^{-3}$ . Here we have used Eq. (23) and  $g_\gamma = g_\nu = 2$ . Right-handed neutrino, even if exist and light, are not in thermal equilibrium and are excited in small amounts, see Ref. [11] and Section 6.3.4 about "sterile" neutrino as dark matter.

Assume that by now neutrino became non-relativistic, i.e. their masses are larger than the present day temperature. In this case, neutrino energy density is given by  $\rho_\nu = \sum_i m_{\nu i} n_{\nu i}$ . Since it has to be smaller than  $\Omega_m \rho_c$ , we already have the constraint  $\sum_i m_{\nu i} < 93 \Omega_m h^2 \text{ eV} \approx 10 \text{ eV}$ . Modern cosmological constraints on neutrino masses are almost two orders of magnitude stronger actually (see later in the lectures).

Now, using Eqs. (24) and (27) we find that after "freeze-out" and  $e^+e^-$  annihilation (but at  $T > m_\nu$ ) the cosmological radiation background is parametrized as

$$\rho_r = \rho_\gamma + \rho_\nu = \frac{\pi^2}{15} T_\gamma^4 \left[ 1 + \frac{7}{8} \left( \frac{4}{11} \right)^{4/3} N_{eff} \right]. \quad (28)$$

At face value  $N_{eff}$  stands here for the number of active neutrino flavours and should be equal to three. But actually according to conventions used by cosmologists,  $N_{eff} \neq 3$ , neither it is integer. The reasons are as follows:

- When  $e^+e^-$  annihilate, neutrino are not decoupled completely yet since neutrino "freeze-out" temperature  $\approx 1$  MeV is too close to  $e^+e^-$  annihilation temperature. This leads to slight neutrino heating with distorted distribution (20) and to somewhat larger neutrino energy density  $\rho_\nu$  [12], which in applications is parametrised simply as larger  $N$ , to account only for the increase of  $\rho_\nu$ ,

$$N_{eff} = 3.046.$$

- There can be other contributions into radiation, e.g. light sterile neutrinos, Goldstones, or some other very light particles. These contributions are called "dark radiation". They are also included into  $N_{eff}$  and "dark radiation" is searched for in modern data as a signal that  $N_{eff} > 3.046$ .

Therefore,  $N_{eff}$  is another important cosmological parameter, potentially signalling new physics.

Now we can also understand why  $g_*$  splits on  $g_s$  and  $g_\rho$  at  $T \lesssim \text{MeV}$ . At these temperatures, radiation consists of two fractions with different temperatures each, gas of photons and gas of neutrino. Therefore, when

writing Eqs. (24) and (25) we have two options. We could have two terms in each of these equations, one for photons, another for neutrino, each term would have different temperatures. Or we can do the same way as in Eq. (28), including ratio of temperatures into  $g_*$  instead. And this latter approach has been decided to be more convenient by cosmologists. Since temperature enters in different powers to energy and entropy densities, we have splitting of  $g_*$  on  $g_s$  and  $g_\rho$ . Asymptotic values of these functions, which can be used at  $T < 100$  keV, are shown below, assuming there is no dark radiation

$$g_\rho(0) = 2 + 3.046 \left(\frac{7}{4}\right) \left(\frac{4}{11}\right)^{4/3} = 3.38, \quad (29)$$

$$g_s(0) = 2 + 3.046 \left(\frac{7}{4}\right) \left(\frac{4}{11}\right) = 3.94. \quad (30)$$

### 3.2.1.1 Matter-radiation equality

Radiation energy density scales with expansion as  $\rho \propto a^{-4}$ , while matter energy density scales as  $\rho \propto a^{-3}$ . It follows that the Universe was radiation dominated at the early stages of the evolution. Let us find now at which cosmological redshift and temperature the very important event happens: namely, when the energy density of radiation becomes equal to the energy density of matter. Using Eq. (28) with  $N_{eff} = 3.046$  and present day photon temperature, Eq. (19), we obtain  $\rho_r = 4.41 \times 10^{-10}$  GeV cm<sup>-3</sup> for the current radiation energy density. Recall now the value of critical density,  $\rho_c = 1.05 \times 10^{-5} h^2$  GeV cm<sup>-3</sup>, Eq. [13], to get  $\Omega_r = 4.2 \times 10^{-5} h^{-2}$ . Since radiation scales as  $\rho_r = \rho_c \Omega_r (1+z)^4$  while matter as  $\rho_m = \rho_c \Omega_m (1+z)^3$ , we find for the redshift of matter-radiation equality

$$1 + z_{eq} = \frac{\Omega_m}{\Omega_r} = 3400,$$

and for the corresponding temperature  $T_{eq} = 0.8$  eV. Deriving this I used values for  $\Omega_m$  and  $h$  from Table 2. Keeping those as free parameters we have  $T_{eq} = 5.6 \Omega_m h^2$  eV.

At higher temperatures the Universe was radiation dominated and its expansion was governed by the following Hubble parameter

$$H = \sqrt{\frac{8\pi G \rho_r}{3}} = \sqrt{\frac{8\pi^3 g_* T^4}{90 M_{Pl}^2}} \simeq 1.66 \sqrt{g_*} \frac{T^2}{M_{Pl}}. \quad (31)$$

Since during radiation dominated stage  $H = 1/2t$ , we obtain the Universe age (in seconds) as a function of temperature

$$t(\text{s}) = \frac{2.42}{\sqrt{g_*}} \left(\frac{\text{MeV}}{T}\right)^2.$$

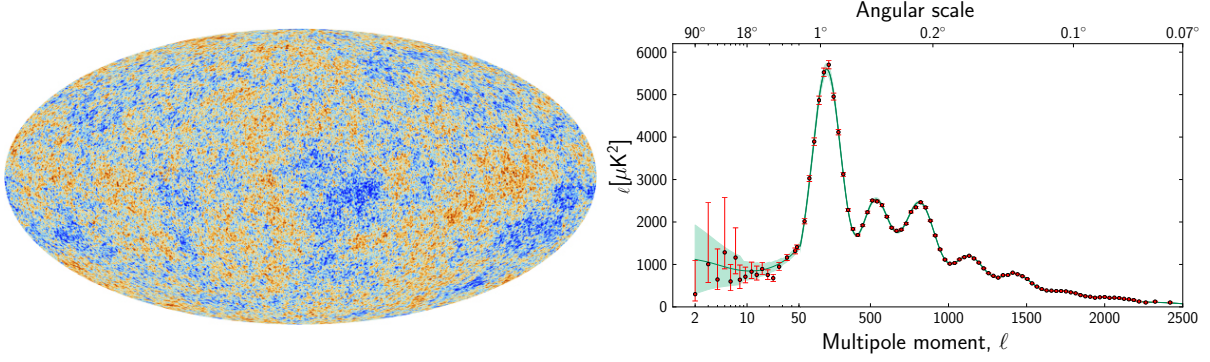
Stretching this time-temperature relation to equality temperature, and using expression (29) for  $g_*$ , we find that at equality the Universe was 65 thousand years old.

## 3.3 Last scattering of light

Baryonic matter is ionized at temperatures higher than the hydrogen ionization energy  $E_{ion} = 13.6$  eV and photons are in thermal equilibrium with primordial plasma. They cannot propagate large distances and the plasma is not transparent. With expansion the Universe cools down. At some point protons and electrons of primordial plasma recombined into neutral hydrogen and the Universe became transparent for radiation. This happens when the mean free path of photons becomes comparable to the size of the Universe at that time. Corresponding temperature is called "last scattering". After that photons are travelling without being affected by scattering. We see this light as Cosmic Microwave Background Radiation (CMBR). More precisely, the CMBR comes from the surface of the last scattering. We cannot see past this surface. Let us determine here when the last scattering had occurred in the early Universe.

Fraction of ionized hydrogen as function of temperature can be described by the Saha equation. It is derived by simply making ratios of number densities, Eq. (26), of different fractions in question in thermal equilibrium. For the case of hydrogen recombination

$$\frac{n_e n_p}{n_H} = \left(\frac{m_e T}{2\pi}\right)^{3/2} e^{-E_{ion}/T}. \quad (32)$$



**Fig. 3:** Sky map of primordial temperature fluctuations in Galactic coordinates (left panel) and temperature power spectrum (right panel) as measured by Planck space observatory [2].

Here  $n_e$ ,  $n_p$  and  $n_H$  are the number densities of electrons, protons, and neutral hydrogen respectively. Plasma is electrically neutral, i.e.  $n_e = n_p$ . To find closed relation for the fraction of ionized atoms,  $X \equiv n_p/(n_p + n_H) = n_p/n_B$ , we need the relation between the baryon number density,  $n_B$ , and temperature. This relation can be parameterized with the help of an important cosmological parameter called **baryon asymmetry**

$$\eta = \frac{n_B}{n_\gamma} = \frac{n_p + n_H}{n_\gamma} = (6.1 \pm 0.05) \times 10^{-10}, \quad (33)$$

where  $n_\gamma$  is the number density of photons, Eq. (23). Baryon asymmetry can be estimated by an order of magnitude by simply counting the number of baryons, or comparing element abundances predicted by the theory of Big Bang Nucleosynthesis to observations. Those are not most precise methods, though; the value presented in Eq. (33) was obtained from fitting the spectrum of CMBR fluctuations, see below. Nowadays, this is the most precise baryometer.

Defining recombination as the temperature when  $X = 0.1$ , we find  $T_{\text{rec}} \approx 0.3$  eV. The Universe became transparent for radiation when the mean free path of photons became comparable to the size of the Universe at that time. Photons scatter mainly on electrons and we find that the Universe became transparent when

$$(\sigma_{\gamma e} n_e)^{-1} \sim t. \quad (34)$$

Here,  $\sigma_{\gamma e} = 8\pi\alpha^2/3m_e^2$  is the Compton cross-section. For the temperature of last scattering we find  $T_{\text{ls}} \approx 0.26$  eV. Taking the ratio to the current CMBR temperature we find  $z_{\text{ls}} \approx 1000$ .

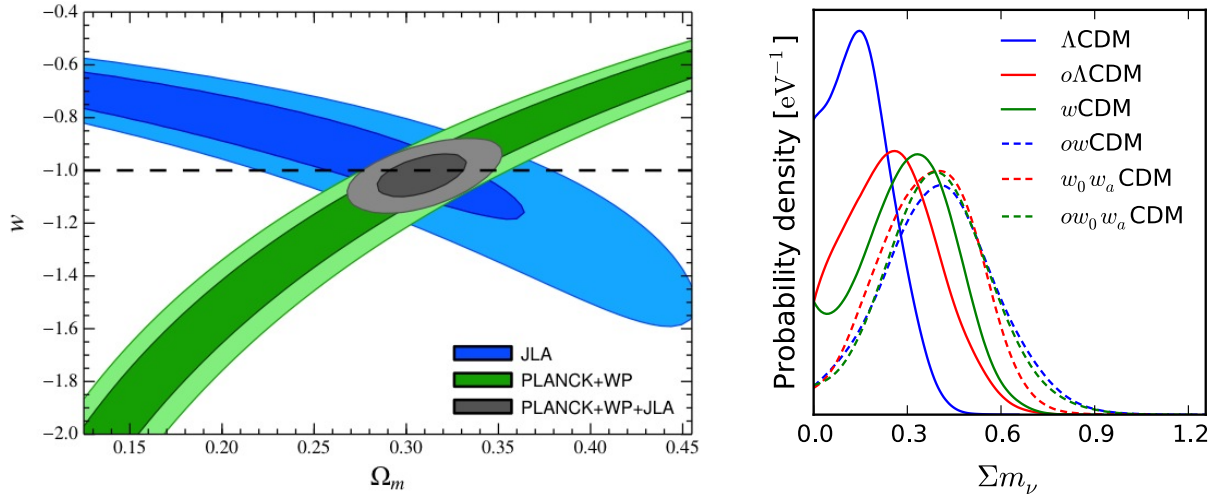
CMBR is the oldest light in the Universe. When registering it, we are looking directly at the deepest past we can, using photons. These photons had traveled the longest distances without being affected by scattering, and geometrically came out almost from the universe Horizon. Therefore the CMBR gives us a snapshot of the baby Universe at the time of last scattering.

#### 4 CMB power spectrum: tool of Precision Cosmology

The temperature of CMBR is slightly different in different patches of the sky - to 1 part in 100,000. These temperature deviations are shown in the sky map Fig. 3, left panel. Measurements of these tiny fluctuations (anisotropy) in CMBR temperature give us a wealth of cosmological information at an unprecedented level of precision and became a most powerful probe of cosmology. The functional form of the CMBR power spectrum is very sensitive to both the various cosmological parameters and to the shape, strength and nature of primordial fluctuations. This spectrum is shown in Fig. 3, right panel. In fact, the values of cosmological parameters listed in Table I largely came out from fitting model predictions to data as in this figure.

The temperature anisotropy,  $T(\mathbf{n})$ , as a function of viewing direction vector  $\mathbf{n}$ , as shown in Fig. 3, left panel, is naturally expanded in a basis of spherical harmonic,  $Y_{lm}$

$$T(\mathbf{n}) = \sum_{l,m} a_{lm} Y_{lm}(\mathbf{n}). \quad (35)$$



**Fig. 4:** Left panel: Combination of SN Ia (blue contours) and Planck data (green contours) tell us that the equation of state of the dark energy is consistent with that of a vacuum,  $w = -1$ , from Ref [19]. Right panel: Constraints on neutrino mass from combined Planck and BAO data, from Ref [20].

Coefficients  $a_{lm}$  in this decomposition define the angular power spectrum,  $C_l$

$$C_l = \frac{1}{2l+1} \sum_m |a_{lm}|^2. \quad (36)$$

Assuming random phases, the r.m.s. temperature fluctuation associated with the angular scale  $l$  can be found as

$$\Delta T_l = \sqrt{C_l l(l+1)/2\pi} \equiv \sqrt{D_l}. \quad (37)$$

Spectrum,  $D_l$ , as measured by Planck collaboration, is shown in Fig. 3, right panel. In fact, it was realized already right after the discovery of CMBR, that fluctuations in its temperature should have fundamental significance as a reflection of the seed perturbations which grew into galaxies and clusters. In a pure baryonic Universe it was expected that the level of fluctuations should be of the order  $\delta T/T \sim 10^{-2} - 10^{-3}$ . Measurements of the CMBR anisotropy with ever-increasing accuracy have begun. Once the temperature fluctuations were shown to be less than one part in a thousand, it became clear that baryonic density fluctuations did not have time to evolve into the nonlinear structures visible today. A gravitationally dominant dark matter component was invoked. For explanations why it is necessary, see Section 6. Eventually, fluctuations were detected [13] at the level of  $\delta T/T \sim 10^{-5}$ , consistent with the structure formation in Cold Dark Matter models with the Harrison-Zel'dovich spectrum of primordial perturbations motivated by cosmological Inflation, see Section 5 and Appendices.

The foundations of the theory of CMBR anisotropy were set out by Sachs & Wolfe [14], Silk [15], Peebles & Yu [16], Syunyaev & Zel'Dovich [17]. The measured spectrum of CMBR power has a characteristic shape of multiple peaks. Positions of these peaks and their relative amplitudes are sensitive to many cosmological parameters in a non-trivial way. Fitting the data to model predictions gives very accurate values for many of these parameters (though there are some degeneracies between different sets). Numerical calculations for different models were done already in Ref. [18], and power spectra exhibiting acoustic peaks (similar to those in Fig. 3, right panel) were presented. It was realized, in particular, that positions of the peaks are shifted with respect to each other for adiabatic and isentropic primordial fluctuations.

