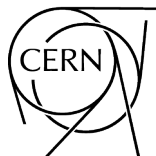


## **2017 CERN–Latin-American School of High-Energy Physics**

San Juan del Rio, Mexico  
8 – 21 March 2017

Editors: M. Mulders  
G. Zanderighi



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

ISBN 978-92-9083-502-8 (paperback)

ISBN 978-92-9083-503-5 (PDF)

ISSN 2519-8041 (Print)


ISSN 2519-805X (Online)

DOI <https://doi.org/10.23730/CYRSP-2018-004>

Accepted for publication by the CERN Report Editorial Board (CREB) on 14 November 2018

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

Copyright © CERN, 2018

 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 2017 CERN–Latin-American School of High-Energy Physics, San Juan del Rio, Mexico, 8 – 21 March 2017, edited by M. Mulders and G. Zanderighi, CERN Yellow Reports: School Proceedings, Vol. 4/2018, CERN-2018-007-SP (CERN, Geneva, 2018), <https://doi.org/10.23730/CYRSP-2018-004>

A contribution in this volume should be cited as:

[Author name(s)], in Proceedings of the 2017 CERN–Latin-American School of High-Energy Physics, San Juan del Rio, Mexico, 8 – 21 March 2017, edited by M. Mulders and G. Zanderighi, CERN Yellow Reports: School Proceedings, Vol. 4/2018, CERN-2018-007-SP (CERN, Geneva, 2018), pp. [first page]–[last page], <https://doi.org/10.23730/CYRSP-2018-004>. [first page]



## **Abstract**

The CERN–Latin-American 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 Standard Model of electroweak interactions, quantum chromodynamics, flavour physics, new physics beyond the standard model, heavy-ion physics, cosmology, and statistics for particle physicists.



## Preface

The ninth Event in the series of CERN–Latin-American Schools of High-Energy Physics took place from 8 to 21 March 2017 in San Juan del Rio, Mexico. It was organized by CERN with the support of Mexican colleagues from BUAP, CINVESTAV, MCTP, SMF, UAQ, UCOL, UGTO, UMSNH, UNAM and USON.

The School received financial support from: CERN; CIEMAT, Spain; and the Mexican national funding agency *Consejo Nacional de Ciencia y Tecnología* (CONACyT). Financial and in-kind contributions were also received from the following institutes and universities: UAQ; Facultad de Ciencias Fisico Matematicas, BUAP; Instituto de Ciencias Nucleares, UNAM; Division de Ciencias Exactas y Naturales, USON; and Secretaria de Turismo de Queretaro. Our sincere thanks go to all of the sponsors for making it possible to organize the School with a large number of young participants from Latin-American countries, many of whom would otherwise not have been able to attend. We would particularly like to thank CONACyT for their generous financial support, and UAQ and UNAM for their important in-kind contributions and practical help.

The School was hosted in the beautiful and traditional Hotel Hacienda Galindo near to San Juan del Rio. We are indebted to the hotel and its friendly staff for their help in making the Event such a success. In particular, we would like to mention Luis de la Barrera Martínez and Karina Piña Hernández who helped us greatly in preparing the School, as well as during the Event itself.

Malena Tejeda Yeomans from USON acted as local director for the School, assisted by members of the local organising committee. We are extremely grateful to Malena and her colleagues from BUAP, CINVESTAV, UAQ, UCOL, UGTO, UMSNH, UNAM and USON for their excellent work in organizing the School and for creating such a wonderful atmosphere for the participants.

Seventy-five students of 18 different nationalities attended the School, including 26 from Mexico. Following the tradition of the School the students shared twin rooms mixing nationalities, and in particular the Europeans mixed with Latin Americans.

The 12 lecturers came from Europe, Latin America and the USA. The lectures, which were given in English, were complemented by daily discussion sessions led by five physicists coming from Latin America. The lectures and the discussion sessions were all held using the conference facilities of the hotel. The students displayed their own research work in the form of posters in a special evening session during the first week. The posters were left on display until the end of the School. The students from each discussion group also performed a project, studying in detail the analysis of a published paper from an LHC experiment. A representative of each group gave a brief summary talk during a special evening session during the second week of the School.

Our 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 undoubtedly appreciated their personal contributions in answering questions and explaining points of theory.

The opening ceremony of the School included a question-and-answer session with Julia Tagüeña Parga, Deputy Director for Scientific Development, CONACyT, and Fabiola Gianotti, Director General of CERN, who joined via a live video link.

Hosting the School was an important event for the physics community in Mexico and many Outreach activities were arranged around it, benefitting from the presence of high-level scientists who were teaching at the School. In particular, nine public lectures were arranged at the UAQ and UNAM campuses in nearby Queretaro, and there was substantial coverage published online via the CONACyT Press Office.

A highlight of the Outreach programme associated with the School was a public lecture in Spanish on 8 March, International Women's day. The lecture, *Einstein, agujeros negros y ondas gravitacionales* (Einstein, black holes and gravitational waves), was given by Gabriela Gonzalez, spokesperson of the LIGO collaboration. The programme included introductory talks by Julia Tagüeña Parga and Fabiola Gianotti, illustrating the role of women in top-level scientific functions.

We are very grateful to Kate Ross, the administrator for the CERN Schools of Physics, for her efforts in the lengthy preparations for the School and during the Event itself. Her efficient work, friendly attitude, and continuous care of the participants and their needs were highly appreciated.

The participants will certainly remember the two interesting excursions: an afternoon visit to the nearby town of Bernal, followed by a wine tasting and dinner at the Bodega Cote Vineyard; and a full-day excursion to the outstanding archaeological site of the Teotihuacan pyramids, followed by dinner at a traditional gorditas restaurant in Queretaro. They also greatly appreciated evenings spent together in the hotel, especially the farewell party on the last night.

The success of the School was to a large extent due to the students themselves. Their poster session and group projects were very well prepared and highly appreciated, 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	Martijn Mulders	47	Giacomo Caria
2	Matheus Hostert	48	Arturo Rodriguez Rodriguez
3	Jorge Castano	49	Juan Carlos De Haro Santos
4	Pedro Roberto Mercado Lozano	50	Elias Polanco Euan
5	Ron Shellard	51	Alejandro Szykman
6	Rose Ardell	52	Alejandro Ayala
7	Karla Pena Rodriguez	53	Stoyan Trilov
8	Joaquin Hoya	54	Laura Calcagni
9	Otto Carlos Lippmann	55	Mariel Estevez
10	Kate Ross	56	Annie Meneses Gonzalez
11	Jose Andres Monroy Montanez	57	Ignacio Alberto Monroy Canon
12	Juan Cristobal Rivera	58	Antonio Pich
13	Omar Vazquez Rueda	59	Freja Thoresen
14	Camilo Salazar	60	Denys Yen Arrebato Villar
15	Antonio Rojas Ramos	61	Jordi Salinas
16	Manuel Morgado	62	Paulina Valenzuela Coronado
17	Sofia Escobar	63	Marcus O'Flaherty
18	Adriana Vazquez	64	Simen Hellesund
19	Julia Pena	65	Jose Carlos Castillo Fallas
20	Edgardo Marbello Santrich	66	Maria Di Domenico Franco
21	Jose Manuel Quimbayo Garcia	67	Luis Hernandez
22	Carlos Adrian Jimenez Carballo	68	Ivania Maturana
23	Rogelio Rosenfeld	69	Gino Marceca
24	Maria Rosario Fraire Bojorquez	70	Edgar Chavarria
25	Jorge Luis Gutierrez Santiago	71	David Zhu
26	Edgar Dominguez Rosas	72	Hector Bello Martinez
27	Javier Cobos	73	Luis Valenzuela Cazares
28	Saul Hernandez	74	Pablo Rivadeneira
29	Adriana Perez Martinez	75	Selomit Ramirez
30	Luiz Gustavo Silva De Oliveira	76	Santeri Laurila
31	Gilson Correia Silva	77	Jorge Venzor
32	Marxil Sanchez Garcia	78	Martha Cecilia Duran Osuna
33	Jose Carlos Jimenez Apaza	79	Jorge Stalin Martinez Armas
34	Rafael Aoude	80	Anthony Mard Calatayud Cadenillas
35	Jamerson Rodrigues	81	Roger Hernandez-Pinto
36	Paris Sphicas	82	Ignacio Fabre
37	Roberto Padron Stevens	83	Christian Laner
38	Danilo Silva De Albuquerque	84	Federico Von Der Pahlen
39	Suxer Lazara Alfonso Garcia	85	Iraq Rabadan
40	David Rivera Rangel	86	Daniel Perez
41	Malena Tejada Yeomans	87	Gero Von Gersdorff
42	Federico Sanchez	88	Rogelio Reyes Almanza
43	Mayra Alejandra Rivera Ruiz	89	Azarael Yebra Perez
44	Matias Fernandez Lakatos	90	Rosa Luz Zamora Peinado
45	Nick Ellis	91	Nhell Heder Cerna Velazco
46	Vinicius Franco Lima		