To improve significantly constraints on some cosmological parameters a combination of CMBR with other data is needed. For example, combining Planck data with Supernova data we find that the Dark energy equation of state is close to a vacuum,  $w = -1.02 \pm 0.06$  [19], while each of these sets alone would give weak constraints, see Fig. 4, left panel. Combination of Planck data with data on correlation properties of galaxy clustering, namely data on Baryon Acoustic Oscillations (BAO) tells us that the Universe is spatially flat,  $1 - \sum \Omega_i = 0.000 \pm 0.005$ , see Ref. [2]. That same data set improves many other constraints. An example of constraints on the sum of neutrino masses in this joint data set is shown in Fig. 4, right panel [20]. Solid blue line corresponds to the case of

$\Lambda$ CDM model, which means zero spatial curvature and  $w = -1$ . Other curves on this figure correspond to a Dark energy models with equation of state different from that of a vacuum. For the  $\Lambda$ CDM the constraint on the neutrino masses is  $\sum m_\nu < 0.22$  eV with positive  $2\sigma$  detection if Dark energy is more complicated substance than vacuum. However,  $\Lambda$ CDM is a good model and is consistent with all cosmological data to date.

Therefore  $\Lambda$ CDM can be safely assumed, and then other cosmological parameters can be determined quite well from the CMBR data alone. For example, parameters presented in Table 2 (except spatial curvature) were obtained from the CMBR data alone assuming  $\Lambda$ CDM model. Note that the constraint on  $\Omega_\Lambda$  from the supernova luminosity distance relations, Section 2.2.3, I also obtained assuming the  $\Lambda$ CDM model. Otherwise from the SN data alone we would only know for sure that the Universe expansion is accelerating, but the fraction of dark energy,  $\Omega_{DE}$ , would depend upon assumed equation of state  $w$ , as it is shown by blue contours in Fig. 4, left panel.

#### 4.1 Acoustic oscillations

As we could see already, large amount of cosmological information is encoded in the functional form of the CMBR power spectrum. To get feeling of physics which is behind, let us give a qualitative picture of why the CMBR power spectrum has a specific shape of a sequence of peaks, and explain how it depends on the values of particular cosmological parameters. Insight, sufficient for the purposes of these lectures, can be gained with the idealization of a perfect radiation fluid. In complete treatment, one has to follow the evolution of coupled radiation and metric fluctuations, i.e. to solve the linearized Einstein equations. However, essential physics of radiation (or matter) fluctuations can be extracted without going into the tedious algebra of General Relativity. It is sufficient to consider the energy-momentum conservation, Eq. (2). To solve for metric perturbations, full treatment based on Einstein equations, Eq. (1), is needed of course. We will not do that here, but simply quote results for the evolution of the gravitational potentials (coincident in some important cases with the solutions for the Newtonian potentials).

Perturbations of the ideal radiation fluid,  $p = \rho/3$ , can be separated into perturbations of its temperature, velocity and gravitational potential. In the general-relativistic treatment gravitational potential appears as a fractional perturbation of the scale factor in the perturbed metric

$$ds^2 = a^2(\eta) [(1 + 2\Psi)d\eta^2 - (1 - 2\Phi)dx^i dx_j] . \quad (38)$$

Two equations contained in the energy-momentum conservation,  $T^{\mu\nu}{}_{;\nu} = 0$  (i.e. temporal  $\mu = 0$  and spatial  $\mu = i$  parts of this equation), written in metric (38), can be combined to exclude the velocity perturbations. The resulting expression is simple

$$\ddot{\theta}_k + \frac{k^2}{3}\theta_k = -\frac{k^2}{3}\Phi_k + \ddot{\Phi}_k . \quad (39)$$

Note that this equation is the exact result for a pure radiation fluid. Here,  $\theta_k$  are Fourier amplitudes of  $\delta T/T$  with wavenumber  $k$ , and  $\Phi_k$  is a Fourier transform of gravitational potential. Analysis of solutions of the Einstein equations for  $\Phi$  shows that  $\Phi_k$  do not depend upon time in two important cases:

1. For superhorizon scales, which are defined as  $k\eta \ll 1$ .
2. For all scales in the case of matter dominated expansion,  $p = 0$ .

In these situations the last term in the r.h.s. of Eq. (39), namely,  $\ddot{\Phi}_k$ , can be neglected. The Einstein equations also restrict the initial conditions for fluctuations. For the adiabatic mode in the limit  $k\eta \ll 1$  one finds

$$\delta_{0k} = -2\Phi_{0k} , \quad (40)$$

where  $\delta \equiv \delta\rho/\rho$ , and subscript 0 refers to the initial values. The adiabatic mode is defined as a perturbation in the total energy density. For the one component fluid, which we consider here, only the adiabatic mode can exist. Note that fractional perturbation of the scale factor in metric (38),  $a(\eta, \mathbf{x}) = a(\eta) + \delta a(\eta, \mathbf{x}) \equiv a(\eta)(1 - \Phi)$ , can be expressed as perturbation of spatial curvature, see Eq. (4). Therefore, adiabatic perturbations are also called curvature perturbations. Let us re-write Eq. (40) for temperature perturbations:

- Radiation domination,  $\delta = 4\delta T/T$ , and we find

$$\theta_{0k} = -\frac{\Phi_{0k}}{2} . \quad (41)$$

- Matter domination,  $\delta = 3\delta T/T$ , and we find

$$\theta_{0k} = -\frac{2\Phi_{0k}}{3} . \quad (42)$$



Recall now that in the limit  $k\eta \ll 1$  the gravitational potential is time-independent,  $\Phi = \text{const}$ . Therefore, Eq. (39) has to be supplemented by the following initial conditions in the case of the adiabatic mode:

$$\theta_{0k} \neq 0, \quad \dot{\theta}_{0k} = 0. \quad (43)$$

#### 4.1.0.1 Temperature fluctuations on largest scales.

Let us consider the modes which had entered the horizon after matter-radiation equality,  $k\eta_{\text{eq}} < 1$ . For those modes,  $\dot{\Phi} = 0$  all the way from initial moments till present, and the solutions of Eq. (39) with adiabatic initial conditions is

$$\theta_k + \Phi_k = (\theta_{0k} + \Phi_{0k}) \cos\left(\frac{k\eta}{\sqrt{3}}\right). \quad (44)$$

As gravity tries to compress the fluid, the radiation pressure resists. As in everyday physics, this leads to acoustic oscillations. But here, it is important that oscillations are synchronized. All modes have the same phase regardless of  $k$ . This is a consequence of  $\dot{\theta}_{0k} = 0$ , which is valid for all  $k$ . At the last scattering, the universe becomes transparent for the radiation and we see a snapshot of these oscillations at  $\eta = \eta_{\text{ls}}$ .

To get its way to the observer, the radiation has to climb out of the gravitational wells,  $\Phi$ , which are formed at the last scattering surface. Therefore the observed temperature fluctuations are  $\theta_{\text{obs}} = \theta(\eta_{\text{ls}}) + \Phi$ , or

$$\theta_{k,\text{obs}} = \frac{1}{3}\Phi_{0k} \cos\left(\frac{k\eta_{\text{ls}}}{\sqrt{3}}\right), \quad (45)$$

where we have used Eq. (42), which relates initial values of  $\theta$  and  $\Phi$ . Note that overdense regions correspond to cold spots in the temperature map on the sky, since the gravitational potential is negative. This is famous Sachs-Wolfe effect [14].

#### 4.1.0.2 Acoustic peaks in CMBR.

Modes caught in the extrema of their oscillation,  $k_n\eta_{\text{ls}}/\sqrt{3} = n\pi$ , will have enhanced fluctuations, yielding a fundamental scale, or frequency, related to the universe sound horizon,  $s_* \equiv \eta_{\text{ls}}/\sqrt{3}$ . By using a simple geometrical projection, this becomes an angular scale on the observed sky. In a spatially flat Universe, the position of the first peak corresponds to  $l_1 \approx 200$ , see below. Both minima and maxima of the cosine in Eq. (45) give peaks in the CMBR power spectrum, which follow a harmonic relationship,  $k_n = n\pi/s_*$ , see Fig. 3, right panel.

The amplitudes of the acoustic peaks are recovered correctly after the following effects are taken into account:

1. Baryon loading. The effect of added baryons is exactly the same for the oscillator equation Eq. (44), as if we had increased the mass of a load connected to a spring, which oscillates in a constant gravitational field and with the starting point on the top of an uncompressed coil at rest. The addition of baryons makes a deeper compressional phase, and therefore increases every other peak in the CMBR power spectrum. (First, third, fifth, . . .) The CMBR power spectrum is a precise baryometer.
2. Time-dependence of  $\Phi$  after horizon crossing in radiation dominated universe. Gravitational potentials are not constant, but decay inside the horizon during radiation domination. This decay drives the oscillations: it is timed in such a way that compressed fluid has no gravitational force to fight with, when the fluid turns around. Therefore, the amplitudes of the acoustic peaks increase as the cold dark matter fraction decreases, which allows to measure  $\Omega_m$ .
3. Dissipation. This leads to a dumping of higher order peaks in the CMBR power spectrum.

#### 4.1.0.3 Position of the first peak.

Position of the first peak is determined by the angular size of the sound horizon at last scattering. Let us calculate here a similar quantity: the causal horizon (which is larger by a factor of  $\sqrt{3}$  in comparison with the sound horizon). The comoving distance traveled by light,  $ds^2 = 0$ , from the ‘‘Big Bang’’ to redshift  $z$  is determined by a relation similar to Eq. (17), but with different integration limits

$$\chi(z) = \int_z^\infty \frac{dz'}{H(z')}, \quad (46)$$

where  $H(z)$  is given by Eq. (16). One has to integrate this relation with a complete set of  $\Omega_i$ . However, from the last scattering to  $z \sim 1$ , the Universe was matter dominated. Therefore, the causal horizon in a matter dominated Universe  $\chi(z) = 2/H_0\sqrt{1+z}$  should give a reasonable first approximation to the true value of integral in Eq (46). Consider now two light rays registered at  $z = 0$  which were separated by a comoving distance  $\chi(z_{\text{ls}})$  at the moment of emission. Since both propagate in the metric  $ds^2 = a^2(d\eta^2 - d\chi^2 - \chi^2 d\theta^2) = 0$ , we find for the angular size of horizon at last scattering

$$\theta_h = \frac{\chi(z_{\text{ls}})}{\chi(0) - \chi(z_{\text{ls}})} \approx \frac{1}{\sqrt{1+z_{\text{ls}}}} = \sqrt{\frac{T_0}{T_{\text{ls}}}} \approx 1.7^\circ. \quad (47)$$

Note that this is an approximate relation since we had neglected the dark energy contribution into expansion of the Universe at late stages. To get sound horizon we have to divide Eq. (47) by  $\sqrt{3}$ . Observationally, the sound horizon angular scale is tightly constrained by Planck from the position of the first peak:  $\theta_* = 0.59648^\circ \pm 0.00018^\circ$  [2]. This is important direct observable, which can be used to set constraints on cosmological parameters entering Eq. (47).

#### 4.1.0.4 Horizon problem.

Relation (47) tells us that regions separated by more than  $> 2^\circ$  on the map of microwave sky, Fig. 3, have not been in the causal contact prior to the last scattering in the standard Friedmann cosmology. We should see  $10^4$  causally disconnected regions at the surface of last scattering. Temperature could vary wildly from point to point which are further away from each other than  $2^\circ$ . Yet, CMBR temperature is the same to better than  $10^{-4}$  accuracy all over the sky. Observations tell us that all sky regions were somehow synchronized according to the adiabatic initial conditions, Eq. (43), with only small initial perturbations present,  $\Phi_i \ll 1$ . This constitutes the so-called ‘‘Horizon problem’’ of standard cosmology. In Section 5 we will see how this problem is solved in frameworks of inflationary cosmology.

## 5 Inflationary Cosmology

In frameworks of ‘‘classical’’ cosmology and assuming no fine-tuning, one concludes that a typical universe should have had Plankian size, live Plankian time and contain just a few particles. This conclusion is based on the observation that Friedmann equations contain single dimension-full parameter  $M_{\text{Pl}} \sim 10^{19}$  GeV, while dimensionless parameters naturally are expected to be of order unity. Yet, the observable Universe contains  $10^{90}$  particles in it and had survived  $10^{65}$  Plankian times. Where does it all come from? In other words, why is the Universe so big, flat ( $\Omega_0 \approx 1$ ) and old ( $t > 10^{10}$  years), homogeneous and isotropic ( $\delta T/T \sim 10^{-5}$ ), why does it contain so much entropy ( $S > 10^{90}$ ) and does not contain unwanted relics like magnetic monopoles? These puzzles of classical cosmology were solved with the invention of Inflation [21–26]. All these questions are related to the initial conditions and one can simply postulate them. The beauty of Inflation is that it generates these unnatural initial conditions of Big Bang, while the pre-existing state (which can be arbitrary to a large extent) is forgotten. Moreover, with development, Inflationary theory delivered unplanned bonuses. Not only does the Universe become clean and homogeneous during inflation, but also the tiny perturbations necessary for the genesis of galaxies are created with the correct magnitude and spectrum. Below we consider the basics of Inflationary cosmology.

### 5.1 Big Bang puzzles and Inflationary solutions

By definition, Inflation is a period in the Universe evolution when  $\ddot{a} > 0$ . Using the second Friedmann equation, Eq. (7), we find that the inflationary stage is realized when  $p < -\rho/3$ . In particular, if  $p = -\rho$  the energy density remains constant during expansion in accord with the first law of thermodynamics, Eq. (8), and the physical volume expands exponentially fast,  $a(t) = e^{Ht}$ , see Eq. (6). Let us see now how the condition  $\ddot{a} > 0$  during some early stage solves problems of classical cosmology.

#### 5.1.0.1 Horizon problem

In Section 4.1.0.4 we have found that the angular size of horizon at the moment of last scattering is  $\approx 2^\circ$ , see Eq. (47), which tells us that we observe  $10^4$  causally disconnected regions at the surface of last scattering. The question then arises, why is the Universe so homogeneous at large scales?

This problem can be solved if during some period of time the Universe expansion was sufficiently fast. To find quantitative requirement, let us consider a power law for the Universe expansion,  $a(t) \propto t^\gamma$ . The physical size

of a given patch grows in proportion to the scale factor,  $R_P \propto a(t)$ . On the other hand, Eq. (46) tells us that the physical size of a causally connected region (horizon) grows in proportion with time,  $R_H = a\chi = t/(1 - \gamma)$ .

The exponent  $\gamma$  depends upon the equation of state,  $\gamma = 1/2$  for radiation and  $\gamma = 2/3$  for the matter dominated expansion. In any case, for the “classical” Friedmann Universe  $\gamma < 1$  and the horizon expands faster than volume. Take the largest visible patch today. It follows that in the past its physical size should have been larger than the horizon size (since they are equal today) and therefore this patch should have contained many casually disconnected regions. On the other hand, if during some period of evolution  $\gamma > 1$ , the whole visible Universe could have been inflated from one (“small”) causally connected region. In such cosmology, any given patch in the Universe passes the boundary of causally connected region twice. First when it is inflated and becomes bigger than horizon, and second, when the inflationary stage changes to “Big Bang” and casually connected region at some future point in time becomes larger than this given patch. Note that  $\gamma > 1$  means  $\ddot{a} > 0$ .

### 5.1.0.2 Curvature problem

The first Friedmann equation (6) can be re-written as

$$k = a^2 \left( \frac{8\pi G}{3} \rho - H^2 \right) = a^2 H^2 (\Omega - 1) = \dot{a}^2 (\Omega - 1). \quad (48)$$

Since  $k$  is a constant, we immediately see the problem: during matter or radiation dominated stages  $\dot{a}^2$  decreases (this happens for any expansion stage with  $\ddot{a} < 0$ ), and  $\Omega$  is driven away from unity. However, at present we observe  $\Omega \approx 1$ . Therefore, initially the Universe has to be extremely fine-tuned, say at the epoch of nucleosynthesis, when temperature was  $T \sim 1$  MeV, one should have  $|\Omega(t_{\text{NS}}) - 1| < 10^{-15}$ , and even stronger tuning is required at earlier epochs. A possible solution is obvious: accelerated expansion  $\ddot{a} > 0$  increases  $\dot{a}$  and therefore drives  $\Omega(t)$  to unity prior to radiation dominated stage. A robust prediction of inflationary cosmology is a flat Universe,  $\Omega = 1$ .

### 5.1.0.3 The problem of Entropy

As we know already, the energy of a vacuum stays constant despite the expansion. In this way, room for matter full of energy could have been created. If there is mechanism to convert vacuum energy into particles and radiation at some later stage, then the observed huge entropy will be created and the problem of entropy will be solved. Potentially, this mechanism works for any inflationary scenario, since the product  $\rho a^3$  is guaranteed to grow whenever  $\ddot{a} > 0$ . However, the important question is whether a graceful exit out of the inflationary stage and successful reheating is indeed possible. In practice, this issue has killed a number of inflationary models. Remarkably, the original model by A. Guth [23] had being ruled out precisely on these grounds [27].

### 5.1.0.4 For how long the inflationary stage should last?

Inflation has to continue for a sufficiently long time for the problems of horizon, curvature and entropy to be solved. All these requirements give roughly the same condition on the number of “e-foldings” of inflation [23] and we present here a (simplified) derivation based on entropy. Multiplying the current temperature in the universe by its visible size we find  $T_0 a_0 \chi_0 \sim 10^{30}$ , where  $\chi_0$  is the comoving size of the present horizon. We also want the whole visible universe to be inflated out of a single causally connected patch. A given wave mode is in vacuum state when its wavelength is smaller than the size of Hubble parameter during inflation and becomes frozen as a classical fluctuation when it becomes larger. This is just a consequence of the quantum field theory in a universe expanding with acceleration,  $\ddot{a} > 0$ , see Appendices. Therefore, important inflationary period, which can be linked to observations, is from the moment when the patch corresponding to the whole visible Universe goes out of the Hubble scale and to the moment when inflation ends, i.e. from the moment  $a_i \chi_0 = H^{-1}$  until  $a_r = e^{N_*} a_i$ , where the number of e-foldings,  $N_* \equiv H \Delta t$ , parametrises duration of this inflationary period. At the end of inflation the vacuum energy goes to radiation with temperature  $T_r$  which is related to the present day temperature as  $T_r a_r = T_0 a_0$ , see Eq. (12), and we neglect here the change in the number of relativistic degrees of freedom from  $T_r$  to  $T_0$ . This gives

$$\frac{T_r}{H} e^{N_*} \approx 10^{30}. \quad (49)$$

Number of e-foldings  $N_*$  depends upon reheating temperature. In popular models of Inflation the ratio  $T_r/H_i$  is within a couple orders of magnitude from unity, and we find  $50 \lesssim N_* \lesssim 60$ . I stress again that  $N_*$  is not the duration of inflation. The latter cannot be smaller than  $N_*$ , but inflation can last longer of course, and then our Universe is homogeneous to scales much much larger than its visible part today.

## 5.2 Models of Inflation

Consider energy-momentum tensor  $T_{\mu\nu}$  for a scalar field  $\varphi$

$$T_{\mu\nu} = \partial_\mu\varphi \partial_\nu\varphi - g_{\mu\nu} \mathcal{L}, \quad (50)$$

with the Lagrangian

$$\mathcal{L} = \partial_\mu\varphi \partial^\mu\varphi - V(\varphi). \quad (51)$$

In a state when all derivatives of  $\varphi$  are zero, the stress-energy tensor of a scalar field simplifies to  $T_{\mu\nu} = V(\varphi) \delta_{\mu\nu}$ . This corresponds to a vacuum state. Indeed, comparing with Eq. (5), we find  $V = \rho = -p$ . A large number of inflationary models exists where  $\varphi \approx \text{const}$  during some period of evolution and vacuum-like state is imitated. Such field is called inflaton.

1. *False vacuum inflation.* Conceptually simple and easily understandable scenario was suggested by A. Guth [23]. Consider potential  $V(\phi)$  which has a local minimum with a non-zero energy density separated from the true ground state by a potential barrier. A universe which happened to be trapped in the meta-stable minimum will stay there for a while (since such a state can decay only via subbarrier tunneling) and expansion of the universe will diminish all field gradients. Then the Universe enters a vacuum state and Inflation starts. Subsequent phase transition into the true minimum ends inflationary stage and creates the radiation phase. Today the model of Guth and its variants based on potential barriers is good for illustration purposes only. It did not stand up to observations since inhomogeneities which are created during the phase transition into the radiation phase are too large [27]. But the model gives easily understandable answer to the frequently asked question: how can it be that the energy density stays constant despite the expansion?

2. *Chaotic inflation.* Andrei Linde was first to realize that things work in the simplest possible setup [26]. Consider potential

$$V(\phi) = \frac{1}{2} m_\phi^2 \phi^2. \quad (52)$$

Field equation in an expanding Universe and for the homogeneous mode is  $\ddot{\phi} + 3H\dot{\phi} + m_\phi^2\phi = 0$ . If  $H \gg m_\phi$ , the “friction”  $3H\dot{\phi}$  dominates and the field does not move (almost). Therefore, time derivatives in  $T_{\mu\nu}$  can be neglected, and inflation starts (in a sufficiently homogeneous patch of the Universe). A Hubble parameter in this case is determined by the potential energy,  $H \approx m_\phi\phi/M_{\text{Pl}}$ , and we see that inflation starts if the initial field value happened to satisfy  $\phi > M_{\text{Pl}}$ . During inflationary stage the field slowly rolls down the potential hill. This motion is very important in the theory of structure creation, see Appendices. Inflation ends when  $\phi \sim M_{\text{Pl}}$ . At this time, field oscillations start around the potential minimum and later decay into radiation. In this way all matter content has been likely created in our Universe. In general, this model generalises to arbitrary monomial  $V(\phi) \propto \phi^\alpha$  as field potential at large  $\phi$ .

3.  *$R^2$  - inflation.* Historically, this is the first model of inflation. It was invented by A. Starobinsky [21, 22]. Einstein-Hilbert action, leading to Einstein equations (1) should be modified inevitably in quantum field theory on a curved space-time. In particular, counter-terms proportional to the squares of different curvature tensors should be added to cancel divergences. Starobinsky considered the simplest form of extended gravitational Lagrangian

$$L = \frac{M_{\text{Pl}}^2}{2} R + \beta R^2, \quad (53)$$

where  $R$  - scalar curvature and  $\beta$  - some dimensionless constant. Universe inflates in this model. It can be understood as follows. After conformal rotation this model is equivalent to the usual Einstein gravity plus a scalar field with potential

$$V(\phi) = \Lambda^4 \left( 1 - e^{-\sqrt{2/3} \phi/M_{\text{Pl}}} \right)^2.$$

This potential has very flat plateau at  $\phi > M_{\text{Pl}}$ , and with such initial value of  $\phi$  the Universe will inflate. The Universe will be heated up in the same way is in chaotic inflation after  $\phi$  will slowly reach  $M_{\text{Pl}}$ .

## 5.3 Unified theory of Creation

During Inflation the Universe was in a vacuum-like state. We have to figure out how this “vacuum” had been turned into the matter we observe around us, and how primordial fluctuations which gave rise to galaxies were created. Solution to all these problems can be understood in a single unified approach. Basically, everything reduces to

a problem of particle creation in a time-dependent classical background. On top of every “vacuum” there are fluctuations of all quantum fields which are present in a given model. This bath of virtual quanta is indestructible, and even Inflation cannot get rid of it. Being small, fluctuations obey an oscillator equation

$$\ddot{u}_k + [k^2 + m_{\text{eff}}^2(\eta)] u_k = 0, \quad (54)$$

here  $u_k$  are amplitudes of fluctuating fields in Fourier space. Effective mass becomes time dependent through the coupling to time-dependent background. Because  $m_{\text{eff}}$  is time dependent, it is not possible to keep fluctuations in a vacuum. If oscillators with momentum  $k$  happened to be in the vacuum at one time, they will not be in the vacuum at a latter time because positive and negative frequency solutions mix, see Appendices. Several remarks are in order.

- Eq. (54) is valid for all particle species.
- The equation looks that simple in a conformal reference frame  $ds^2 = a(\eta)^2 (d\eta^2 - dx^2)$ . (And a “dot” means derivative with respect to  $\eta$ .)
- Of particular interest are ripples of space-time itself: curvature fluctuations (scalar fluctuations of the metric) and gravity waves (tensor fluctuations of the metric).
- Effective mass  $m_{\text{eff}}$  can be non-zero even for massless fields. Gravitational waves give the simplest example [28], with  $m_{\text{eff}}^2 = -\ddot{a}/a$ . The effective mass for curvature fluctuations has a similar structure  $m_{\text{eff}}^2 = -\ddot{z}/z$ , but with  $a$  being replaced by  $z \equiv a\dot{\phi}/H$ , see Refs. [29–32].
- For a scalar field which does not couple to the inflation, the effective mass is given by Eq. (67). For conformally coupled, but massive scalar it reduces to  $m_{\text{eff}} = m_0 a(\eta)$ .

Note that creation in Inflationary theory is possible because nature is not conformally-invariant. Otherwise,  $m_{\text{eff}}$  would be time-independent and vacuum would remain vacuum forever. There are two important instances of time varying classical background in cosmology: expansion of space-time,  $a(\eta)$ , and motion of the inflaton field,  $\phi(\eta)$ . Both can be operational separately or together at any epoch of creation:

- During inflation. This is when superhorizon size perturbations of metric are created, which give seeds for the formation of galaxies and Large Scale Structure in general.
- After inflation while the inflaton oscillates. This is when matter itself is created out of energy generated from the vacuum.

There are several primary observables which can be calculated out of  $u_k$  and further used for calculation of quantities of interest. Most useful are:

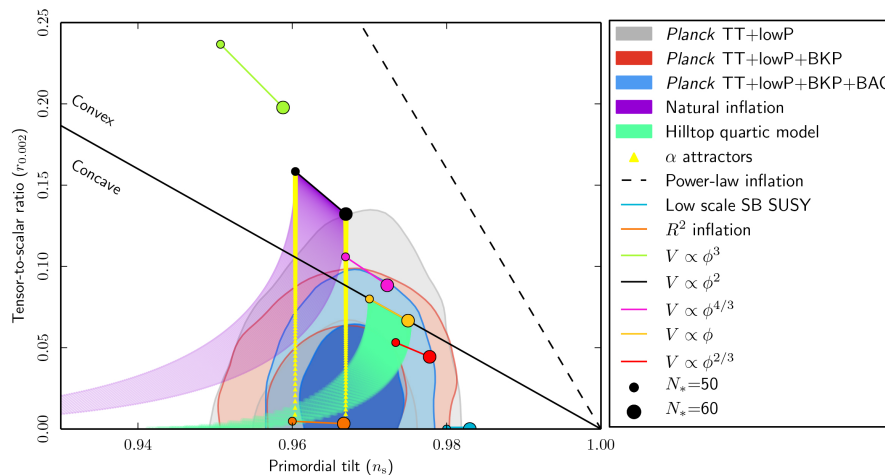
- The particle occupation numbers,  $n_k$ . Integration over  $d^3k$  gives the particle number density.
- The power spectrum of field fluctuations,  $P(k) \equiv u_k^* u_k$ . Integration over  $d^3k$  gives the field variance.

Depending on physical situation, only one or the other may have sense. The particle number in a comoving volume is useful because it is adiabatic invariant on sub-horizon scales (or when  $m > H$ ) and allows to calculate the amount of created matter and abundances of various relics, e.g. dark matter. But it has no meaning at super-horizon scales when  $m < H$ . Then the power spectrum of field fluctuations is used which allows to calculate density perturbations and gravitational waves generated during inflation. Necessary details of such calculation are given in Appendices.

### 5.3.1 Testing Inflationary predictions

Typically, the spectrum of curvature perturbations generated during inflations has a form  $P_\zeta(k) = A_s k^{n_s}$ , where  $A_s$  and  $n_s$  are constants (i.e. weakly depend upon scale  $k$ ). Similarly, for gravitational waves  $P_T(k) = A_T k^{n_T}$ . To the first approximation, the Hubble parameter  $H$  during inflation is constant. Then, power spectra do not depend on  $k$  and  $n_s = 1$ ,  $n_T = 0$ . This case is called the Harrison-Zel’dovich spectrum [35, 36] of primordial perturbations which has been suggested on general grounds before inflationary theory was invented. However, in reality,  $H$  is changing and these constants take different, model dependent values. Nevertheless, there is model independent relation between the slope of tensor perturbations and the ratio of power in tensor to curvature modes

$$r \equiv \frac{P_T(k)}{P_\zeta(k)} = -8n_T. \quad (55)$$



**Fig. 5:** Constraints on inflationary models in the  $(n_s, r)$  plane, from Ref. [34]. Coloured line segments with circles at ends correspond to predictions of different inflationary models with different inflation potentials. Within each segment  $N_*$  varies in the interval  $50 \leq N_* \leq 60$ , see Eq. (49).

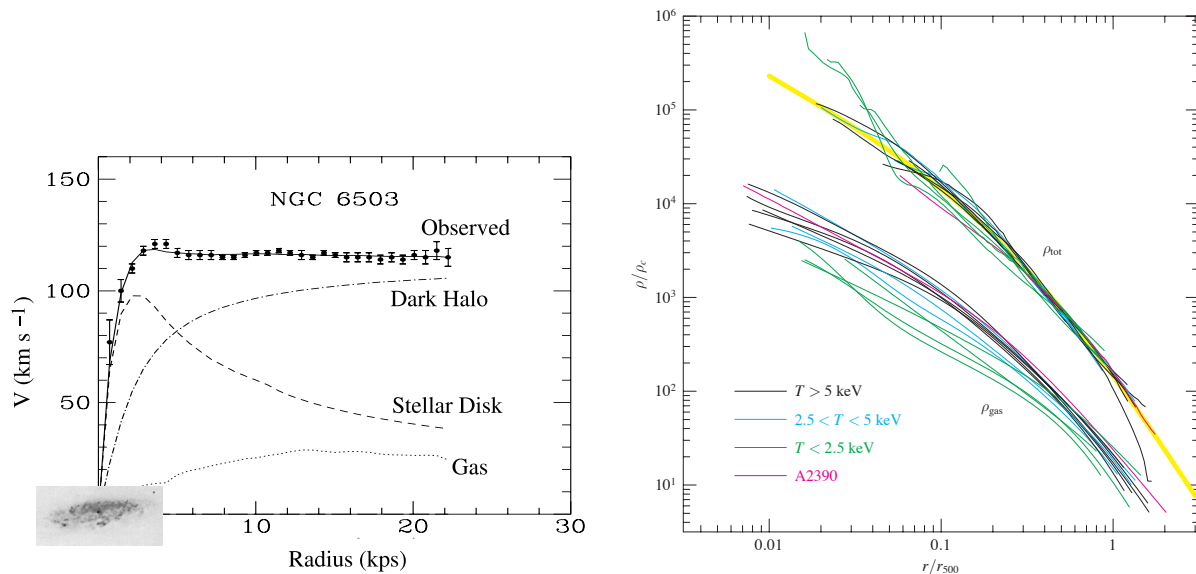
This is called the *consistency relation* to which (simple) inflationary models should obey. It will be robust and ultimate test of inflationary theory when imprint of gravitational waves in a CMBR will be discovered.