# Photographs (montage)



**9TH CERN LATIN-AMERICAN SCHOOL  
OF HIGH-ENERGY PHYSICS**

**San Juan del Rio, Mexico, 8-21 March 2017**









# Contents

Preface	
<i>N. Ellis</i> .....	v
Photograph of participants .....	vii
Photographs (montage) .....	x
Field Theory and the Electro-Weak Standard Model	
<i>C.A. García Canal</i> .....	1
QCD and the physics of hadronic collisions	
<i>M.L. Mangano</i> .....	27
Flavour dynamics and violations of the CP symmetry	
<i>A. Pich</i> .....	63
Beyond the Standard Model	
<i>M. Mondragon</i> .....	101
Heavy-Ion Physics	
<i>A. Ayala</i> .....	125
A sketchy introduction to Cosmology	
<i>R. Rosenfeld</i> .....	151
Probability and Statistics for Particle Physics	
<i>C. Mañá</i> .....	171
Organizing Committee .....	333
Local Organizing Committee .....	333
List of Lecturers .....	333
List of Discussion Leaders .....	334
List of Students .....	335
List of Posters .....	336





# Field Theory and the Electro-Weak Standard Model

C.A. García Canal

IFLP/CONICET and Departamento de Física

Universidad Nacional de La Plata, C.C.67, 1900, La Plata, Argentina

## Abstract

These lectures present an introduction to Quantum Field Theory and its particular application to the building of the Standard Model of Electro-Weak interactions.

## Keywords

Quantum Field Theory, Standard Model, Electroweak interactions

*”One should never underestimate the pleasure we feel  
from hearing something we already know”*

*Enrico Fermi*

## 1 Introduction to Quantum Field Theory

Even if Quantum Mechanics was extremely successful in describing and predicting the atomic physics phenomena, it clearly appears that the inclusion of Relativity in its treatment produces serious difficulties.

We had learned that as soon as the characteristic distances of a given problem are so small as the atomic size ( $1\text{Å} = 10^{-8}\text{ cm}$ ) or smaller, the Newton Mechanics has to be replaced by Quantum Mechanics with all its consequences. We have to remember that in these conditions we are into the validity domain of the Heisenberg Uncertain Principle, the basic brick of Quantum Theory.

$$\Delta x \Delta p \simeq \hbar \quad (1)$$

The Planck constant  $\hbar$  provides the scale where quantum effects are protagonists. This principle implies that it is not possible to define a trajectory for the quantum objects because you cannot provide simultaneously the position and the velocity of the particle, necessary data for the integration of Newton equation.

The uncertain principle has still further and important consequences. In fact, if the distances implied are smaller than the Angstrom, then  $\Delta p$  should be larger. In other words, the velocities implied will be larger and even comparable to the light velocity. In this case it is essential to take into account the Special Relativity rules. In particular, the well known equation  $E = mc^2$  has as a consequence that with sufficient energy, mass can be generated in the form of new quantum particles. The number of particles can change, is not more a constant of movement. It is necessary to build up a formalism able to treat a variable number of particles, of quantum objects. This formalism is the Quantum Field Theory (QFT). This is, at present, the best starting point to study and to describe matter and interaction at the most elementary level. It contains the possibility of creation and/or annihilation of quantum particles as for example electrons, photons and quarks.

QFT provides a set of formal strategies and mathematical tools that give rise to an image of the micro-world completely different to the classical conception of particles and fields (as the classical electromagnetic and gravitational ones).

As the paradigm of QFT is the quantum version of the electromagnetic field: Quantum Electrodynamics. Then, it is worth discussing, even briefly, the concept of a classical field that one uses for describing electricity, magnetism and the gravitational macroscopic forces.

Let us remember the Coulomb interaction between point charges that implies that electric charges interact even if they are not in contact, situation known as an action at a distance. To avoid this situation, the concept of field was born that among its important consequences it is mandatory to mention the prediction of electromagnetic waves. Electric and magnetic fields are propagated in space and time as waves. The situation of the electromagnetic field is reproduced by the gravitational interaction with only attractive forces in this case.

We should realize that at this point we are living together with two different phenomena related to the dynamics. From one side we have the wave process that implies propagation of a perturbation without a net displacement of matter and from the other side we have the displacement of concentrated elements as the classical particles are.

From the beginning of XX century one has fundamental advances in the understanding of the intimal constitution of matter, namely, the quantum hypothesis of Planck to take care of the black body radiation and the photon concept introduced by Einstein to explain the photoelectric effect. The quantum presence is not only in the process of emission of radiation, but also in the way energy travels. On the other side, the association of a characteristic wave length, proportional to its momentum, to a particle was proposed by de Broglie. This was confirmed by the experimental detection of diffraction of electrons by crystals. These phenomena gave rise to the wave-particle duality dilemma. Today we know that this duality is not more than the result of prejudicially pretending to keep the language of classical design (appropriate for the human scale) for the description of the atomic and subatomic phenomena.

Consequently, for a coherent treatment of the observed phenomena related to a wave character connected with the probability of presence of the quantum object, that takes into account the possibility of a variable number of the quantum particles according to the energy, it was necessary to develop the quantum formalism of QFT.

Quantum Electrodynamics (QED) is the paradigm of QFT and explain with astonishing precision the interaction between electrons (the electron field) and the quantum electromagnetic field (the field of photons). The quantization of the electromagnetic field implies the presence of photons as quanta or quantum excitations of the field. On the other hand, it contains a relativistic treatment of the electron.

The process of quantizing a field contains two steps. First one produces an analysis of the classical fields in terms of normal modes, namely a Fourier analysis corresponding to infinity degrees of freedom. Then each mode (each Fourier component) is independently described as a quantum harmonic oscillator. As a result, the energy of the system is expressed as the sum of terms corresponding to the energy of each oscillator, weighted by the number of oscillators with each particular energy. This number is a quantum operator and counting the number of objects in each mode is synonymous of the quantum particles (quanta) corresponding to the field in the state characterized by each possible energy. This is clearly the interpretation of the QFT in terms of quantum particles. Notice again that these quantum particles are completely different objects to the classical particles. They only share the property of being discrete entities that can be counted.

The following step in the development of the formalism of Quantum Field Theory is to couple different fields following precise rules that are mainly based on symmetry.

## 2 Introduction to the Formalism of QFT

In QFT, to each quantum particle species one assigns a **field**  $\chi(t, \vec{x})$  so that, this correspondence with classical mechanics is performed.

**Classical Mechanics** Point mass  $m$  in one dimensional space  $\Rightarrow$  a generalized coordinate  $q(t)$

Lagrangian:  $L = L(q(t), \dot{q}(t))$

**Classical Field Theory** Field in three dimensions  $\Rightarrow$  a generalized coordinate  $\chi(t, \vec{x})$  at each space-time point

Lagrangian:  $L = \int d^3x \mathcal{L}$ ; where the Lagrangian density is  $\mathcal{L} = \mathcal{L}(\chi, \dot{\chi}, \vec{\nabla}\chi)$   
 Now, from the action  $S = \int dt L \Rightarrow$  equations of motion (Euler-Lagrange):

### Classical Mechanics

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = 0$$

### Classical Field Theory

$$\frac{\partial \mathcal{L}}{\partial \chi} - \frac{\partial}{\partial t} \frac{\partial \mathcal{L}}{\partial \dot{\chi}} + \vec{\nabla} \frac{\partial \mathcal{L}}{\partial \vec{\nabla}\chi} = 0$$

Let us first consider free fields ( $\mathcal{L}_{int} = 0$ ). Through their behavior under Lorentz group transformations, one distinguishes i) scalar fields; ii) spinor fields; iii) vector fields; ...

#### i) Scalar field

It corresponds to spin 0. The field equation is the Klein-Gordon one and the Lagrangian density reads:

$$\mathcal{L}_0^\phi = \frac{1}{2} \left[ \left( \frac{1}{2} \frac{\partial \phi}{\partial t} \right)^2 - (\vec{\nabla} \phi)^2 \right] - \frac{1}{2} m \phi^2 \quad (2)$$

#### ii) Spinor field

It corresponds to spin 1/2. The corresponding field equation is de Dirac equation and the Lagrangian density:

$$\mathcal{L}_0^\psi = \bar{\psi} (\gamma^\mu \partial_\mu - m) \psi \quad (3)$$

#### iii) Vector field

It corresponds to spin 1. Maxwell equations are the corresponding ones. The Lagrangian density is:

$$\mathcal{L}_0^A = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \quad (4)$$

The next step is to take into account interactions between fields. We postpone the discussion of this point for a while.