However, both  $A_s$  and  $n_s$  are measured, extracted from CMBR observations and can be compared to model predictions. The most recent constraints in the  $(n_s, r)$  plane, obtained by Planck collaboration [34] are presented in Fig. 5. We see that in chaotic inflationary model, Eq. (52), the gravitational waves would have been already discovered by Planck, and this model is ruled out nowadays. Best runner is  $R^2$  model of A. Starobinsky, Eq. (53), which is a perfect fit. However, observation of the imprint of gravity waves in this model will be very difficult task, if possible at all.

To summarise, all predictions of Inflationary cosmology, which could have being tested so far, have being confirmed. In particular, the Universe is spatially flat (within experimental errors), see Table I. The primordial perturbations are of superhorizon size and adiabatic. The spectral index is close to the Harrison-Zeldovich case, see Fig. 5. Crucial test of inflationary paradigm would be detection of gravity waves and verification of the consistency relation.

## 6 Dark Matter

We have seen already in Section 4 that CMBR observations accurately measure the nature and spectrum of the primordial fluctuations, the geometry of the Universe, its present expansion rate and its composition, see also Table 2, which is based on these observations. In particular, these measurements tell us that in addition to baryonic matter there should be also dark matter which so far had been seen only through its gravitational influence. This "sterility" leaves open the possibility that in fact we should look for modification of gravity, not for dark matter, in order to explain the missing mass problem. While both possibilities are exciting and beyond contemporary physics, a successful modified gravity theory was not constructed yet. Therefore, I will not discuss numerous attempts and various models of modified gravity here, instead I'll just give two original references, the early one [37], and the most recent one [38]. It is difficult to construct such a theory for several reasons. In particular, the evidence for missing mass exists at various scales and epochs while modification should explain everything. Contrary to that, e.g. simple variants of MOND [37] do explain the "missing mass" on galactic scales without invoking dark matter, but fail to explain other evidence. Moreover, MOND is phenomenological, non-relativistic prescription, not a theory. Therefore, other cosmological tests, beyond CMBR, are also important. Below we consider cosmological observations that are independent of the CMB but also point to the existence of non-baryonic dark matter. At the end of the section I briefly discuss some popular models of dark matter and present status of dark matter searches in corresponding models.



**Fig. 6:** *Left panel:* Rotational curve of the galaxy NGC6503. I superimposed the optical image of corresponding galaxy with the rotational curve, approximately to scale in radius. *Right panel:* Scaled cluster density profiles extracted from X-ray observations of different clusters, from Ref. [45].

## 6.1 Dark Matter: the evidence

Missing mass is seen on all cosmological scales. In particular, it reveals itself as

- Flat rotational curves in galaxies;
- Gravitational potential which confines galaxies and hot gas in clusters;
- Gravitational lenses in clusters;
- Gravitational potential which allows structure formation from tiny primeval perturbations;
- Gravitational potential which creates CMBR anisotropies.

In this subsection I shortly review this overwhelming evidence for the unseen, but gravitating mass.

### 6.1.1 Dark Matter in Galaxies

Consider a test particle which is orbiting a body of mass  $M$  at a distance  $r$ . Within the frameworks of Newtonian dynamics the velocity of a particle is given by

$$v_{\text{rot}} = \sqrt{\frac{GM(r)}{r}}. \quad (56)$$

Outside of the body, the mass does not depend on distance, and the rotational velocity should obey the Kepler law,  $v_{\text{rot}} \propto r^{-1/2}$ . Planets of the Solar system obey this law. However, this is not the case for stars or gas which are orbiting galaxies. Far away from the visible part of a galaxy, rotational curves are still rising or remain flat. An example is shown in Fig. 6, left panel. An optical image of the NGC6503 galaxy is superimposed with its rotational curve, approximately to correct scale. The contribution of visible baryons in the form of stars and hot gas can be accounted for, and the expected rotational curve can be constructed. The corresponding contributions are shown in Fig. 6, left panel. One can see that the data-points are far above the contribution of visible matter. The contribution of missing dark mass, which should be added to cope with data, is also shown and is indicated as Dark Halo. For the rotational velocity to remain flat, the mass in the halo should grow with the radius as  $M(r) \propto r$ , i.e., the density of dark matter in the halo should decrease as  $\rho(r) \propto r^{-2}$ .

### 6.1.2 Dark matter density profiles.

To interpret what is seen in the data, in particular, to interpret the results of direct and indirect dark matter searches, and to plan for further strategy, it is important to know the expected phase-space structure of the dark halo and corresponding dark matter density profiles. For interacting particles a thermal distribution over energies is eventually established. However, in conventional cold dark matter models, particles are non-interacting, except gravitationally. Binary gravitational interactions are negligible for elementary particles, and resulting phase-space distributions are not unique, even for a stationary equilibrium states, and even if constraint to the flat rotational curves is enforced. Below I highlight several such distributions, which are often discussed in the literature and are used in applications.

1. The simplest self-gravitating stationary solution which gives flat rotational curves corresponds to an “isothermal sphere” with Maxwellian distribution of particles over velocities:

$$n(\vec{r}, \vec{v}) = n(r) e^{-v^2/v_0^2}. \quad (57)$$

Solution of the equation of hydrostatic equilibrium can be approximated by the density profile

$$\rho(r) = \frac{\rho_0}{(1+x^2)}, \quad \text{where } x \equiv r/r_c. \quad (58)$$

It should be stressed that the distribution Eq. (57), in contrast to a distribution in real thermal equilibrium, depends on particle velocities, not on their energies. Such distributions may arise in time-dependent gravitational potential as a result of collisionless relaxation.

2. There exist several density profiles which are empirical fits to numerical simulations, most often used is Navarro, Frenk & White (NFW) profile [39].

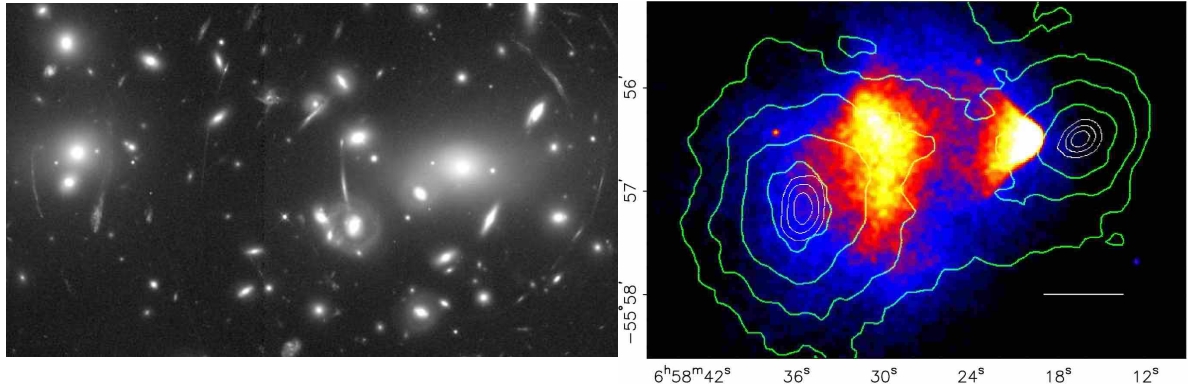
$$\rho(r) = \frac{\rho_0}{x(1+x)^2}. \quad (59)$$

3. In the CDM model, the distribution of dark matter particles in the phase space during initial linear stage prior to structure formation corresponds to thin hypersurface,  $\mathbf{v} = H\mathbf{r}$  (this is just Hubble law). Since during collisionless evolution the phase-space density conserves as a consequence of the Liouville theorem, then even at the non-linear stage the distribution will still be a thin hypersurface. It can be deformed in a complicated way and wrapped around, but it cannot tear apart, intersect itself, puff up or dissolve. The resulting idealised phase-space distribution describing isolated dark halo has been studied in Ref. [40] for the case of spherical symmetry. Initial thin hypersurface gets wrapped around indeed, forming large and ever increasing number of folds in the phase space in the inner galaxy. Existence of such a folded structure is a topological statement. Corresponding model is also called “infall model”. It reproduces flat rotational curves, but there are several interesting differences to other distributions, though. Rotational curves of the infall model have several small ripples which appear near caustics in the dark matter distribution. (Those are places in the phase-space where particles turn-around and have zero velocity). It is important to note that caustics may be observable and their discovery in real data will prove that the missing mass is dark matter indeed, not a modification of gravity, even if dark matter particles will not be directly identified. The energy spectrum of dark matter particles at a fixed position also deviates from other distributions. This may be important for the direct dark matter searches. Also, the infall model gives the insight [40] on why the empirical fit, the NFW profile, has this particular form, Eq (59). Observationally, signature of the infall is seen e.g. in our local group of galaxies [41], but at largest distances, outside of caustics. Caustics which are furthest away from the galaxy centre were resolved recently in the high-resolution N-body modelling of galaxy formation [42]. To understand how far the folded structure of the infall model continues into the inner halo in a galaxy like our own will require even larger simulations with better resolution.

### 6.1.3 Dark Matter in Clusters of Galaxies

Already in 1933, F. Zwicky [43] deduced the existence of dark matter in the Coma cluster of galaxies. Nowadays, there are several ways to estimate masses of clusters: based on the kinetic motion of member galaxies, on X-ray data, and on gravitational lensing. These methods are different and independent. In the dynamical method, it is assumed that clusters are in virial equilibrium, and the virialized mass is easily computed from the velocity dispersion. In X-ray imaging of hot intracluster gas, the mass estimates are obtained assuming hydrostatic equilibrium. Mass estimates based on lensing are free of any such assumptions. All methods give results which are consistent with each other, and tell that the mass of the luminous matter in clusters is much smaller than the total mass.





**Fig. 7:** *Left panel:* Image of the cluster Abell 2218 taken with the Hubble space telescope (see Ref. [46]). Spectacular arcs resulting from strong lensing of background galaxies by dark matter in the cluster are clearly seen. *Right panel:* Deep Chandra image of the Bullet cluster. The X-ray brightness of the gas component is coded in yellow, red and blue colours. Distribution of the gravitating mass, obtained from weak lensing reconstruction, is shown by green contours, from Ref. [50].

Recent review on basic properties of clusters and their role in modern astrophysics and cosmology can be found e.g. in [44].

*Kinetic mass estimates.* Those are based on the virial theorem,  $\langle E_{\text{pot}} \rangle + 2\langle E_{\text{kin}} \rangle = 0$ . Here  $\langle E_{\text{kin}} \rangle$  is averaged kinetic energy of a constituents in the gravitationally bound object (e.g. cluster of galaxies) and  $\langle E_{\text{pot}} \rangle$  is their averaged potential energy. Measuring the velocity dispersion of galaxies in the clusters and its geometrical size gives an estimate of the total mass,

$$M \sim \frac{2\langle r \rangle \langle v^2 \rangle}{G}. \quad (60)$$

The result can be expressed as mass-to-light ratio,  $M/L$ , using the Solar value of this parameter. For the Coma cluster, which consists of about 1000 galaxies, Zwicky [43] has found

$$M/L \sim 300 h (M_{\odot}/L_{\odot}). \quad (61)$$

Modern techniques end up with very much the same answer.

*Mass estimates based on X-rays.* Mass is also traced in clusters of galaxies by the hot gas which is visible in X-rays. Assume hot gas is in thermal equilibrium in a gravitational well created by a cluster. Then, cluster mass profiles can be derived from the gas density and temperature as functions of the distance to a cluster centre. This independent method has its own advantages and disadvantages. With respect to galaxy dynamics (see above) or lensing mass estimates (see below), this method has the advantage of being less sensitive to projection effects along the line of sight through the cluster. However, validity of the assumptions of ICM hydrostatic equilibrium and spherical symmetry of the cluster gravitational potential wells may depend on the evolutionary state of the cluster.

As an example, the radial density profiles derived in Ref. [45] from the Chandra X-ray satellite data are shown in Fig. 6, right panel. We see that dark matter density exceeds gas density by an order of magnitude at any value of the radius. Dark matter density as a function of radius is well fitted by NFW profile, Eq. (59), which is shown by thick yellow line. Total gas mass fractions varies between 5 and 15 percent from cluster to cluster and systematically depends upon cluster mass. These values are somewhat lower than the Universal baryon fraction suggested by the CMB observations, but approaches it for the heaviest clusters.

*Gravitational Lensing.* As photons travel from a background galaxy to the observer, their trajectories are bent by mass distributions. This effect of gravitational lensing allows direct mass measurement without any assumptions about the dynamical state of the cluster. The method relies on the measurement of the distortions that lensing induces on the images of the background galaxies, an example of such distortions is shown in Fig. 7, left panel. A reconstruction of lens geometry provides a map of the mass distribution in the deflector. For a review of the method see e.g. Ref [47]. The images of extended sources are deformed by the gravitational field. In some

cases, the distortion is strong enough to be recognized as arcs produced by a galaxy cluster serving as a lens, see Fig. 7, left panel. For the cluster A 2218, shown in this figure, Squires et al. [48] compared the mass profiles derived from weak lensing data and from the X-ray emission. The reconstructed mass map qualitatively agrees with the optical and X-ray light distributions. A mass-to-light ratio of  $M/L = (440 \pm 80)h$  in solar units has been derived. The gas to total mass ratio was found to be  $M_{\text{gas}}/M_{\text{tot}} = (0.04 \pm 0.02)h^{-3/2}$ . The radial mass profile agrees with the mass distribution obtained from the X-ray analysis. For a recent study of mass density profiles of galaxy clusters derived from the gravitational lensing see e.g. Ref. [49]. A sample of 50 galaxy clusters at  $0.15 < z < 0.3$  has been studied. Again, dark matter density as a function of radius is perfectly fitted by the NFW profile, Eq. (59), but "isothermal" profile is a bad fit.

*Dark matter or modification of gravity?* In principle, the excess gravitational force, undoubtedly observed in galaxies and clusters of galaxies, could be not a manifestation of the Dark Matter, but may have origin in some modification of Einstein gravity. Gravitational lensing studies of the Bullet Cluster 1E 0657-56 are claimed [50] to provide the best evidence to date for the existence of dark matter, as opposed do modifications of gravity. The Bullet Cluster consists of two colliding clusters of galaxies. Reconstructed distribution of the gas, stars and gravitating matters shown in Fig. 7, right panel. The X-ray brightness of the hot gas is coded in yellow, red and blue colours. Distribution of the gravitating mass is shown by green contours and was obtained from weak lensing reconstruction. It coincides with distribution of stars, but counts of stars gives small contribution to the overall mass balance. The hot gas of the two colliding components, seen in X-rays, represents most of the baryonic, i.e. ordinary, matter in the cluster pair. The hot gas in this collision was slowed down by a drag force. In contrast, the dark matter or stars were not slowed by the impact, because they do not interact strongly with itself or the gas except through gravity. This produced the separation of gravitating matter and gas seen in the data. If hot gas was the most massive component in the clusters, and dark matter would be absent, as proposed by alternative gravity theories, such a separation would not have been seen. Therefore, dark matter is required to explain what is seen here.