The third step is to make the theory quantal. The process of quantizing the field theory is done, in our presentation, by "copying" Quantum Mechanics. Namely, we define the canonical momentum through the field version of

$$p(t) \equiv \frac{\partial L}{\partial \dot{q}}$$

and impose the canonical commutation relation, now with  $q$  and  $p$  read as operators, equivalent to

$$[q(t), p(t)] = i\hbar \quad (5)$$

that guarantees the validity of the uncertainty principle.

In this way, for

#### i) Spin 0:

$$\pi(t, \vec{x}) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(t, \vec{x})}$$

and the quantization is imposed through the equal time commutation relation

$$[\phi(t, \vec{x}), \pi(t, \vec{x}')] = i\hbar \delta^{(3)}(\vec{x} - \vec{x}') \quad (6)$$

ii) **Spin 1/2:**

$$\pi_\alpha(t, \vec{x}) = \frac{\partial \mathcal{L}}{\partial \dot{\psi}_\alpha(t, \vec{x})}$$

with the equal time anti-commutators

$$\{\psi_\alpha(t, \vec{x}), \pi_\beta(t, \vec{x}')\} = i\hbar \delta_{\alpha\beta} \delta^{(3)}(\vec{x} - \vec{x}') \quad (7)$$

i) **Spin 1:**

$$\pi^\mu(t, \vec{x}) = \frac{\partial \mathcal{L}}{\partial \dot{A}_\mu(t, \vec{x})}$$

with the quantization imposed by

$$[A^\mu(t, \vec{x}), \pi^\nu(t, \vec{x}')] = i\hbar g^{\mu\nu} \delta^{(3)}(\vec{x} - \vec{x}') \quad (8)$$

Notice that in this case there are some difficulties because gauge invariance of electromagnetism implies  $A^3 = 0$  and consequently, for these components  $A^3$  and  $\pi^3$  one cannot satisfy the previous relation. This problem needs a particular treatment for the quantization of the electromagnetic field (See Bibliography).

After the quantization procedure we have ended with a series of field operators. The natural question now is where are the quantum particles?

Let us answer this question by analyzing the scalar field, whose equation of movement is

$$(\square + m^2) \phi(t, \vec{x}) = 0$$

Taking profit of the fact that it is a linear equation, we write its general solution in terms of the Fourier transform

$$\phi(t, \vec{x}) \propto \int dE d^3 p \delta(E^2 - \vec{p}^2 - m^2) \left[ a(E, \vec{p}) e^{-i(Et - \vec{p} \cdot \vec{x})} + a^\dagger(E, \vec{p}) e^{+i(Et - \vec{p} \cdot \vec{x})} \right]$$

Due to the quantization condition (6), the Fourier coefficients  $a(E, \vec{p})$  and  $a^\dagger(E, \vec{p})$  must be operators, clearly satisfying the commutation relations

$$[a(E, \vec{p}), a^\dagger(E, \vec{p}')] = 2E\hbar \delta^{(3)}(\vec{p} - \vec{p}') \quad (9)$$

$$[a(E, \vec{p}), a(E, \vec{p}')] = 0 \quad (10)$$

$$[a^\dagger(E, \vec{p}), a^\dagger(E, \vec{p}')] = 0 \quad (11)$$

Consequently, the Hamiltonian of the system can be written as

$$\begin{aligned} H &= \int d^3 x (\pi \dot{\phi}) \\ &\simeq \int dE d^3 p \delta(E^2 - \vec{p}^2 - m^2) E a^\dagger(E, \vec{p}) a(E, \vec{p}) a(E, \vec{p}) \end{aligned}$$

where appears the "number operator"

$$N(p) = N(E, \vec{p}) \equiv a^\dagger(E, \vec{p}) a(E, \vec{p}) \quad (12)$$

acting in the multiparticle-states Fock space, verifying the eigenvalue equation

$$N(E, \vec{p}) |n(E, \vec{p})\rangle = n(E, \vec{p}) |n(E, \vec{p})\rangle$$

that allows one to interpret  $n(E, \vec{p})$  as the number of quanta with spin 0, mass  $m$ , energy between  $E$  and  $E + dE$  and momentum between  $\vec{p}$  and  $\vec{p} + d\vec{p}$ . Certainly the validity of the relationship  $E = \sqrt{\vec{p}^2 + m^2}$  is implicit. The number operator has the property

$$N(p) a^{(\dagger)}(p) |n(p)\rangle = \left[ n(p) \begin{matrix} - \\ + \end{matrix} 1 \right] a^{(\dagger)}(p) |n(p)\rangle$$



that induce the names: creation operator for  $a^{(\dagger)}(p)$  and annihilation operator for  $a(p)$ , with the corresponding eigenvalue equations.

The previous analysis clearly shows that the quantization procedure provides the connection between quantum fields and quantum particles.

One can now consider multiparticle states. To this end one has to recall that the indistinguishability between identical quantum particles makes necessary to introduce the corresponding statistic. Namely Bose-Einstein for integer spin and Fermi-Dirac for half integer spin.

In the case of Bose-Einstein, a multi-boson state reads

$$|n_1(p_1), \dots, n_m(p_m)\rangle \propto [a^{(\dagger)}(p_1)]^{n_1} \dots [a^{(\dagger)}(p_m)]^{n_m} |0\rangle$$

that presents a total symmetry under the interchange of any pair of particles.

For the case Fermi-Dirac one has to ensure the validity of the Pauli exclusion principle that implies using anticommutators (as we have already used in the process of quantization) instead of commutators because

$$\{a^{(\dagger)}(p), a^{(\dagger)}(p)\} = 0 \Rightarrow (a^{(\dagger)}(p))^2 = 0 \Rightarrow n_i = 0, 1$$

and the multiparticle fermion state reads

$$|p_1, \dots, p_m\rangle = a^{(\dagger)}(p_1) \dots a^{(\dagger)}(p_m) |0\rangle$$

Notice that the use of anticommutators for fermions guarantees also that the energy of the system is bounded from below.

Let us now go to discuss the *parameters* and the *observables* of a quantum field theory.

In general, to specify a field theory is equivalent to give a Lagrangian. For example, for the scalar case one gives  $\mathcal{L}(\phi, \partial\phi)$ . Lets take the simple case with a quartic self-interaction, namely

$$\mathcal{L} = \frac{1}{2} [\partial_\mu \phi_B(x)]^2 - \frac{1}{2} m_B^2 \phi_B^2 - \frac{1}{4} g_B \phi^4 \quad (13)$$

where the index  $B$  is there for "bare", the initial value before any interaction, measured by  $g_B$  has occurred. Remember that the field and the corresponding canonical momentum verify (6). If we compute now physical observables like cross-sections, or decay rates, or the physical mass, all of them result functions of the bare parameters  $m_B$  and  $g_B$ . Consequently, any perturbative calculation one can do (following the similar mechanism we use in Quantum Mechanics) will end with a series in powers of  $g_B$ . The construction of the perturbative series has a protocol based on clear rules known as Feynman diagrams. The reader can consult, for example, these specific recent references where Feynman diagrams are treated: i) K. Kumerički, "Feynman diagrams for beginners", arXiv: 1602.04182; ii) S.M. Bilenky, "Introduction to Feynman diagrams", vol.65 in International series in natural philosophy, ISBN: 978-0-08-017799-1.

Feynman diagrams are a pictorial representation of probability amplitudes at a given order of the perturbation theory. Every line and every cross in the picture has a mathematical interpretation, in a similar way as in an electrical circuit.

In any case, calculations could depend on other parameter. In fact, it is valid to perform the quantization as

$$[A^\mu(t, \vec{x}), \pi^\nu(t, \vec{x}')] = \frac{i}{Z_\phi} \hbar g^{\mu\nu} \delta^{(3)}(\vec{x} - \vec{x}')$$

where  $Z_\phi$  is an arbitrary number. Certainly,  $Z_\phi = 1$  corresponds to the bare situation. It is also clear that any magnitude you can compute, say the  $S$ -matrix, will start depending on  $Z_\phi$ .

The important fact to notice is that both,  $m_B$  and  $g_B$  do not contain any physics. In fact, one can ask, for example, which is the physical mass of the quantum of the field  $\phi$ . It is the value of  $p^2$ , the square

of the momentum, where the propagator of the quantum particle has a pole. If there were no interaction ( $g_B = 0$ ) then of course  $P_{pole}^2 = m_B^2$ . But if interactions are present, the pole of the propagator is at a value  $p_{pole}^2 = m^2 \neq m_B^2$ . It acquires infinite corrections!

This simple example shows the necessity of *renormalization*.

We present briefly the idea of renormalization. If  $m_b$ ,  $g_B$  and  $Z_\phi = 1$  are maintained fixed, the perturbative contributions, coming from the calculation of Feynman diagrams, ends with divergent integrals. Take as an example the electron selfenergy in QED. It has at the lower order a logarithmic divergence. If a cut-off  $\lambda$  in momentum is introduced, one gets in this example a result  $I \propto e_B^2 \ln \lambda$  (in this case the bare electron electric charge plays the role of  $g_B$ ).

So under this cut-off intent, in general one ends with S-matrix elements  $S_B = S_B(p_i, m_B, g_B, Z_\phi; \lambda)$  and for many of these contributions, the (physical) limit  $\lambda \rightarrow \infty$  gives rise to the non sense result  $S_B \rightarrow \infty$ .

Just to try to get sensible results from the perturbation theory approach to quantum field theory, the renormalization scheme was proposed. Namely, to allow that

$$\begin{aligned} m_B &= m_B(\lambda) \\ g_B &= g_B(\lambda) \\ Z_\phi(\lambda) &\neq 1 \end{aligned}$$

and to adjust the functional dependence on  $\lambda$  in order to cancel the divergencies that appear when  $\lambda \rightarrow \infty$ . This is achieved computing some  $S_B$  and comparing the results with experimental data in order to deduce the dependence of the above functions, for example  $m_B = m_B(\lambda)$ . Clearly, a priori there exist a big problem, namely, in principle there is an infinite number of potentially divergent contributions and only three functions to be adjusted. But extraordinary cases exist... Some theories are certainly renormalizable and three function to adjust are enough. Those theories where the number of terms that are independently divergent (primitive divergent as they are called) is equal or less than the functions of  $\lambda$  to adjust.

The practice of renormalization goes through the replacement

$$\begin{aligned} \phi_B &= \sqrt{Z_\phi(\lambda)} \phi \\ g_B &= Z_\phi^{-2}(\lambda) Z_g(\lambda) g \\ m_B^2 &= Z_\phi^{-1}(\lambda) Z_m(\lambda) m^2 \end{aligned}$$

and the adjustment of the  $Z_i(\lambda)$  so to get  $g$  and  $m$  finite and independent of  $\lambda$ . These  $Z_i(\lambda)$  are treated as power series in  $g$ . The relations before seems capricious but they are chosen in this way because the Lagrangian written in terms of the new magnitudes reads simple

$$\mathcal{L} = \frac{1}{2} Z_\phi [\partial_\mu \phi(x)]^2 - \frac{1}{2} Z_m m^2 \phi^2 - \frac{1}{4} Z_g g \phi^4 \quad (14)$$

that certainly can be treated with the Feynman diagrams technique, now in terms of 'dressed' constants.

### 3 Standard Model

What is what one understands for a Standard Model? It is a theoretical framework that starting from observation allows to predict and correlate new data. In general it is an excellent "approximation" at a given (energy) scale. Consequently, one is always driven to go beyond. In Physics, the Standard Model has had a nice evolution. We can first mention the proposal of Empedocles of considering the four elements: earth, water, air and fire, as the compounds of everything, linked by love and hate. The next step in the evolution of the Standard Model is Mendeleev's periodic table, the base of Chemistry. After this, Quantum Mechanics, the theory of atoms, is the following Standard Model. Today Standard Model

is  $3!$ , the Glashow-Salam-Weinberg electroweak model plus Quantum Chromodynamics that includes quantum particles and quantum fields with validity down to distances of the order of  $10^{-18} m$ .

The present Standard Model contains, in an extremely economical way, the ingredients to describe (almost) everything. It is based upon *Gauge Symmetry*. This symmetry is the engine of the present physics development allowing a unified treatment of the fundamental forces. This treatment conforms the *Gauge theories*. A Gauge theory is a synthesis of Quantum Field Theory with a particular symmetry. The idea of a gauge theory was introduced by Hermann Weyl in 1919. At that time, only the electron and the proton were known...It was an idea for future times.

#### 4 Symmetries

Wigner, referring to Einstein's relativity, stated: "Einstein's work sets the inversion of a tendency: before it, invariance principles came from the dynamical laws. Now, is natural for us, to obtain the laws of Nature from invariance principles". This is the way our present knowledge of the fundamental interactions is obtained.

Certainly, people love symmetry and also Nature loves symmetry, or better said: our model of Nature loves symmetry. In the same way that we act in front of a framed picture that is not in the right position because we are compelled to exert a force in order to restore the broken axial symmetry, fundamental forces of Nature are present just to ensure the validity of a symmetry, the gauge symmetry. In order to arrive to this concept, let us start by recalling which are the symmetries in Physics.

One can divide the symmetries in Physics into two types: i) Geometrical symmetries, related to transformation on the space-time coordinates that have to do with observational situations and ii) Internal symmetries, acting on the dynamical variables and operators, having to do with the quality of observables. These internal symmetries could be global if they have no contact with space-time, or local when the parameters of the internal transformations are connected with space-time. These local internal symmetries are *gauge symmetries*.

The important point for the present discussion is that in order to guarantee a gauge symmetry, particularly in QFT, it is mandatory to introduce new fields into the game. These extra fields produce the interactions. This is the framework for formalizing the fundamental interactions in Nature.

In particular, it is the main ingredient in building the  $3!$ -model, the Standard Model. We call it  $3!$  because  $3! = 3 \times 2 \times 1$ , that remember us that the Standard Model is based on the requirement of the Gauge Symmetry

$$3! = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$$

where the symmetry  $SU(3)_C$  is responsible for the strong interaction among quarks that carry color charge and  $SU(2)_L \otimes U(1)_Y$  is related to weak and electromagnetic interactions that are treated in a unified way and named electroweak.

The gauge symmetry is guaranteed through the inclusion of gauge fields, mediators of the interactions. All the interaction between fermion fields (leptons or quarks) are carried by the exchange of vector-boson gauge fields (gluons, weak-bosons, photon).

The known fermions, the characters of the "drama", appear in three generations and are related among them via gauge transformations. Vertically are connected by  $SU(2)_L$  and quarks, horizontally, via  $SU(3)_C$ . We summarize them in this table

$$\left[ \begin{array}{c|ccc} \nu_e & u_1 & u_2 & u_3 \\ e & d_1 & d_2 & d_3 \end{array} \right]; \left[ \begin{array}{c|ccc} \nu_\mu & c_1 & c_2 & c_3 \\ \mu & s_1 & s_2 & s_3 \end{array} \right]; \left[ \begin{array}{c|ccc} \nu_\tau & t_1 & t_2 & t_3 \\ \tau & b_1 & b_2 & b_3 \end{array} \right]$$

*Fermion Generations or Families*

We include below the mediators of the interactions that we shall relate briefly with gauge fields.

<i>gluons</i>	$g_1, g_2, \dots, g_8$
<i>photon</i>	$\gamma$
“ <i>weakons</i> ”	$W^+, W^-, Z$

*Intermediate Bosons*

Notice that these are the "characters" at the level of elementarity we recognize at present and that is why they bear today the name of "elementary (?) particles (?)".

We also should add the Higgs field  $H$ , responsible, as we will see, of the masses of the massive particles.

## 5 Symmetries in Field Theory

Hopefully, we have convinced the audience that the natural language for Particle Physics is Quantum Field Theory (QFT). The fundamental magnitude to start with is a quantum field  $\chi_\alpha(x)$  where  $\alpha$  summarized all the internal indices needed (as the components, etc), and now symmetries need to be incorporated in the corresponding Lagrangian. For example in

$$\mathcal{L} = \mathcal{L}(\phi, \partial_\mu \phi)$$

We start considering continuous symmetries whose transformations are represented by elements of a continuous group  $G$ . Then the field  $\chi_\alpha$  will be a member of an irreducible multiplet that under a transformation  $a$  belonging to the group  $G$  transform as

$$\chi_\alpha(x) \xrightarrow{a} \chi'_\alpha(x) = \mathcal{R}_{\alpha\beta}(a) \chi_\beta(x)$$

where  $\mathcal{R}_{\alpha\beta}$  is the matrix representation of  $G$  that verifies

$$\mathcal{R}_{\alpha\beta}(a') \mathcal{R}_{\beta\gamma}(a) = \mathcal{R}_{\alpha\gamma}(a'')$$

in order to satisfy the group product equivalence

$$\chi_\alpha(x) \xrightarrow{a} \chi'_\alpha(x) \xrightarrow{a'} \chi''_\alpha(x) \equiv \chi_\alpha(x) \xrightarrow{a''} \chi''_\alpha(x)$$

Going now to the Hilbert space, transformations are represented by unitary operators  $U(a)$  and one has, being  $\chi_\alpha$  a member of an irreducible multiplet, that

$$U^{-1}(a) \chi_\alpha(x) U(a) = \chi'_\alpha(x) = \mathcal{R}_{\alpha\beta}(a) \chi_\beta(x) \quad (15)$$

and

$$U(a) U(a') = U(a'') \quad (16)$$

As we are treating continuous symmetries, it is possible to study their infinitesimal version and write

$$U(\delta a) = 1 + i \delta a_i G_i$$

with  $\delta a$  real.  $G_i$  are the generators of the group  $G$ , being hermitian operators. Due to the composition law (16) one has

$$[G_i, G_j] = i c_{ijk} G_k \quad (17)$$

namely, the Lie algebra of the group.  $c_{ijk}$  are the structure constants of the group. It is clear that the representation of the group for infinitesimal transformation can be written as

$$\mathcal{R}_{\alpha\beta}(\delta a) = \delta_{\alpha\beta} + \iota \delta a_i (g_i)_{\alpha\beta}$$

where  $g_i$  are the representations of the generators  $G_i$  and obey the Lie algebra.