#### 6.1.4 Structure formation and DM

By present time the structures in the Universe (i.e. galaxies and clusters) are formed already, in other words perturbations in matter have entered non-linear regime,  $\delta\rho/\rho \gtrsim 1$ . However, the initial perturbations were small  $\delta\rho/\rho \sim 10^{-5}$ , as we know from measurements of temperature fluctuations in CMBR, see Section 4. Perturbations do not grow significantly in the radiation dominated epoch, they can start growing only during matter domination and are growing then in proportion to the scale factor,  $\delta\rho/\rho \sim a = 1/z$ . Moreover, baryonic plasma is tightly coupled to radiation, therefore perturbations in baryonic matter start to grow only after recombination. For the same reason, initial perturbations in baryons at the time of recombination are equal to fluctuations in CMBR. If baryons were to constitute the only matter content, then perturbations in matter at present time would be equal to

$$\frac{\delta\rho}{\rho}|_{\text{today}} = z_{\text{rec}} \frac{\delta\rho}{\rho}|_{\text{rec}} \sim 10^{-2}, \quad (62)$$

where  $z_{\text{rec}} \approx 1100$  is the redshift of recombination. This apparent contradiction is resolved by the dark matter. In our Universe structure has had time to develop only because perturbations in non-baryonic dark matter have started their growth prior to recombination. Baryonic matter then "catch up" simply by falling into already existing gravitational wells. This is one of the strongest and simplest arguments in favour of non-baryonic dark matter.

## 6.2 Dark Matter: particle candidates

Cosmology tells us that the Standard Model of particle physics is incomplete. The model which will extend it should contain particles which would constitute non-baryonic dark matter. And there should exist some mechanism to produce it with correct abundance,  $\Omega_{\text{DM}} \approx 0.27$ , see Table 2. Also, trusted and popular DM candidates appear naturally in the models whose origin is unrelated to the dark matter problem. There is no shortage of particle physics models which obey those requirements, with the huge range of DM particle masses and very different production mechanisms. Some dark matter particle candidates are listed in Table 3. Given concrete model of particle physics, a theorist should first calculate the cosmological abundance of DM produced in the model in hands. Below, in the subsection 6.3, I give some examples of such calculations to highlight various mechanisms of DM production. Then, in the subsection 6.4, I briefly describe vast topic of direct and indirect searches for most popular DM candidates, with corresponding derived constraints.

**Table 3:** Dark Matter particle candidates

| candidate        | mass          | some refs |
|------------------|---------------|-----------|
| Graviton         | $10^{-21}$ eV | [51]      |
| Axion            | $10^{-5}$ eV  | [52]      |
| Sterile neutrino | 10 keV        | [53]      |
| Mirror matter    | 1 GeV         | [54]      |
| WIMP             | 100 GeV       | [55]      |
| WIMPZILLA        | $10^{13}$ GeV | [56]      |

### 6.3 Production mechanisms

Depending upon production mechanism, the resulting dark matter can appear as 'cold', 'warm' or 'hot'. Loosely speaking, velocities of cold dark matter are so small that they are not influencing the large scale structure formation at all. Velocities of hot dark matter particles are too big. Their kinetic energy does not allow particles to clump galaxy halos and may smear out even clusters of galaxies. Such DM is ruled out. Warm dark matter is the intermediate case. It may wash out structure at smallest observable scales of dwarf galaxies but does not influence formation of big haloes like our Milky Way. Cold dark matter models have some problems explaining observations at small scales, the warm dark matter models have some advantages here, see below.

Further, dark matter particle candidates can be divided into several classes according to a mechanism of their production in the early universe. We start with popular class of DM candidates referred to as "thermal relics".

#### 6.3.1 Cosmological abundance of thermal relics

By definition, a thermal relic is assumed to be in thermodynamic equilibrium at early times. At some point in the evolution particles go out of equilibrium and after that their number in a comoving volume remains constant. The process is called "freeze-out". For thermal relics it is just the value of particle mass which determines if it will be hot, warm or cold. To see this, let us define the *free streaming length* for a given DM particle species with mass  $M_X$  as a horizon size at a temperature when particles are still relativistic, i.e. at  $T \sim M_X$ . Clearly, structure will be washed out at all scales smaller than this. Later on, particles are non-relativistic and cannot move much farther away. Structure is preserved at larger scales. Horizon size at  $T \sim M_X$  expanded to present epoch is given by

$$L_{fs} \sim \frac{M_{Pl}}{T_0 M_X}.$$

For  $M_X \sim 1$  eV this gives  $L_{fs} \sim 100$  Mpc. Clearly, models with such a big free streaming length are ruled out. On the other hand, for  $M_X \sim 1$  keV we find  $L_{fs} \sim 0.1$  Mpc. This corresponds to the size of a dwarf galaxy. Therefore, this gives the lower bound for the warm DM particle mass:

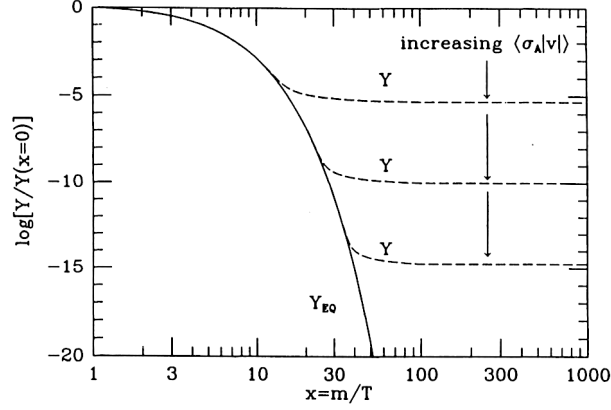
$$M_X > 1 \text{ keV}.$$

For thermal relics the resulting dark matter will be definitely cold if freeze-out occurs when particles are non-relativistic, i.e. at temperatures smaller than particle mass,  $T < M_X$ . WIMP, and in particular neutralino, appearing in supersymmetric models, belong to this class of dark matter.

A given particle species will track the equilibrium abundance as long as reactions which keep them in chemical equilibrium can proceed rapidly enough. Here, "rapidly enough" means that the mean free time between interactions is smaller than the age of the universe,  $\tau < t_u$ . This condition can also be written as  $n\sigma v > H$ . In thermal equilibrium, the number density of non-relativistic particles is given by Eq. (26). In this regime the number density decreases exponentially fast with decreasing temperature  $T$ . When the rate of reactions becomes lower than the expansion rate, the particles can no longer track the equilibrium value and thereafter particle concentrations in a comoving volume remain constant. Clearly, the more strongly interacting particles are, the longer they stay in equilibrium, and the smaller their freeze-out abundance will be, see Fig. 8. Here we defined particle abundance as the ratio of particle number to the entropy density,  $Y \equiv n/s$ .

Freeze-out concentration  $n$  is determined by the condition  $n_f \sigma v \approx H$ , or (neglecting numerical factors)

$$n_f \approx \frac{H}{\langle \sigma v \rangle} \approx \frac{T_f^2}{M_{Pl} \langle \sigma v \rangle}.$$



**Fig. 8:** A schematic view of comoving number density of a stable species as they evolve through the process of thermal freeze-out.

After freeze-out the ratio of  $n$  and entropy density  $s$  remains constant. In particular, present density is given by  $n_0 = n_f s_0 / s_f$ . Therefore

$$\Omega_{DM} \equiv \frac{mn_0}{\rho_c} = \frac{mn_f s_0}{s_f \rho_c} \sim \frac{m}{T_f} \frac{1}{\langle\sigma v\rangle} \frac{T_0^3}{M_{Pl}\rho_c}.$$

Freeze-out temperature  $T_f$  cannot go much below particle mass  $m$ , see Fig. 8. One gets  $x_f \equiv m/T_f = 20 - 30$  for all practically interesting values of annihilation cross-section. Restoring now all numerical factors in the above estimate we obtain

$$\Omega_{DM} = \frac{16\pi^2}{3} \sqrt{\frac{\pi}{45}} \frac{x_f g_0}{\sqrt{g_*(T_f)}} \frac{T_0^3}{M_{Pl}^3 H_0^2} \frac{1}{\langle\sigma v\rangle}. \quad (63)$$

For the  $s$ -wave annihilation  $\sigma = \sigma_0/v$  and we have numerically

$$\Omega_{DM} \approx 0.2 \frac{\text{pb}}{\sigma_0}.$$

Note that a picobarn crosssections are in the ballpark of the electroweak scale,  $\text{pb} \approx \alpha^2/(100 \text{ GeV})^2$ . That is why the weakly interacting massive particles, appearing e.g. in supersymmetric extensions of the Standard Model are considered to be natural candidates for the dark matter. Another useful parametrisation of this result is given by

$$\Omega_{DM} \approx 0.2 \frac{3 \cdot 10^{-26} \text{ cm}^3 \text{ s}^{-1}}{\langle\sigma v\rangle}. \quad (64)$$

This expression is used for the discussion of dark matter direct and indirect search results and strategies.

### 6.3.2 Cosmological abundance of ultra-light bosons

Dark matter particles can be very light and still very cold if they did not originate from the thermal bath. Of course this holds for bosons only, since the phase-space restrictions will not allow light fermions to saturate required energy density in galaxy halos. Corresponding constraint on fermions is called Tremaine-Gunn limit [58] and reads  $M_F \gtrsim 1 \text{ keV}$ .

To illustrate the general idea, let us consider a scalar field with potential  $V(\phi) = m^2\phi^2/2$ . The field equations for the Fourier modes with a momentum  $k$  in an expanding Universe are

$$\ddot{\phi}_k + 3H\dot{\phi}_k + (k^2 + m^2)\phi_k = 0. \quad (65)$$

Since the term  $\propto H$  can be understood as a friction, amplitude of modes with  $9H^2 \gg (k^2 + m^2)$  (almost) does not change with time. Then, the oscillations of modes with a given  $k$  commence when  $H$  becomes sufficiently small,  $9H^2 \ll (k^2 + m^2)$ . Oscillating modes behave like particles, and their amplitude decreases with expansion. Since modes with the largest  $k$  start oscillations first, they will have the smallest amplitude and the field becomes

homogeneous on a current horizon scale. This holds while mass term is unimportant, i.e. till  $3H > m$ . Modes with all  $k < 3H$  will start oscillations simultaneously when  $3H \approx m$ , and will behave like cold dark matter since then.

Resulting abundance of dark matter will depend upon initial amplitude of modes with  $k < 3H$ . Why the initial amplitude of such modes is non-zero in the first place? Such fields are generated during inflation if  $m$  is smaller than the value of the Hubble constant during inflation, see Eq. (A.3). In this way e.g. massive gravitons are created as a dark matter, see Ref. [51].

Situation in the case of axions is even easier to understand. Potential for the axion field  $a$  has the following form

$$V(a) = m_a^2 f_a^2 (1 - \cos(a/f_a)).$$

Axion mass is temperature dependent,  $m_a = m_a(T)$  and at  $T \gg 1$  GeV it is zero. Therefore, at this temperatures  $V(a) = 0$  and the axion field takes arbitrary values in the range  $0 < a/f_a < 2\pi$ . Field oscillations start with amplitude  $a \sim f_a$  at  $T \approx 1$  GeV when  $3H(T) = m_a(T)$ . Correct axion abundance is obtained for  $10^{-5} \lesssim m_a \lesssim 10^{-3}$ . Note that the field will be homogeneous on the horizon scale at  $T > 1$  GeV, but may be inhomogeneous on larger scales. This may lead to formation of dense clumps, ‘‘axion miniclusters’’ of the mass  $M \sim 10^{-12} M_\odot$  [59].

### 6.3.3 Cosmological abundance of superheavy particles

Superheavy particles can be created purely gravitationally. As we have seen in Section 5.3, generically, a quantum field cannot be kept in a vacuum in the expanding universe. This can be understood on the example of a scalar field, Eq. (65). In conformal time  $\eta$ , Eq. (14), and for rescaled field,  $u_k \equiv \phi_k a$ , the mode equations take form of an oscillator equation

$$\ddot{u}_k + [k^2 + m_{\text{eff}}^2(\eta)] u_k = 0, \quad (66)$$

with time-dependent mass

$$m_{\text{eff}}^2(\eta) = a^2 m^2 - \frac{\ddot{a}}{a} (1 - 6\xi). \quad (67)$$

This is one particular case of the general situation described by Eq. (54). The constant  $\xi$  describes the coupling to the scalar curvature, the corresponding term in the Lagrangian is  $\xi R\phi$ . The value  $\xi = 0$  corresponds to minimal coupling (Eq. (65) was written for this case), while  $\xi = 1/6$  is the case of conformal coupling. Equations for massless, conformally coupled quanta are reduced to the equation of motion in Minkowski space-time. Particle creation does not occur in this case. For massive particles, conformal invariance is broken and particles are created regardless of the value of  $\xi$ . Let us consider the case of  $\xi = 1/6$ . It is the particle mass which couples the system to the background expansion and serves as the source of particle creation in this case. Therefore, we expect that the number of created particles in comoving volume is  $\propto m^3$  and the effect is strongest for the heaviest particles. In inflationary model (52) the abundance of created particles,  $\Omega_{\text{SH}}$ , will match observations if  $m \sim 10^{13}$  GeV [56], precise value of required superheavy particle mass depends upon reheating temperature and the value of  $\xi$ . Therefore, a dark matter can be created in the early Universe even if it has no couplings at all, the only condition reads: be superheavy.

### 6.3.4 Cosmological abundance of sterile neutrino

Active neutrino are massive, this fact signifies a new physics beyond the Standard Model. Other fermions have masses because they exist as left handed and right handed states with coupling to the Higgs field  $H$ . However, active neutrinos are left-handed. Therefore, a natural way to generate masses for the neutrino would be to consider them at the same footing as other fermions and to add right handed neutrinos,  $N_j$ , to the Standard Model Lagrangian,

$$\mathcal{L} = \mathcal{L}_{\text{SM}} + i\bar{N}_j \partial_\mu \gamma^\mu N_j - \left[ \lambda_{ji} (\bar{L}_i H) N_j + \frac{M_j}{2} \bar{N}_j^c N_j + \text{h.c.} \right]. \quad (68)$$

Flavour indexes  $j$  may run from one to three, but not necessarily. In what follows I omit explicit writing of indexes. In the first term in square brackets  $L_i$  stands for a doublet of left-handed leptons. This term generates Dirac masses for the neutrino,  $m_D = \lambda \langle H \rangle$ . In general, right handed neutrino may have Majorana masses,  $M$ , as well. Such term is forbidden for other fermions since their right-handed components have charges, but right handed neutrino are neutral.