By using the relation (15), one immediately gets

$$[G_i, \chi_\alpha(x)] = -(g_i)_{\alpha\beta} \chi_\beta(x)$$

that shows how the field transforms under the group  $G$ .

Clearly, the invariance of a field theory under the group  $G$  implies that the action is invariant. Consequently, via the Noether theorem, there are currents

$$J_i^\mu(x) = \frac{\partial \mathcal{L}}{\partial \partial_\mu \chi_\alpha(x)} \frac{1}{\iota} (g_i)_{\alpha\beta} \chi_\beta(x)$$

that are conserved:

$$\partial_\mu J_i^\mu(x) = 0$$

allowing the identification of charges

$$G_i = \int d^3x J_i^0(x)$$

that are also conserved, they commute with the Hamiltonian.

Let us now consider quantum states as  $|p; \alpha\rangle$  corresponding to a one particle with  $p^2 = -m^2$ .

The particle states corresponding to a field  $\chi_\alpha(x)$  transform as the field if and only if the vacuum is  $G$ -invariant, namely

$$U(a) |0\rangle = |0\rangle$$

that implies that the generators annihilate the vacuum

$$G_j |0\rangle = 0$$

In general this is not the case and for this reason one divides the realization of a given symmetry according to the behavior of the vacuum being  $G$ -invariant or not.

### Wigner-Weyl realization

The vacuum is  $G$ -invariant, or in other words, it is annihilated by the generators and consequently, it has the same symmetry as the action. Then the field has vacuum expectation value equal to zero

$$\langle 0 | \chi_\alpha(x) | 0 \rangle = 0$$

When this condition is fulfilled, there is a theorem that shows that all the states in a given multiplet have the same mass. This is the case, for example, of proton and neutron that if isospin would be a perfect symmetry, they should have the same mass.

### Nambu-Goldstone realization

In this case the vacuum is not  $G$ -invariant. In other words

$$U(a) |0\rangle \neq |0\rangle$$

and consequently

$$\langle 0 | \chi_\alpha(x) | 0 \rangle \neq 0$$

Consequently, the Goldstone theorem applies. It says that in a theory, for each generator  $G_j$  that does not annihilate the vacuum there is present one massless boson, called Goldstone boson. That is

$$G_j^{N.A.} |0\rangle = |p; j\rangle_G \text{ with } m = 0$$

## 6 Gauge theories

From the original era of gauge invariance, started by Weyl (1919), it survived mainly as Maxwell equations' symmetry until around 1959 when Yang and Mills proposed an extension of gauge symmetry beyond electromagnetism.

The basic ideas of a theory based upon gauge invariance, a gauge theory can be mimicked by the following very simple example of a harmonic oscillator rotating in a plane with period  $T = 2\pi/\omega$ . Referring to standard coordinates, the motion is represented by

$$\begin{aligned} y &= A \sin(\omega t) \\ x &= B \cos(\omega t) \end{aligned}$$

that can be written together by means of a complex variable  $z = x + iy$ . The oscillator equation now reads

$$\frac{d^2 z}{dt^2} + \omega^2 z = 0 \quad (18)$$

In fact, to decide measuring the instant position of the oscillator  $\theta = \omega t$ , from the horizontal  $x$ -axis is arbitrary. One certainly can choose an alternative orthogonal system  $(x', y')$  and measure angles from  $x'$  rotated an angle  $\alpha$  from  $x$ . We are changing  $\theta \rightarrow \theta - \alpha$  and say that  $\theta$  was "regauged". Notice now that multiplying the Eq.(18) by  $\exp(-i\alpha)$  and redefining  $z' = z \exp(-i\alpha)$  the Eq.(18) is covariant (does not change written in terms of  $z'$ ). One can conclude that the absolute value of  $\theta$  is irrelevant or in more precise words, the equation is *global gauge invariant*. Global means that  $\alpha$  is time independent.

What happens if we allow for a *local gauge transformation*  $\alpha = \alpha(t)$ ? It is clear that  $\exp(-i\alpha(t))$  cannot be absorbed in the redefinition of  $z'$  because  $d\alpha(t)/dt \neq 0$ .

One can regain the invariance, now a local one, compensating the time derivative of  $\alpha$  by means of the replacement

$$\frac{d}{dt} \rightarrow \frac{d}{dt} - A(t)$$

with the requirement that: if

$$\theta \rightarrow \theta - \alpha(t)$$

then

$$A(t) \rightarrow A(t) - \frac{d\alpha(t)}{dt}$$

that implies that  $A(t)$  works as a compensator, a *gauge field* allowing for the validity of the invariance under a local gauge transformation.

We clearly notice that when  $\alpha = \alpha(t)$  one is in fact rotating the reference frame with an angular velocity  $d\alpha(t)/dt$  and in a rotating frame, a non inertial frame, there are the so called "fictitious" forces (centrifuge, Coriolis). It is precisely  $A(t)$  the generator of these forces. The origin of new interactions comes from the requirement of local gauge invariance.

To go ahead let us now remember for a while Maxwell electromagnetic gauge invariance and in particular its origin. The vector  $\vec{A}$  is defined via its curl as  $\vec{\nabla} \wedge \vec{A} = \vec{B}$ , but there is no condition to be imposed to its divergence. This freedom in the election of  $\vec{\nabla} \cdot \vec{A}$  is precisely the gauge symmetry of electromagnetism. Taking into account the "vectorial poem" due to Enrique Loedel (1901-1962, former professor at La Plata) that says <sup>1</sup>

Esto el Papa exclamó al firmar la bula

---

<sup>1</sup>This was exclaimed by the Pope/ when furious excommunicates Luther:/ The divergency of a curl is zero/ and the curl of a gradient is always zero/ The great German priest pleaded god/ and exclaimed with his usual vehemence/ the curl of a curl plus nabra two/ gives the gradient of any divergence

con que furioso excomulgó a Lutero:  
 La *divergencia* de un *rotor* es nula  
 y el *rotor* de un *gradiente* es siempre cero.  
 El gran fraile alemán invocó a dios  
 y exclamó con su habitual vehemencia:  
 El *rotor* de un *rotor* más *nabla* *dos*  
 da el *gradiente* de toda *divergencia*.

one understand that the electromagnetic equations do not change if the replacement, with  $\Lambda$  a scalar function

$$\begin{aligned}\vec{A} &\rightarrow \vec{A}' = \vec{A} + \vec{\nabla}\Lambda \\ \phi &\rightarrow \phi' = \phi - \frac{\partial\Lambda}{\partial t}\end{aligned}$$

is performed, because both  $\vec{E}$  and  $\vec{B}$  are not changed. The gauge symmetry is clear. Moreover, gauge invariance of electromagnetism guarantees the charge conservation.

The previous comments open the possibility of using the field  $\vec{A}$  as compensator when one ask for local gauge invariance of a field theory. In fact, the quantum electromagnetic interaction mediated by photons, can be formalized starting from the requirement of local (abelian) gauge invariance. This framework can be generalized, as Yang and Mills proposed, to the case of a non-abelian group of symmetry.

In summary, the fundamental interactions can be described by theories with local gauge symmetry and consequently mediated by the corresponding compensator field, the corresponding gauge field.

Even if we will come back later to this important point, let us state a fundamental property of gauge fields: they are massless fields. In fact, a typical mass term  $\mathcal{L}_{mass} = \frac{1}{2} mA_\mu A^\mu$  for a vector field is clearly not gauge invariant.

## 7 Constructing a Gauge Theory for matter fields

### 7.1 "Standard" way

First of all, one chooses an appropriate unitary Lie group  $G\{g\}$ . Then one propose an action with global invariance under  $G$  as the matter symmetry group. Having then an action that is invariant under constant phase transformation, there appears the conservation of a Noether current. After this one promotes global invariance to a local one:  $g \rightarrow g(x) \in G$ , i.e. phases that are space-time dependent. Now one needs to introduce gauge fields to compensate, via a gauge transformation, the changes produced by the presence of local phases. Finally, a kinetic term for gauge fields has to be included in the resulting Lagrangian.

In all this procedure, as soon as  $g(x)$  are well behaved, the conserved Noether current of the global case is unchanged.

Let us consider a typical example. A complex scalar field and an one parameter group of symmetry.

$$\mathcal{L}(\phi, \partial_\mu \phi) = \partial_\mu \phi^* \partial^\mu \phi - m^2 \phi^* \phi$$

that is invariant under a global phase transformation (constant phase  $\alpha$ )

$$\phi(x) \rightarrow \phi'(x) = e^{i\alpha} \phi(x)$$

corresponding to the group  $U(1)$ : implements rotations on a unit circle.

Now we ask for a local phase transformation  $\alpha = \alpha(x)$ . The motivation for this requirement was given by Yang and Mills, stated that: "The concept of field and the concept of local interactions imply

a spreading of information to neighbouring points and eliminate the action at a distance. Then global phase invariance seems to contradict the generalized idea of locality”.

Performing the local gauge transformation

$$\phi(x) \rightarrow \phi'(x) = e^{i\alpha(x)} \phi(x) \quad (19)$$

in the Lagrangian above, one sees that the mass term stays without changes when written in terms of  $\phi'(x)$  while the kinetic term does not because

$$\partial^\mu \phi(x) \rightarrow e^{i\alpha(x)} [\partial_\mu + i\partial_\mu \alpha(x)] \phi(x)$$

The question is how to turn  $\mathcal{L}$  local invariant. Taking profit of the freedom of the electromagnetic field  $A_\mu$ , one replaces  $\partial_\mu$  by the gauge covariant derivative

$$D_\mu \phi(x) \equiv (\partial_\mu - iqA_\mu(x)) \phi(x) \quad (20)$$

with the constraint that when  $\phi(x)$  changes with the local phase  $\alpha(x)$  as in (19), the gauge field  $A_\mu$  suffer the gauge transformation

$$A_\mu(x) \rightarrow A'_\mu(x) = A_\mu(x) - \frac{1}{q} \partial_\mu \alpha(x) \quad (21)$$

where  $q$  clearly measures the coupling between de scalar field and the gauge field. In this way, the covariant derivative transforms exactly as the field. Consequently, the invariant theory under local phase transformations has the new Lagrangian

$$\mathcal{L} = [\partial_\mu + iqA_\mu(x)] \phi^*(x) [\partial_\mu - iqA_\mu(x)] \phi(x) - m^2 \phi^*(x)\phi(x) \quad (22)$$

that explicitly includes interaction, exactly the so called minimal interaction. The only way to have local phase invariance is including interaction with the gauge field that plays the dual action of a compensating field to restore the symmetry and of a comparative field that distinguishes charges in the case of the original complex field. As we have said before, one has to add the kinetic term of the gauge field, namely

$$-\frac{1}{4} F^{\mu\nu} F_{\mu\nu} \quad (23)$$

but one cannot add a mass term for  $A_\mu$ , because it breaks the phase symmetry.

Certainly, one can make the same treatment for fermion fields, to obtain

$$\mathcal{L}_f = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + i\bar{\psi} \gamma^\mu (\partial_\mu - ieA_\mu(x)) \psi - m\bar{\psi}\psi \quad (24)$$

and the same procedure could be performed in connection with non-abelian phase transformations. We will be back to this point below.

## 8 ”Upsidedown” way

It is possible and ”natural” to invert the process above avoiding the ”ukase”<sup>2</sup> that promotes global to local symmetries on the matter Lagrangian. The upsidedown proposal we have presented (C.A. García Canal and F.A. Schaposnik, ”Building Gauge Theories: The Natural Way”, Fundamental Journal of Modern Physics, 2 (2011) 15) starts with the symmetry of the gauge field, where it is natural. We build the gauge theory of fundamental interactions starting from the interaction mediators: the gauge fields. The sources (matter) are added imposing local gauge invariance and Lorentz invariance.

<sup>2</sup>ukase: have the power of laws but may not alter the regulations of existing laws



Let us explain the procedure starting with electrodynamics. The Maxwell equations, without sources are contained in

$$\partial_\mu F^{\mu\nu} = 0$$

the other two equations are given by Bianchi identity. Remember that

$$F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu$$

that under the gauge transformation (21) is invariant and of course also Maxwell equations are. The Maxwell Lagrangian (23) is gauge invariant and Lorentz invariant.

Let us include now an external (non-dynamical) source  $j_{ext}^\mu$  to write

$$\partial_\mu F^{\mu\nu} = e j_{ext}^\nu$$

Now the natural and simplest Lorentz invariant term in the Lagrangian formalism is

$$\mathcal{L}_{int} = e A_\mu(x) j_{ext}^\mu(x)$$

and in order to get a gauge invariant total Lagrangian, one can require that under a  $\Lambda(x)$  gauge transformation

$$j_{ext}^\mu(x) \rightarrow j_{ext}^{\mu\Lambda}(x) = j_{ext}^\mu(x)$$

and, forced by Maxwell

$$\partial_\mu j_{ext}^\mu(x) = 0$$

Notice that the Lagrangian including  $j_{ext}^\mu(x)$ , under a gauge transformation change at most as a total derivative.

As the structure of Maxwell equations is prescribed by Lorentz symmetry, one can anticipate the coupling of matter to the gauge field. To this end, let us consider a dynamical Dirac fermion field  $\psi(x)$  at the origin of the current  $j^\mu(x)$  (in principle different to  $j_{ext}^\mu(x)$ ). Certainly the most economic is a bilinear form of fermions, that due to Lorentz requirements reads

$$j^\mu(x) = \bar{\psi}(x) \gamma^\mu \psi(x)$$

that together with the gauge invariance of  $j^\mu(x)$ , implies that under a gauge transformation  $\Lambda(x)$

$$\psi(x) \rightarrow \psi^\Lambda(x) = e^{iq\Lambda(x)} \psi(x)$$

where  $q$  is a real number. Finally, in order to have dynamical fermions, one adds

$$\mathcal{L}_D = i \bar{\psi}(x) \gamma^\mu \partial_\mu \psi(x)$$

and we end with a gauge invariant total Lagrangian  $\mathcal{L} = \mathcal{L}_M + \mathcal{L}_D + \mathcal{L}_{int}$ , if  $q \equiv e$ . We can introduce the usual covariant derivative  $D_\mu = \partial_\mu - e A_\mu(x)$  that allows one to write the well known minimal electromagnetic coupling

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \bar{\psi}(x) \gamma^\mu D_\mu \psi(x) \quad (25)$$

Certainly, the same procedure, under the same requirements can easily be done for the non-abelian case.

Summarizing this section we remember that the usual way of building a gauge theory starts promoting a global unitary symmetry of the matter Lagrangian to a local one and this requires the inclusion of a gauge field that introduces an unavoidable, but wanted, interaction. The upsidedown way starts from the pure gauge theory that is coupled to matter by imposing Lorentz and gauge invariance in constructing the matter current.

Finally notice that the upsidedown way goes parallel to the geometrical approach that starts defining a connection in a principal fiber bundle and introduce matter fields as sections in the associated vector bundle.

## 9 Weak Interactions

Let us remember that the weak interaction is responsible for i) nuclear  $\beta$ -decay; ii) many hadron decays; iii) everything induced by neutrinos. They are characterized by an effective coupling  $G_F \simeq 10^{-5}/m_p \simeq 1.4 \cdot 10^{-49} \text{ erg cm}^3$  and have the property of breaking the symmetry of parity ( $P$ ). They also suggested (to Pauli) the existence of neutrino.

The first phenomenological analysis due to Fermi was inspired by a comparison between the electrodynamical vertex  $e \rightarrow e + \gamma$  and neutron beta decay  $n \rightarrow p + e + \bar{\nu}_e$ . It ends with the Lagrangian corresponding to a point interaction

$$\mathcal{L}_F = G_F J_{np}^\mu(x) \cdot J_{\mu e \bar{\nu}}(x) = G_F [\bar{\psi}_p(x) \gamma^\mu \psi_n(x)] [\bar{\psi}_e(x) \gamma_\mu \psi_{\nu_e}(x)]$$

This proposal is parity conserving, but after Wu experiment to test the Lee and Yang prediction of the violation of parity in weak interactions, this Lagrangian was modified changing from a vector like interaction to a vector minus axial interaction, namely with currents of the form

$$J^\mu(x) = \bar{\psi}_f \gamma^\mu (1 - \gamma^5) \psi_{f'}(x)$$

This  $V - A$  property of the weak currents, implies that the neutrino is left handed.