Right-handed components are also called *sterile neutrino* since they do not interact directly with particles of the Standard Model. However, they are not really sterile since interact with other particles via mixing. Indeed,

to get neutrino mass eigenstates we have to diagonalise mass matrix in square brackets of Lagrangian (68). This gives mixing of active and sterile neutrino

$$\theta = \frac{m_D}{M}. \quad (69)$$

Therefore, sterile neutrino interaction matrix elements are the same as for the active neutrino except they are multiplied by  $\theta$ . If  $m_D \ll M$  the masses of heavy states nearly coincide with  $M$  and the lightest among sterile neutrinos is a good candidate for dark matter if its mass  $M \gtrsim 1$  keV. (But not heavier than 50 keV, otherwise its decays to  $\gamma$  will contradict observed X-ray astrophysical backgrounds, see Section 6.4.)

Sterile neutrino can be produced in the early Universe directly in the inflation decays [60], or via mixing, Eq. (69), with active neutrino [?]. Production rate of sterile neutrinos in the latter case can be obtained multiplying production rate for the active neutrinos in primordial plasma by mixing angle squared

$$\Gamma \approx \theta^2 \sigma_W n \sim \theta^2 G_F^2 T^2 \cdot T^3.$$

Multiplying this rate by time,  $t \sim H^{-1} \sim M_{\text{Pl}}/T^2$  we obtain number density of sterile neutrinos produced

$$\frac{n_s}{n_\gamma} \sim \theta^2 G_F^2 T^3 M_{\text{Pl}}.$$

To close this estimate we note that active-sterile neutrino mixing is temperature dependent [61]

$$\theta \rightarrow \theta_M = \frac{\theta}{1 + 2.4(T/200 \text{ MeV})^6 (\text{keV}/M_1)^2},$$

which gives for the production temperature of sterile neutrino

$$T \sim 130 \left( \frac{M}{1 \text{ keV}} \right)^{1/3} \text{ MeV},$$

and resulting abundance [?]

$$\Omega_s \sim \Omega_m \frac{\sin^2(2\theta)}{10^{-7}} \left( \frac{M}{1 \text{ keV}} \right)^2, \quad (70)$$

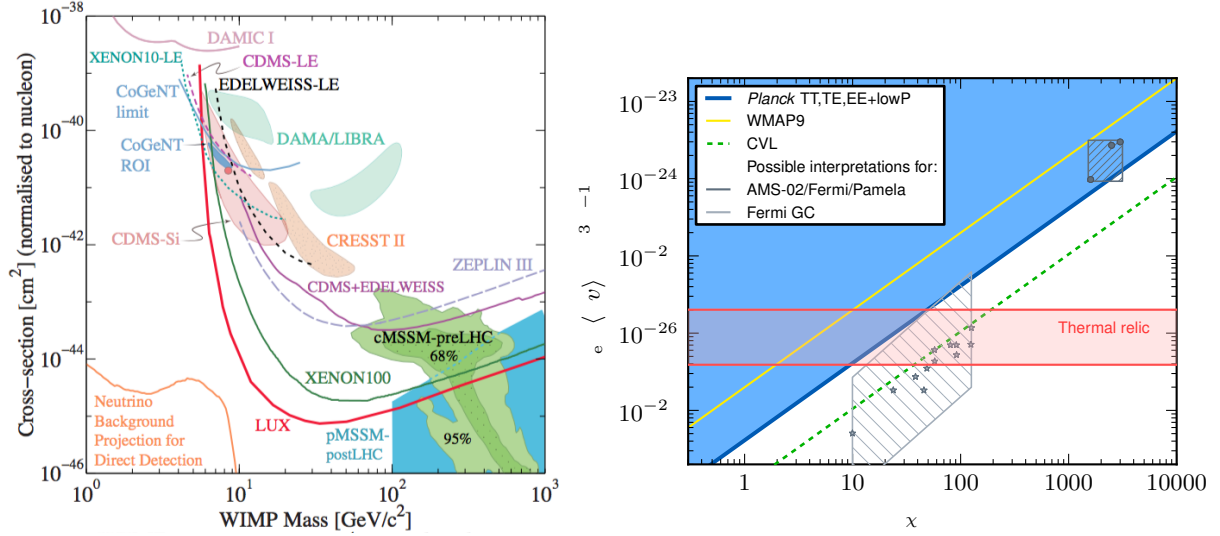
where  $\Omega_m$  is observed dark matter abundance. Proper calculation involves solution of Boltzmann equations. Details and the list of references can be found in the recent review [62]. Quoted result, Eq. (70), corresponds to zero lepton asymmetry. With maximum asymmetry the required  $\theta$  can be two orders of magnitude smaller [63] at the same mass of sterile neutrino, see Fig. 10, right panel.

## 6.4 Dark matter searches and constraints

Dark matter particles, in majority of suggested models, can be discovered in direct dedicated searches in laboratories. Dark matter can also leave trace and be identified in indirect searches, e.g. analysing data on cosmic ray, X-ray, gamma-ray and neutrino telescopes.

1. *Neutralino.* WIMP particles have tiny but phenomenologically important elastic cross-section with usual baryonic matter. For WIMPs heavier than nuclei,  $m_X \gg m_N$  and a typical velocities in the Galaxy halo,  $v \sim 300$  km/s, typical recoil energy is  $E_R \sim m_N v^2 \sim 1 - 100$  keV. The recoil can be measured studying ionization, scintillation, heat or sound waves it creates in a detector. Different experiments use different techniques, or their combinations. But, of course, it should be done deep underground, in low background laboratories. Current observational bounds on the scattering cross-section exclude a lot of the WIMP parameter space of MSSM but do not test the bulk of it, see Fig. 9, left panel. Intriguingly, crystal-based experiments CDMS Si, CoGeNT ROI, CRESST II and DAMA/LIBRA claim some hints of a positive dark matter signal. These claims are mutually exclusive, and cluster in the mass region of tens of GeV and at cross sections between  $10^{42}$  and  $10^{39}$  cm<sup>2</sup>, see Fig. 9, left panel. However, the noble-gas experiments ZEPLIN, XENON and most recent LUX, exclude this parameter region. Remaining expectation for supersymmetric models, after all constraint are taken into account, including LHC results, are shown by shaded area in the lower right corner, marked as MSSM, in the same figure. The uncertainty for the expected signal arises because the scattering cross-section is not directly related to the annihilation cross-section.

However, that same self-annihilation that plays a central role in the freeze-out, see Section 6.3.1, leads also to the dark matter annihilation in the Galaxy halo. It can give rise to a significant flux of  $\gamma$ -rays, neutrinos,



**Fig. 9:** Left panel: Constraints on neutralino from different underground experiments are shown by correspondingly marked colored lines. Similarly, claimed hints of detection are represented by shaded areas. Remaining expectation for supersymmetric models are shown by shaded area in the lower right corner, marked as MSSM. Right panel: Constraints on neutralino self-annihilation cross-section from CMBR. The blue area shows the parameter space excluded by the Planck data. The yellow line indicates the constraint using WMAP9 data. The dark grey circles and light grey stars correspond to various claims of indirect detection as cosmic or  $\gamma$ -ray excesses. The horizontal red band corresponds to correct neutrino abundance, Eq. (64). From Ref. [2].

and even antimatter such as antiprotons and positrons, especially from regions with large dark-matter density. This creates prospective signal for the indirect WIMP detection. It is searched for, as an excess over conventional astrophysical backgrounds, by the orbital cosmic ray observatories, ground based atmospheric Cherenkov and neutrino telescopes. Though annihilation cross-section for indirect searches is fixed, some uncertainty arises here because of certain uncertainty in dark matter density profiles.

Dark-matter annihilation with a non-vanishing branching ratio into the electromagnetic channel leads also to distortions of the CMB which has been probed with WMAP and Planck data [2]. WIMPs lighter than 10 GeV originating in thermal freeze-out scenario are excluded by these observations and the advantage of CMB-based limits lies in the absence of astrophysical uncertainties, see Fig. 9, right panel. Dark matter annihilation interpretation of the cosmic ray excess detected by AMS, Fermi and Pamela satellites (shown by dark grey circles) are also excluded now by Planck. However, the interpretation of the  $\gamma$ -ray excess from the Galactic centre measured by Fermi (corresponding parameter regions are indicated by light grey stars) is still viable. Intriguingly, it intersects with the horizontal red band which corresponds to the correct neutralino abundance as thermal relic.

Dark matter particles escape direct detection at colliders such as the Large Hadron Collider (LHC) at CERN, however, they would produce a characteristic signal of missing energy. Arising constraints on WIMP-nucleon cross-section are model dependent, but are powerful and competitive with direct searches in underground labs, especially in the region of low masses. Recent detailed review on direct, indirect and collider WIMP searches can be found e.g. in Ref. [65].

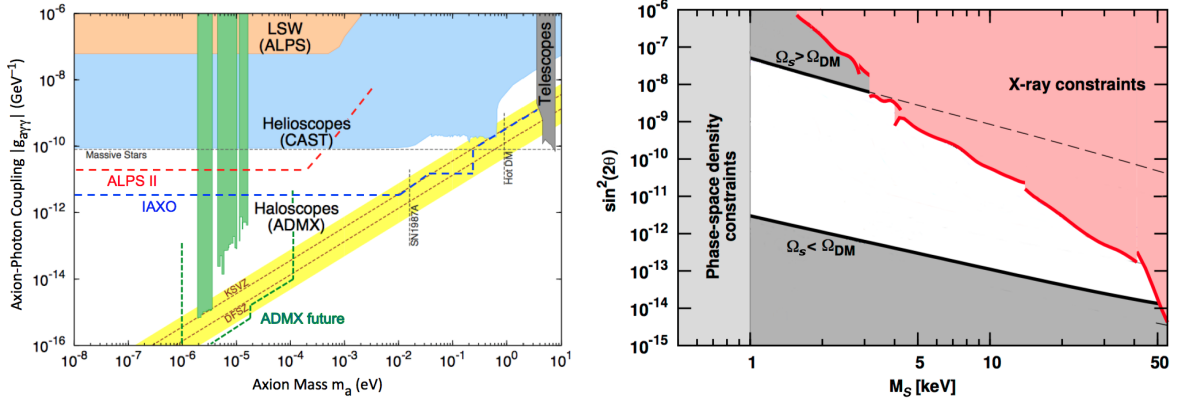
2. *Axions.* Axion interactions with photons and fermions can be parametrised as

$$L_{\text{int}} = -\frac{1}{4}g_{a\gamma} a F_{\mu\nu} \tilde{F}^{\mu\nu} - \sum_{\text{fermions}} g_{ai} a \bar{\psi}_i \gamma_5 \psi_i, \quad (71)$$

where

$$g_{a\gamma} \equiv \frac{\alpha}{2\pi} \frac{C_{a\gamma}}{f_a}, \quad g_{ai} \equiv \frac{m_i}{f_a} C_{ai},$$

and  $C_{a\gamma}$ ,  $C_{ai}$  are model dependent parameters (in simple models of order unity). Direct axion searches in the laboratory are based on interactions with  $\gamma$ . Namely, axions constituting the Milky Way dark matter halo would



**Fig. 10:** Left panel: Parameter space for the axion and axion-like dark matter models. Yellow band corresponds to the correct cosmological abundance of QCD axions if they make all of dark matter. Regions excluded by ADMX and CAST are shown by green and blue shaded areas. Dashed lines show prospective limits of planned experiments. From Ref. [68]. Right panel: Unshaded white region represents allowed parameter space for DM sterile neutrinos. The upper and lower thick black lines correspond to correct abundances for zero and maximal lepton asymmetry. Red region in the upper right corner is forbidden by the X-ray constraints. The region below 1 keV is ruled out by the phase-space density arguments. Adapted from Ref. [64].

resonantly convert into a monochromatic microwave signal in a high-Q microwave cavity permeated by a strong magnetic field [66]. Such axion search experiments (ADMX is the most recent one) are called *haloscopes*. Similarly, axions or axion-like particles, emitted by the Sun will be converted in the strong magnetic field into X-ray photons. Axion experiments which search for this signal (CAST is the most recent one) are called *helioscopes*. Constraints obtained by ADMX and CAST experiments are shown in Fig. 10, left panel.

Interaction Lagrangian Eq. (71) leads also to a new observable astrophysical phenomena, which may lead to indirect axion detection and give constraints on axion parameters. Extra energy losses by stars is one of those effects. Corresponding constraints practically coincide with the bound obtained by CAST. It is shown by the dotted line marked by the label "Massive Stars" in the Fig. 10, left panel. No accident, along the same line we can find models capable to explain several claims hinting for the axion effects in the astrophysical data, for a review see e.g. Ref. [67]. Recent detailed review on direct and indirect axion searches can be found also in Refs. [65, 68].

3. *Sterile neutrino*. Recent detailed review on direct and indirect sterile neutrino searches can be found in Ref. [62]. In every process where active neutrino appears, sterile neutrino can appear as well, again via mixing, Eq. (69). This opens the way for a laboratory sterile neutrino searches. For example, in the keV mass range, appearance of sterile neutrinos changes kinematics and the spectrum of nuclear decays. Most recent searches of sterile neutrinos in tritium  $\beta$ -decay has started in Troitsk [69], and will be continued at KATRIN experiment [62].

Also, at one loop level sterile neutrino are decaying into active neutrino and photon. Loop diagrams for this process are the same as for the electromagnetic form factors of a massive neutrino in the Standard Model with one external neutrino leg being connected to sterile neutrino via mixing (69). The decay width can be easily obtained using e.g. results of Ref. [70] and is given by

$$\Gamma_{N \rightarrow \gamma \nu_a} = \frac{9 \alpha G_F^2}{256 \cdot 4\pi^4} \sin^2 2\theta m_s^5 = 5.5 \times 10^{-22} \theta^2 \left[ \frac{m_s}{1 \text{ keV}} \right]^5 \text{ s}^{-1}. \quad (72)$$

Because of that, sterile neutrino dark matter is not completely dark. It can be detected by searching for an unidentified X-ray line, which would appear at a frequency  $\omega = m_s/2$ . Intensity of this line should follow dark matter density profiles. Dwarf satellite galaxies are a good places to search for such a signal because they are dark matter dominated and usual astrophysical X-ray background is small there [71].

To conclude this section. A large number of various claims exists in the literature with a hints of indirect dark matter signal for all of the candidates described above: neutralino, axion-like particles and sterile neutrino. Do describe these hints in detail would require separate volume, interested reader can consult recent reviews [62, 65, 67]. As usual, hints appear at a boundary of allowed parameter space where observational capabilities are



stretched. Moreover, indirect dark matter signal can be confused with conventional astrophysical backgrounds or effects. Clearly, these claims are mutually exclusive and it is not possible for all of them to be precursors of the true signal, since dark matter is either neutrino, or axions, or sterile neutrino, or something else. On the other hand, one of those may turn out to be true and it is not excluded that we see already the tip of a real iceberg.

## 7 Conclusions

As we have seen, cosmology and astrophysics gave us solid evidence that the Standard Model of particle physics is incomplete. We have to extend it to explain neutrino masses, baryogenesis, and dark matter. Dark energy can be explained by the Einstein's  $\Lambda$ -term, but we do not know why it exists, and there seems to be too many coincidences between numerical values of cosmological parameters. On the other hand, a form of dark energy explains the Universe origin within inflationary paradigm, which increasingly finds support in cosmological data.

Cosmology just recently became a precision science but is full of surprises already, helping to build true model of microphysics. It is up to high energy physicists to find out what this new physics is. With advances of this program we, in turn, will have better understanding of the Universe origin, of its evolution, of its current state, and of its future fate.

## Acknowledgements

I would like to use the opportunity to thank CERN and Dubna for the organization and support of this annual summer school series for the young high-energy physicists.