It is necessary to be aware that the Fermi proposal should be considered as a phenomenological effective Lagrangian because it works well as a first order approximation and fails at higher orders and is an example of a non-renormalizable theory. It becomes harmful at the so-called limit of unitarity, around  $300 \text{ GeV}$ .

Consequently one tries to improve the treatment of weak interactions taking inspiration in QED. To this end one looks for a theory with a dimensionless coupling constant and a carrier of the interaction playing the role of the photon in QED.

The resulting model is the Intermediate Boson Theory that includes two charged vector bosons mediators of the weak interaction,  $W^\pm$ . This is done with the constraint that the Fermi theory is the limit of low energy ( $q^2 \ll M_W^2$ ). Here also a problem with convergence appears because the very short range of weak interaction requires that the vector bosons must be massive (of the order of  $80 \text{ GeV}$ ). Consequently there are contributions including the massive propagator of the  $W^\pm$  that diverge for large momenta and the intermediate boson theory is again only a phenomenological model valid below its own limit of unitarity.

## 10 Latent Symmetry

We present now the way to the *electroweak* theory that allows one to deal simultaneously with both interactions: electromagnetic plus weak. The way to the theory is via gauge theories as the ones we have discussed previously. The problem to be solved is the necessary mass the mediators of weak interaction must have. This is in conflict with gauge symmetry that does not allow a mass term for the gauge fields. In any case, there is a way out of this conundrum, the possibility of having the symmetry in a latent version. This is no more than to have the symmetry realized à la Nambu-Goldstone. The symmetry is present in the Lagrangian but is not respected by the vacuum expectation value of the fields. This fact has non-trivial consequences in systems of infinite extension.

We start the discussion of this topic by presenting a simple mechanical example. This is the case of the bent rod.

Consider a cylindrical rod along the  $z$ -axis, charged with a force  $F$  along its axis. The system is obviously symmetric under rotations around the  $z$ -axis. The system has the symmetric solution  $x = y = 0$  as soon as the force  $F$  is sufficiently small. In fact, there is a critical force  $F_{cr}$  given by  $F_{cr} = \pi^2 EI/\rho^2$  where  $E$  is the modulus of elasticity,  $I$ , the moment of inertia of the rod and  $\rho$  its density. When  $F > F_{cr}$ , the rod is bent, given rise to an asymmetric solution from a symmetric equation of motion. Certainly one cannot predict the direction in the  $(x, y)$  plane where the rod is going to bend. However, a symmetry

transformation on an asymmetric solution goes into another asymmetric solution. The symmetry is *latent*. In a language nearer to our field theory, we say that the ground state (the vacuum) is degenerate. The symmetry is said to be *spontaneously broken*, it is realized à la Nambu-Goldstone.

Another nice example is the Heisenberg model for a magnet based on the Hamiltonian with exchange interaction

$$H = K \sum_{\langle ij \rangle} \vec{s}_i \cdot \vec{s}_j$$

that clearly is rotational,  $O(3)$ , invariant. This model presents, for a temperature below a critical one ( $T < T_c$ ) a breaking of the  $O(3)$  symmetry because elementary magnetic moments (spins if you want) tend to align. The well known ferromagnetic transition. In any case, the symmetry is maintained latent. In fact (if you can eliminate the effect of the magnetic field of Earth, or others), there is no preferred direction for the alignment. Each possible direction is connected with any other by an  $O(3)$  transformation. Certainly, to this happens it is necessary to be in the thermodynamical limit, a system of infinite extension, infinite degeneracy of the angular momenta. The Goldstone mode of this spontaneously broken symmetry is the spin wave.

We go now towards the Higgs model. Let us consider first a model with:

### Continuous global latent symmetry

This model is also called Goldstone model and is defined by the complex scalar field Lagrangian

$$\mathcal{L}(\phi, \partial_\mu \phi) = -\partial_\mu \phi^\dagger(x) \partial^\mu \phi(x) - \lambda [\phi^\dagger(x) \phi(x) - v]^2 \quad (26)$$

that is invariant under a  $U(1)$  global transformation:  $\phi(x) \rightarrow \phi'(x) = e^{i\alpha} \phi(x)$  with  $\alpha$  constant. The behavior of the interaction  $V(\phi^\dagger, \phi) = \lambda [\phi^\dagger(x) \phi(x) - v]^2$  depends on the sign of  $v$ .

case i)  $v < 0$

consequently,  $V$  has one minimum at  $\phi = 0$  and the  $U(1)$  symmetry is Wigner-Weyl realized.

case ii)  $v > 0$

now  $V$  has an infinite number of minima, defined by  $\phi^\dagger \phi = v$  or, in terms of the real  $\phi_1$  and the imaginary part  $\phi_2$  of  $\phi$ : the circle of minima  $\phi_1^2 + \phi_2^2 = v$ . The picture of the interaction energy looks like a Mexican hat or a "culo de botella". Then we have a Nambu-Goldstone realization of the  $U(1)$  symmetry. The symmetry is latent and there is a Goldstone boson present ( $U(1)$  has one generator).

Clearly, in this case it has no sense to develop a perturbation theory around  $\phi = 0$  (an unstable point). One has to develop around some point  $\phi = \sqrt{v} e^{i\theta}$ , with  $\theta$  arbitrary due to the non-uniqueness of vacuum. One can take  $\theta = 0$  and define the shift

$$\phi(x) = \sqrt{v} + \chi(x) \Rightarrow \langle 0 | \chi(x) | 0 \rangle = 0 \quad (27)$$

It is worth writing the interaction in terms of  $\chi$

$$V = \lambda v (\chi + \chi^\dagger)^2 + 2\lambda \sqrt{v} \chi^\dagger \chi (\chi + \chi^\dagger) + \lambda (\chi \chi^\dagger)^2$$

that when expressed in terms of  $\chi_1 = (\chi + \chi^\dagger)/\sqrt{2}$  and  $\chi_2 = i(\chi^\dagger - \chi)/\sqrt{2}$  shows that there is a massive degree of freedom  $\chi_1$  with mass  $m_1 = 4\lambda v$  and a massless one,  $\chi_2$  which is the Goldstone boson that is present because the symmetry was spontaneously broken. One can say also that this shows that the  $U(1)$  symmetry is not more present in the spectrum.

We can still get more insight on the model properties by performing the change of variables:  $\chi_1 = \rho \cos \theta$  and  $\chi_2 = \rho \sin \theta$ . It is clear that a  $U(1)$  transformation implies the changes:  $\rho \rightarrow \rho$  and  $\theta \rightarrow \theta + \alpha$ . After the displacement  $\rho' = \rho - \sqrt{v}$ , the Lagrangian reads

$$\mathcal{L} = \frac{1}{2} (\partial_\mu \rho')^2 + \frac{1}{2} (\rho' + \sqrt{v})^2 (\partial_\mu \theta)^2 - v(\rho' + \sqrt{v})$$

and shows that  $\rho$  is a radial excitation (clearly seen on the Mexican hat) and  $\theta$  corresponds to the movement around the circle without energy consumption  $\hbar\omega = 0$ , certainly the massless Goldstone boson excitation.

### Local global latent symmetry: Higgs model

Whenever a local symmetry is Nambu-Goldstone realized, the should be Goldstone boson disappear and the gauge field acquires an effective mass. In other words, the Goldstone boson degrees of freedom are transferred to the longitudinal polarization massive vector bosons have. This is the so called Higgs mechanism.