## References

- [1] D. S. Gorbunov and V. A. Rubakov, Introduction to the Theory of the Early Universe, World Scientific, 2011.
- [2] P. A. R. Ade *et al.* [Planck Collaboration], Astron. Astrophys. **594** (2016) A13, arXiv:1502.01589.
- [3] A. G. Riess *et al.*, Astrophys. J. **826** (2016) 56, arXiv:1604.01424.
- [4] Z. Berezhiani, A. D. Dolgov and I. I. Tkachev, Phys. Rev. D **92** (2015) 061303, arXiv:1505.03644.
- [5] A. A. Penzias and R. W. Wilson, Astrophys. J. **142** (1965) 419.
- [6] R. H. Dicke, P. J. E. Peebles, P. G. Roll, and D. T. Wilkinson, Astrophys. J. **142** (1965) 414.
- [7] T. Shmaonov, Priboi Tekhnika Eksperimenta **1** (1957) 83.
- [8] A. McKellar, Proc. Ast. Soc. Pac. **52** (1940) 187.
- [9] A. G. Doroshkevich and I. D. Novikov, Sov. Phys. Dokl. **9** (1964) 111.
- [10] D.J. Fixsen, Astrophys. J. **707** (2009) 916, arXiv:0911.1955.
- [11] A. D. Dolgov, Phys. Rept. **370** (2002) 333, hep-ph/0202122.
- [12] A. D. Dolgov and M. Fukugita, Phys. Rev. D **46** (1992) 5378.
- [13] G. F. Smoot *et al.*, Astrophys. J. **396** (1992) L1.
- [14] R. K. Sachs and A. M. Wolfe, Astrophys. J. **147** (1967) 73.
- [15] J. Silk, Astrophys. J. **151** (1968) 459.
- [16] P. J. E. Peebles and J. T. Yu, Astrophys. J. **162** (1970) 815.
- [17] R. A. Syunyaev and Y. B. Zel'Dovich, Astrophysics and Space Science **7** (1970) 3.
- [18] A. G. Doroshkevich, Y. B. Zel'Dovich, and R. A. Syunyaev, Soviet Astronomy **22** (1978) 523 .
- [19] M. Betoule *et al.* [SDSS Collaboration], Astron. Astrophys. **568** (2014) A22, arXiv:1401.4064.
- [20] M. Pellejero-Ibanez *et al.* [BOSS Collaboration], arXiv:1607.03152.
- [21] A. A. Starobinsky, JETP Lett. **30** (1979) 682.
- [22] A. A. Starobinsky, Phys. Lett. **B91** (1980) 99.
- [23] A. H. Guth, Phys. Rev. **D23** (1981) 347.
- [24] A. D. Linde, Phys. Lett. **B108** (1982) 389.
- [25] A. Albrecht and P. J. Steinhardt, Phys. Rev. Lett. **48** (1982) 1220.
- [26] A. D. Linde, Phys. Lett. **B129** (1983) 177.
- [27] A. H. Guth and E. J. Weinberg, Phys. Rev. **D23** (1981) 876.
- [28] L. P. Grishchuk, Sov. Phys. JETP **40** (1975) 409.

- [29] V. N. Lukash, Sov. Phys. JETP **52** (1980) 807.
- [30] M. Sasaki, Prog. Theor. Phys. **76** (1986) 1036.
- [31] V. F. Mukhanov, Sov. Phys. JETP **67** (1988) 1297.
- [32] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger, Phys. Rept. **215** (1992) 203.
- [33] V. A. Rubakov, M. V. Sazhin, and A. V. Veryaskin, Phys. Lett. **B115** (1982) 189.
- [34] P. A. R. Ade *et al.* [Planck Collaboration], arXiv:1502.02114.
- [35] E. R. Harrison, Phys. Rev. **D1** (1970) 2726.
- [36] Y. B. Zeldovich, Mon. Not. Roy. Astron. Soc. **160** (1972) 1.
- [37] M. Milgrom, Astrophys. J. **270** (1983) 365.
- [38] E. P. Verlinde, arXiv:1611.02269.
- [39] J. F. Navarro, C. S. Frenk, and S. D. M. White, Astrophys. J. **490** (1997) 493.
- [40] P. Sikivie, I. I. Tkachev, and Y. Wang, Phys. Rev. **D56** (1997) 1863.
- [41] G. Steigman and I. Tkachev, Astrophys. J. **522** (1999) 793, astro-ph/9803008.
- [42] K. Dolag, A. D. Dolgov and I. I. Tkachev, JETP Lett. **96** (2013) 754, arXiv:1210.8009.
- [43] F. Zwicky, Helv. Phys. Acta **6** (1933) 110.
- [44] A. A. Vikhlinin *et al.*, Phys. Usp. **57** (2014) 317.
- [45] A. A. Vikhlinin *et al.*, Astrophys. J. **640** (2006) 691, astro-ph/0507092.
- [46] J. P. Kneib *et al.*, Astrophys. J. **471** (1996) 643, astro-ph/9511015.
- [47] M. Bartelmann and P. Schneider, Physics Reports **340** (2001) 291, astro-ph/9912508.
- [48] G. Squires *et al.*, Astrophys. J. **461** (1996) 572, astro-ph/9507008.
- [49] N. Okabe *et al.*, Astrophys. J. **769** (2013) L35, arXiv:1302.2728.
- [50] D. Clowe *et al.*, Astrophys. J. **648** (2006) L109, astro-ph/0608407.
- [51] S. L. Dubovsky, P. G. Tinyakov and I. I. Tkachev, Phys. Rev. Lett. **94** (2005) 181102, [hep-th/0411158; M. Pshirkov, A. Tuntsov, K. A. Postnov, Phys. Rev. Lett. **101** (2008) 261101, arXiv:0805.1519; S. Dubovsky, R. Flauger, A. Starobinsky and I. Tkachev, Phys. Rev. D **81** (2010) 023523, arXiv:0907.1658.
- [52] R. D. Peccei, H. R. Quinn, Phys. Rev. Lett. **38** (1977) 1440; S. Weinberg, Phys. Rev. Lett. **40** (1978) 223; F. Wilczek, Phys. Rev. Lett. **40** (1978) 279; M. Dine, W. Fischler, M. Srednicki, Phys. Lett. **B104** (1981) 199; A. R. Zhitnitsky, Sov. J. Nucl. Phys. **31** (1980) 260; J. E. Kim, Phys. Rev. Lett. **43** (1979) 103; M. A. Shifman, A. I. Vainshtein, V. I. Zakharov, Nucl. Phys. **B166** (1980) 493;
- [53] S. Dodelson and L. M. Widrow, Phys. Rev. Lett. **72** (1994) 17, hep-ph/9303287; X. D. Shi and G. M. Fuller, Phys. Rev. Lett. **82** (1999) 2832, astro-ph/9810076; T. Asaka, S. Blanchet and M. Shaposhnikov, Phys. Lett. B **631** (2005) 151, hep-ph/0503065;
- [54] S. I. Blinnikov and M. Y. Khlopov, Sov. J. Nucl. Phys. **36** (1982) 472; Z. Berezhiani, D. Comelli, and F. L. Villante, Phys. Lett. **B503** (2001) 362; R. Foot and R. R. Volkas, Phys. Rev. **D68** (2003) 021304.
- [55] J. R. Ellis *et al.*, Nucl. Phys. **B238** (1984) 453; G. Jungman, M. Kamionkowski, K. Griest, Phys. Rept. **267** (1996) 195, hep-ph/9506380.
- [56] D. J. H. Chung, E. W. Kolb, and A. Riotto, Phys. Rev. **D59** (1999) 023501, hep-ph/9802238;
- [57] V. Kuzmin and I. Tkachev, JETP Lett. **68** (1998) 271, hep-ph/9802304.
- [58] S. Tremaine and J. E. Gunn, Phys. Rev. Lett. **42** (1979) 407.
- [59] E. W. Kolb and I. I. Tkachev, Phys. Rev. Lett. **71** (1993) 3051, hep-ph/9303313.
- [60] M. Shaposhnikov and I. Tkachev, Phys. Lett. B **639** (2006) 414, hep-ph/0604236.
- [61] A. D. Dolgov and S. H. Hansen, Astropart. Phys. **16** (2002) 339, hep-ph/0009083.
- [62] M. Drewes *et al.*, arXiv:1602.04816.
- [63] M. Laine and M. Shaposhnikov, JCAP **0806** (2008) 031, arXiv:0804.4543.
- [64] A. Boyarsky, O. Ruchayskiy and M. Shaposhnikov, Ann. Rev. Nucl. Part. Sci. **59** (2009) 191, arXiv:0901.0011.
- [65] M. Klasen, M. Pohl and G. Sigl, Prog. Part. Nucl. Phys. **85** (2015) 1, arXiv:1507.03800.
- [66] P. Sikivie, Phys. Rev. Lett. **51** (1983) 1415 Erratum: [Phys. Rev. Lett. **52** (1984) 695].
- [67] A. Ringwald, PoS NEUTEL **2015** (2015) 021, arXiv:1506.04259.
- [68] P. W. Graham *et al.*, Ann. Rev. Nucl. Part. Sci. **65** (2015) 485, arXiv:1602.00039.

- [69] D. N. Abdurashitov *et al.*, JINST **10** (2015) T10005, arXiv:1504.00544.  
 [70] P. B. Pal and L. Wolfenstein, Phys. Rev. D **25** (1982) 766.  
 [71] A. Boyarsky *et al.*, Phys. Rev. Lett. **97** (2006) 261302, astro-ph/0603660.

## Appendices

### A Gravitational creation of metric perturbations

As an important and simple example, let us consider quantum fluctuations of a real scalar field, which we denote as  $\varphi$ . It is appropriate to rescale the field values by the scale factor,  $\varphi \equiv \phi/a(\eta)$ . This brings the equations of motion for the field  $\phi$  into a simple form of Eq. (54). As usual, we decompose  $\phi$  over creation and annihilation operators  $b_{\mathbf{k}}$  and  $b_{\mathbf{k}}^\dagger$

$$\phi(\mathbf{x}, \eta) = \int \frac{d^3k}{(2\pi)^{3/2}} \left[ u_k(\eta) b_{\mathbf{k}} e^{i\mathbf{k}\mathbf{x}} + u_k^*(\eta) b_{\mathbf{k}}^\dagger e^{-i\mathbf{k}\mathbf{x}} \right]. \quad (\text{A.1})$$

Mode functions  $u_k$  satisfy Eq. (54). In what follows we will assume that  $\varphi$  is the inflaton field of the ‘‘chaotic’’ inflationary model, Eq. (52). During inflation  $H \gg m$  and  $H \approx \text{const}$ . So, to start with, we can assume that  $\varphi$  is a massless field on the constant deSitter background. (The massive case can be treated similarly, but analytical expressions are somewhat more complicated and do not change the result in a significant way. Corrections due to change of  $H$  can also be taken into account, and we do that later for the purpose of comparison with observations.) With a constant Hubble parameter during inflation the solution of Friedmann equations in conformal time is

$$a(\eta) = -\frac{1}{H\eta} \quad (\text{A.2})$$

and the equation for mode functions of a massless, conformally coupled to gravity ( $\xi = 0$ ), scalar field takes the form

$$\ddot{u}_k + k^2 u_k - \frac{2}{\eta^2} u_k = 0. \quad (\text{A.3})$$

Solutions which start as vacuum fluctuations in the past ( $\eta \rightarrow -\infty$ ) are given by

$$u_k = \frac{e^{\pm ik\eta}}{\sqrt{2k}} \left( 1 \pm \frac{i}{k\eta} \right). \quad (\text{A.4})$$

Indeed, at  $\eta \rightarrow -\infty$  the second term in the parentheses can be neglected and we have the familiar mode functions of the Minkowski space time. The wavelength of a given mode becomes equal to the horizon size (or ‘‘crosses’’ the horizon) when  $k\eta = 1$ . Inflation proceeds with  $\eta \rightarrow 0$ , so the modes with progressively larger  $k$  cross the horizon. After horizon crossing, when  $k\eta \ll 1$ , the asymptotics of mode functions are

$$u_k = \pm \frac{i}{\sqrt{2k^{3/2}\eta}}, \quad \text{or} \quad \varphi_k = \frac{u_k}{a(\eta)} = \mp \frac{iH}{\sqrt{2k^{3/2}}}. \quad (\text{A.5})$$

The field variance is given by

$$\langle 0 | \phi^2(x) | 0 \rangle = \int \frac{d^3k}{(2\pi)^3} |\varphi_k|^2. \quad (\text{A.6})$$

and we find in the asymptotic (the careful reader will recognize that this is already regularized expression with zero-point fluctuations being subtracted)

$$\langle \varphi^2 \rangle = \frac{H^2}{(2\pi)^2} \int \frac{dk}{k}. \quad (\text{A.7})$$

Defining the power spectrum of the field fluctuations as a power per decade,  $\langle \varphi^2 \rangle \equiv \int P_\varphi(k) d \ln k$ , we find

$$P_\varphi(k) = \frac{H^2}{(2\pi)^2}. \quad (\text{A.8})$$

### A.1 Curvature perturbations

According to Eq. (4), the three-dimensional curvature of space sections of constant time is inversely proportional to the scale factor squared,  ${}^{(3)}R \propto a^{-2}$ . Therefore, the perturbation of spatial curvature is proportional to  $\delta a/a$ , and this ratio can be evaluated as

$$\zeta \equiv \frac{\delta a}{a} = H \delta t = H \frac{\delta \varphi}{\dot{\varphi}}. \quad (\text{A.9})$$

This allows to relate the power spectrum of curvature perturbations to the power spectrum of field fluctuations

$$P_\zeta(k) = \frac{H^2}{\dot{\varphi}^2} P_\varphi(k), \quad (\text{A.10})$$

and we find for the power spectrum of curvature perturbations

$$P_\zeta(k) = \frac{1}{4\pi^2} \frac{H^4}{\dot{\varphi}^2}. \quad (\text{A.11})$$

This very important relation describes inflationary creation of primordial perturbations, and can be confronted with observations. The usefulness of curvature perturbations for this procedure can be appreciated in the following way:

1. Consider the perturbed metric, Eq. (38). The product  $a(1 - \Phi)$  for the long-wavelength perturbations can be viewed as a perturbed scale factor, i.e.  $\delta a/a = -\Phi$ . Comparing this relation with Eq. (A.9) and Eq. (42), we find for the temperature fluctuations which are of the superhorizon size at the surface of last scattering

$$\frac{\delta T}{T} = \frac{2}{3} \zeta_k. \quad (\text{A.12})$$

2. On superhorizon scales the curvature perturbations do not evolve usually. This fact allows to relate directly the observed power spectrum of temperature fluctuations to the power spectrum of curvature fluctuations generated during inflation.

### A.2 Tensor perturbations

Mode functions of gravity waves (after rescaling by  $M_{\text{Pl}}/\sqrt{32\pi}$ ) obey the same equation as mode functions of massless minimally coupled scalar [28]. Using the result Eq. (A.8) we immediately find [33]

$$P_T(k) = 2 \frac{32\pi}{M_{\text{Pl}}^2} P_\varphi(k) = \frac{16}{\pi} \frac{H^2}{M_{\text{Pl}}^2}, \quad (\text{A.13})$$

where the factor of 2 accounts for two graviton polarizations.

### A.3 Slow-roll approximation

During inflation, the field  $\varphi$  rolls down the potential hill very slowly. A reasonable approximation to the dynamics is obtained by neglecting  $\ddot{\varphi}$  in the field equation  $\ddot{\varphi} + 3H\dot{\varphi} + V' = 0$ . This procedure is called the slow-roll approximation

$$\dot{\varphi} \approx -\frac{V'}{3H}. \quad (\text{A.14})$$

Field derivatives can also be neglected in the energy density of the inflaton field,  $\rho \approx V$

$$H^2 = \frac{8\pi}{3M_{\text{Pl}}^2} V. \quad (\text{A.15})$$

This gives for curvature perturbations

$$\zeta_k \equiv P_\zeta(k)^{1/2} = \frac{H^2}{2\pi \dot{\varphi}} = \frac{4H}{M_{\text{Pl}}^2} \frac{V}{V'}. \quad (\text{A.16})$$

#### A.4 Normalizing to CMBR

As an example, let us consider the simplest model  $V = \frac{1}{2}m^2\varphi^2$ . We have

$$\frac{V}{V'} = \frac{\varphi}{2}, \quad \text{and} \quad H = \sqrt{\frac{4\pi}{3}} \frac{m\varphi}{M_{\text{Pl}}}. \quad (\text{A.17})$$

This gives for the curvature fluctuations

$$\zeta_k = \sqrt{\frac{16\pi}{3}} \frac{m\varphi^2}{M_{\text{Pl}}^3}. \quad (\text{A.18})$$

Using the relation between curvature and temperature fluctuations, Eq. (A.12), and normalizing  $\delta T/T$  to the measured value at largest  $l$ , which is  $\delta T/T \sim 10^{-5}$  (see Fig. 3, right panel) we find the restriction on the value of the inflaton mass in this model:

$$m \approx \frac{\delta T}{T} \frac{M_{\text{Pl}}}{30} \approx 10^{13} \text{ GeV}. \quad (\text{A.19})$$

Here I have used the fact that in this model the observable scales cross the horizon when  $\varphi \approx M_{\text{Pl}}$ .

#### A.5 Slow-roll parameters

The number of e-foldings ( $a = e^{Ht} \equiv e^N$ ) of inflationary expansion from the time when  $\varphi = \varphi_i$  to the end can be found as

$$N(\varphi_i) = \int_{t_i}^{t_f} H(t) dt = \int \frac{H}{\dot{\varphi}} d\varphi = \frac{8\pi}{M_{\text{Pl}}^2} \int_{\varphi_e}^{\varphi_i} \frac{V}{V'} d\varphi. \quad (\text{A.20})$$

In particular, in the model Eq. (52) we find that the largest observable scale had crossed the horizon ( $N \sim 65$ ) when  $\varphi_i \approx 3.5M_{\text{Pl}}$ . All cosmological scales which fit within the observable universe encompass a small  $\Delta\phi$  interval within  $M_{\text{Pl}} < \varphi < \varphi_i$ . And inflaton potential should be sufficiently flat over this range of  $\Delta\phi$  for the inflation to proceed. This means that observables essentially depend on the first few derivatives of  $V$  (in addition the the potential  $V(\phi_0)$  itself). From the first two derivatives one can construct the following dimensionless combinations

$$\epsilon \equiv \frac{M_{\text{Pl}}^2}{16\pi} \left( \frac{V'}{V} \right)^2, \quad (\text{A.21})$$

$$\eta \equiv \frac{M_{\text{Pl}}^2}{8\pi} \frac{V''}{V}, \quad (\text{A.22})$$

which are often called the slow-roll parameters.

The power spectra of curvature, Eq. (A.10), and of tensor perturbations, Eq. (A.13), in slow-roll parameters can be rewritten as

$$P_\zeta(k) = \frac{1}{\pi\epsilon} \frac{H^2}{M_{\text{Pl}}^2}, \quad P_T(k) = \frac{16}{\pi} \frac{H^2}{M_{\text{Pl}}^2}. \quad (\text{A.23})$$

Comparing these two expressions we find

$$\frac{P_T(k)}{P_\zeta(k)} = 16\epsilon. \quad (\text{A.24})$$

#### A.6 Primordial spectrum

In general, the spectra can be approximated as power law functions in  $k$ :

$$P_\zeta(k) = P_\zeta(k_0) \left( \frac{k}{k_0} \right)^{n_S-1}, \quad (\text{A.25})$$

$$P_T(k) = P_T(k_0) \left( \frac{k}{k_0} \right)^{n_T}. \quad (\text{A.26})$$

To the first approximation,  $H$  in Eq. (A.23) is constant. Therefore, in this approximation, power spectra do not depend on  $k$  and  $n_S = 1$ ,  $n_T = 0$ . This case is called the Harrison-Zel'dovich spectrum [35, 36] of primordial perturbations. However, in reality,  $H$  is changing, and in Eq (A.23) for every  $k$  one should take the value of  $H$  at the moment when the relevant mode crosses horizon. In slow roll parameters one then finds (see e.g. Ref. [?] for the nice overview)

$$n_S = 1 + 2\eta - 6\epsilon, \quad n_T = -2\epsilon. \quad (\text{A.27})$$

We can re-write Eq. (A.28) as a relation between the slope of tensor perturbations and the ratio of power in tensor to curvature modes

$$\frac{P_T(k)}{P_\zeta(k)} = -8n_T . \quad (\text{A.28})$$

This is called the *consistency relation* to which (simple) inflationary models should obey.

Different models of inflation have different values of slow-roll parameters  $\eta$  and  $\epsilon$ , and therefore can be represented in the  $(\eta, \epsilon)$  parameter plane. Using the relations Eq. (A.27) we see that this plane can be mapped into  $(n_S, n_T)$ , or using also Eq. (A.28) into the  $(n_S, r)$  parameter plane, where  $r$  is the ratio of power in tensor to scalar (curvature) perturbations. In this way, different inflationary models can be linked to observations and constraints can be obtained.

## Organizing Committee

T. Donskova (Schools Administrator, JINR)  
N. Ellis (CERN)  
M. Mulders (CERN)  
A. Olchevsky (JINR)  
K. Ross (Schools Administrator, CERN)  
G. Zanderighi (CERN and Univ. Oxford, UK)

## Local Organizing Committee

T. Buanes (Univ. Bergen, Norway)  
G. Eigen (Univ. Bergen, Norway)  
E. Gramstad (Univ. Oslo, Norway)  
T. Gonzalo (Univ. Oslo, Norway)  
A. Krislock (Univ. Oslo, Norway)  
A. Lipniacka (Univ. Bergen, Norway)  
B. Martin dit Latour (Univ. Bergen, Norway)  
F. Ould-Saada (Univ. Oslo, Norway)  
A. Raklev (Univ. Oslo, Norway)  
A. Read (Univ. Oslo, Norway)  
H. Sandaker (Univ. Oslo, Norway)

## International Advisors

F. Gianotti (CERN)  
R. Heuer (CERN)  
V. Matveev (JINR)  
A. Skrinsky (BINP, Novosibirsk, Russia)  
N. Tyurin (IHEP, Protvino, Russia)

## Lecturers

B. Allanach (Univ. Cambridge, UK)  
G. Barenboim (Univ. Valencia, Spain)  
M. Blanke (KIT, Germany)  
A. Hoecker (CERN)  
A. Kurkela (Univ. Stavanger, Norway and CERN)  
L. Lista (INFN-Napoli, Italy)  
M. Peskin (SLAC, USA)  
I. Tkachev (INR, Russia)  
G. Zanderighi (CERN and Univ. Oxford, UK)

## Discussion Leaders

R. Franceschini (CERN)  
G. Luisoni (CERN)  
R. Sadykov (JINR)  
A. Saponov (JINR)  
M. Schulze (CERN)  
A. Urbano (CERN)

# Students

|                             |                        |                      |
|-----------------------------|------------------------|----------------------|
| Baptiste ABELOOS            | Julian GARCIA PARDINAS | Juska PEKKANEN       |
| Nicola ABRAHAM              | Ksenia GASNIKOVA       | Aleksandr PETROV     |
| John ANDERS                 | Felipe GONZALEZ        | Michael PITT         |
| Aliaksandr ANTOSHKIN        | Daniel GONZALEZ        | Giorgia RAUCO        |
| Tatiana ANTOSHKINA          | Moritz HABERMEHL       | Marilea REALE        |
| Nedaa ASBAH                 | Marco HARRENDORF       | Chiara RIZZI         |
| Alessandra BAAS             | Hadi HASSAN            | Sinan SAGIR          |
| Marton BARTOK               | David HOHN             | Dirk SAMMEL          |
| Ayse BAT                    | Tao HUANG              | Myriam SCHONENBERGER |
| Nicoletta BELLOLI           | Mariya ILIEVA          | Igor SHANDROV        |
| Pirmin BERGER               | Viacheslav KAMINSKIY   | Ksenia SHCHELINA     |
| Danijela BOGAVAC            | Callum KILBY           | Mariana SHOPOVA      |
| Alessio BOLETTI             | Viktar KIREYEU         | Rafal SIKORA         |
| Aurelie BONHOMME            | Suzanne KLAVER         | Miroslav SIMKO       |
| Svende BRAUN                | Denis KORABLEV         | Ismet SIRAL          |
| Angela BURGER               | Dominik KRAUSS         | Radim SLOVAK         |
| Noemi CALACE                | Vit KUCERA             | Juraj SMIESKO        |
| Clement CAMINCHER           | Eloi LE QUILLEUC       | Margherita SPALLA    |
| Ryne CARBONE                | Giuseppe LERNER        | Andre STAHL          |
| Francesco CIROTTO           | Arthur LESAGE          | Rizki SYARIF         |
| Simon CORRODI               | Lucien LO              | Thor TAYLOR          |
| Xavier COUBEZ               | Luigi LONGO            | Boris TEYSSIER       |
| Denys DENYSIUK              | Nicolas LURKIN         | Konstantin TRESKOV   |
| Karlis DREIMANIS            | Steffen MAELAND        | Daniel TURGEMAN      |
| Yi-Ting DUH                 | Artem MAEVSKIY         | Meera VIEIRA MACHADO |
| Giulio DUJANY               | Radoslav MARCHEVSKI    | Maria VIEITES DIAZ   |
| Ana Elena DUMITRIU          | Carla MARIN BENITO     | Luke VINTON          |
| Mai EL SAWY                 | Mareike MEYER          | Alina VISHNEVA       |
| Alberto ESCALANTE DEL VALLE | Emanuele MICHIELIN     | Marek WALCZAK        |
| Oleksiy FEDORCHUK           | Andrea MOGINI          | Andy WHARTON         |
| Cora FISCHER                | Jaime NORMAN           | Stephanie YUEN       |
| Lukasz FULEK                | Tatiana OVSIANNIKOVA   | Ruiqi ZHANG          |
| Katarina GAJDOSOVA          | Rhys OWEN              |                      |



## Posters

| Poster title  | Presenter               |
|---|-------------------------|
| Jet Energy Scale Calibration and Uncertainties  | ABELOOS, B.             |
| Searching for Supersymmetry at the Large Hadron Collider  | ABRAHAM, N.             |
| NOvA test bench at JINR   | ANTOSHKIN, A.           |
| Optical Simulation of PMT   | ANTOSHKINA, T.          |
| Search for new physics predicted by gauge mediated SUSY breaking models with a photon and a boosted Higgs boson as final state at CMS | BARTÓK, M.              |
| $B^0 \rightarrow K^* \mu\mu$ angular analysis with the CMS detector   | BOLETTI, A.             |
| $J/\Psi$ production in 13 TeV pp collisions   | BRAUN, S.               |
| Electron identification efficiency measurement in the ATLAS detector using the probe isolation  | BURGER, A.              |
| The Inclined Layout for the ATLAS Tracking Detector Upgrade   | CALACE, N.              |
| New reconstruction of very boosted top-quarks decaying electronically   | CAMINCHER, C.           |
| WW/WZ resonance production in the $lvqq$ final state at 13 TeV with the ATLAS detector at the LHC                                     | CARBONE, R.             |
| The Mu3e Experiment   | CORRODI, S.             |
| $t\bar{t}H \rightarrow$ multilepton in the CMS experiment for Run II  | COUBEZ, X.              |
| Real-time alignment and calibration of the LHCb Detector in Run II  | DUJANY, G.              |
| Searching for Magnetic Monopoles with the CMS experiment at the LHC   | EL SAWY, M.             |
| Search for $W' \rightarrow l\nu$ production at 13 TeV in CMS  | ESCALANTE DEL VALLE, A. |
| Investigations of the long-term stability of a GEM-TPC  | FEDORCHUK, O.           |
| Search for new phenomena in monojet plus missing transverse momentum final states in pp collisions at 13 TeV using the ATLAS detector | FISCHER, C.             |
| Measurement of the CP-violating phase $\phi_s$ in $B_s^0 \rightarrow (K^+\pi^-)(K^-\pi^+)$ at LHCb                                    | GARCÍA PARDIÑAS, J.     |

| <b>Poster title</b>  | <b>Presenter</b>  |
|--|-------------------|
| WIMP Search at the International Linear Collider   | HABERMEHL, M.     |
| Studies of uncertainties associated with Monte Carlo event generation  | HARRENDORF, M. A. |
| Jet Quenching Measurement with ALICE at the LHC  | HASSAN, H.        |
| A feasibility study of hypernuclei reconstruction at NICA/MPD  | ILIEVA, M.        |
| Search for Higgs Bosons decaying to $B\bar{B}$ produced in association with $t\bar{t}$ at the LHC using the ATLAS Detector | KILBY, C.         |
| Measurement of the CP asymmetry in $B_s^0 - \bar{B}_s^0$ mixing  | KLAVER, S.        |
| Search for displaced dileptons at ATLAS  | KRAUSS, D.        |
| Search for sbottom quarks in $\sqrt{s}=13$ TeV pp collisions with the Atlas detector                                       | LERNER, G.        |
| b-Tagging in the ATLAS experiment in pp collisions at $\sqrt{s}=13$ TeV  | LE QUILLEUC, E.   |
| An inclusive trigger for charm physics at LHCb   | MICHIELIN, E.     |
| Search for rare decays of H and Z Bosons to $\phi + \gamma$ with the ATLAS detector at 13 TeV                              | OWEN, R.          |
| Search for single B' decaying to bH in full hadronic channel   | RAUCO, G.         |
| Fake background estimation in the 13 TeV $H \rightarrow \tau\tau$ analysis at the ATLAS experiment                         | SAMMEL, D.        |
| Search for new physics in fully hadronic final states using the $M_{T2}$ variable with the CMS detector                    | SCHÖNENBERGER, M. |
| Performance of Resistive Plate Chambers installed during the first long shutdown of the CMS experiment                     | SHOPOVA, M.       |
| $\Lambda_c$ baryon production in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV with the STAR Heavy Flavor Tracker          | SIMKO, M.         |
| Search for CP violation in $B^0 \rightarrow \rho^0 K^{*}(892)^0$   | VIEITES DÍAZ, M.  |
| New Neutrino Water-Scintillator Experiment WAGASCI at the J-PARC Beam  | VSIANNIKOVA, T.   |
| Ultra-Peripheral $J/\psi$ Production at CMS  | WALCZAK, M.       |
| New particle-flow based reconstruction of hadronic tau decays and intended use in Higgs CP studies at ATLAS                | YUEN, S. P.       |
| Search for WWW production with 8 TeV data with the ATLAS detector  | ZHANG, R.         |