We present the mechanism for the simple complex scalar field

$$\mathcal{L}(\phi, \partial_\mu \phi) = -\partial_\mu \phi^\dagger(x) \partial^\mu \phi(x) - \lambda [\phi^\dagger(x) \phi(x) - v]^2$$

that, as we saw before, when  $v > 0$  the  $U(1)$  symmetry is latent. Now consider that the symmetry is local. Consequently, one has to replace the derivatives by covariant ones, namely

$$\partial_\mu \rightarrow \partial_\mu - ig a_\mu(x)$$

where  $a_\mu$  is the massless gauge field. It is also necessary to include the kinetic term of this field, to write

$$\mathcal{L}(\phi, \partial_\mu \phi) = -(D_\mu \phi)^\dagger D^\mu \phi - \lambda [\phi^\dagger \phi - v]^2 - \frac{1}{4} F^{\mu\nu} F_{\mu\nu}$$

this Lagrangian includes several interactions. Among them we explicitly show

$$\begin{aligned} \mathcal{L}^{(1)} &= g a_\mu (\phi^\dagger \partial^\mu \phi - \phi \partial^\mu \phi^\dagger) \\ \mathcal{L}^{(2)} &= -g^2 a_\mu a^\mu \phi^\dagger \phi \end{aligned}$$

As the symmetry is spontaneously broken, one has to shift to the field  $\chi$  with zero vacuum expectation value as in Eq.(27). From the second expression above,  $\mathcal{L}^{(2)}$ , it results

$$\mathcal{L}_M = -g^2 v a_\mu a^\mu = -\frac{1}{2} m_a^2 a_\mu a^\mu \quad (28)$$

the gauge field  $a_\mu$  acquires the mass

$$m_a = g \sqrt{\frac{v}{2}}$$

What it is important to notice is the fact that during this process of given mass to the gauge field, the symmetry was not completely lost because as we have said before, it remains latent. Moreover, it is easy to prove that during the process there is a conservation of degrees of freedom. In fact one goes from  $(2 + 2)$  original degrees of freedom (2 of the complex scalar field and 2 from the massless gauge field) to  $(3 + 1)$  degrees of freedom (3 of the now massive gauge field and 1 remaining of the scalar field). The should be Goldstone boson is said to be gauged away. The exchange of degrees of freedom is not a violation of the Goldstone theorem. In gauge theories the Goldstone theorem can be overcome by choosing either the unitary gauge where the initial symmetry is destroyed or a covariant gauge in which the Goldstone theorem is valid but the Goldstone bosons become unphysical particles, they uncouple from the remaining theory. There is another point to be analyzed. Namely, the fact that once the vector field acquires mass, its propagator gains a term that makes divergent the perturbative contributions, as it was discussed when the Intermediate Boson Model was presented. However, in the present case related to the Higgs mechanism there is a gauge equivalence between the original gauge invariant renormalizable Lagrangian and the Lagrangian obtained after the action of the Higgs mechanism. This last one is useful for practical purposes, while the original one, for formal (renormalization) purposes. For this reason a Yang-Mills theory with latent symmetry, with spontaneously broken symmetry, is renormalizable. It maintains sufficient Ward-like identities as to guarantee this property.

### 11 3!: The Standard Model

As it was mentioned before, the 3! model includes the local symmetry under the gauge group  $SU_C(3) \otimes SU(2)_L \otimes U(1)_Y$ . The group  $SU(3)_C$  takes into account the strong interaction between quarks mediated by gluons, the corresponding gauge fields. Here we restrict ourselves to the electromagnetic and weak interactions: electroweak, included in 3! as  $2! \equiv SU(2)_L \otimes U(1)_Y$ .

Let us begin by quoting from one of the authors of the electroweak Standard Model, the S. Weinberg's text in the Physical Review Letters article of 1967 where the  $2! = 2 \times 1$  was proposed: "Leptons interact only with photons and with the intermediate bosons which presumably mediate the weak interaction. What can be more natural than linking those spin one bosons in a multiplet of gauge fields?"

The proposal of the Salam-Weinberg-Glashow model starts by noting that it is necessary to solve several difficulties. Namely, the photon  $\gamma$  is massless while the intermediate boson  $W$  has to be massive. Moreover, the value of the electromagnetic coupling is very different from the weak coupling characterized by the Fermi constant  $G_F$ . In connection with this last point, it is clear that the hierarchy of couplings needs the presence of a mixing angle (sinus versus cosinus).

The first question to be answered to build up a gauge theory it is precisely to decide which gauge group to use for connecting with the weak and electromagnetic interactions. The model must include

- Charged weak current,  $J_\mu^\pm$
- Electromagnetic interaction related to a group  $U(1)$  that should be realized à la Wigner-Weyl to guarantee  $m_\gamma = 0$
- Neutral weak current. Notice that its presence allows for a cancellation mechanism that compensate the divergent contributions coming from a massive and charged  $W$

Consequently, the gauge group  $G$  for weak and electromagnetic interactions has to be a 4 parameter group. The proposal is

$$G \equiv SU(2)_L \otimes U(1)_Y \quad (29)$$

or

$$2! = 2 \times 1$$

where the index  $L$  stands for *left* and the index  $Y$  for (*weak*)-*hypercharge*.

Remember that any fermion spin 1/2 field can be decomposed into left and right components, namely

$$\psi_{L,R} = \frac{1 \mp \gamma_5}{2} \psi$$

The index  $L$  in  $SU(2)$  is because  $\psi_L$  behaves differently of  $\psi_R$  under the transformations of this group. Then we can say that the 2!-model is a chiral theory: distinguishes between left and right.

We also remember that the charged weak current must have a  $(V - A)$  structure that includes a  $(1 - \gamma_5)$  factor. This is why only  $\psi_L$  enters the play. In other words, Fermi phenomenological proposal plus Parity violation implies

$$J_\mu^{cc} = \bar{\psi}_i \gamma_\mu (1 - \gamma_5) \psi_j \equiv 2 \psi_{i,L} \gamma_\mu \psi_{j,L}$$

with  $(\psi_i, \psi_j) = (\nu_e, e), (u, d), \dots$

The simplest representation of  $SU(2)_L$  is a doublet

$$\begin{pmatrix} \psi_i \\ \psi_j \end{pmatrix}$$

Notice that  $SU(2)$  has three generators, say  $T_1, T_2, T_3$ . With  $T_1$  and  $T_2$  one can build  $T_\pm$  related to the charged current. But  $T_3$  cannot be identified with the electric charge  $Q$ , because the electromagnetic charge is pure vector and not  $V - A$ . For this reason the  $G$  group is  $SU(2)_L \otimes U(1)_Y$ .

Which is this  $U(1)_Y$  ? From the currents

$$J_\mu^{e.m.} = q_i \bar{\Psi} \gamma_\mu \Psi \Rightarrow Q = \int d^3x J_0^{e.m.}$$

$$J_\mu^3 = \bar{\Psi}_L \gamma_\mu \frac{\tau_3}{2} \Psi_L \Rightarrow T_3 = \int d^3x J_0^3$$

Then, both members of any doublet as  $(\nu_{e,L}, e_L)$  have the same quantum number ( $Y = Q - T_3$ ), the same hypercharge.  $Y = Q - T_3$  commutes with the generators of  $SU(2)$  and defines the group  $U(1)$  of 2!.

### Matter under $SU(2)_L \otimes U(1)$

In each generation of quarks and leptons, the  $G$  quantum numbers are repeated. Let us consider, as an example, the first generation:  $(e, \nu_e, u, d)$ . The doublets of  $SU(2)_L$  are  $(\nu_e, e)_L$  and  $(u, d)_L$ . As we already said, this election guarantees  $(V - A)$  and the index  $L$  is well understood.

The good assignments of  $SU(2)_L$  and  $U(1)_Y$  values for the first generation is.

	$(\nu_e, e)_L$	$e_R$	$(u, d)_L$	$u_R$	$d_R$
$SU(2)_L$	2	1	2	1	1
$U(1)_Y$	-1/2	-1	1/6	2/3	-1/3

and as was stated above, this quantum numbers repeat for the second and the third generations. In any case, the desired electric charge of fermions verify the relation

$$Q = I_3^W + Y_W$$

where the index  $W$  was introduced just to remember that  $SU(2)_L$  is in some sense the weak version of isospin while  $U(1)_Y$  is seen as the weak hypercharge.

### Building the Lagrangian

As it was mentioned before, the left handed components of each generation of leptons and quarks are assigned into doublets of the fundamental representation of  $SU(2)_L$ . In order to simplify the expressions, let us introduce the notation for the doublets

$$L_\alpha(x) = \frac{(1 - \gamma_5)}{2} \begin{pmatrix} i_\alpha(x) \\ j_\alpha(x) \end{pmatrix}$$

where

$$i_\alpha(x) \equiv \nu_e(x); \nu_\mu(x); \nu_\tau(x)$$

$$\equiv u(x); c(x); t(x)$$

and

$$j_\alpha(x) \equiv e(x); \mu(x); \tau(x)$$

$$\equiv d(x); s(x); b(x)$$

and for the singlets ;

$$R_\alpha(x) = \frac{(1 - \gamma_5)}{2} k_\alpha(x)$$

with  $k_\alpha \equiv u(x); c(x); t(x); e(x); \mu(x); \tau(x); d(x); s(x); b(x)$

Now we go to the four gauge vector bosons needed to guarantee the local gauge invariance under  $SU(2)_L \otimes U(1)_Y$ . There are 3 corresponding to the weak isospin:  $W_\mu^1(x)$ ,  $W_\mu^2(x)$  and  $W_\mu^3(x)$  and 1 corresponding to weak hypercharge:  $B_\mu(x)$ . In writing the corresponding covariant derivatives one introduces the couplings  $g_2$  for  $SU(2)_L$  and  $g_1$  for  $U(1)_Y$ . The gauge invariant Lagrangian is

$$\mathcal{L}_2 = -\frac{1}{4} \vec{F}_{\mu\nu}(x) \vec{F}^{\mu\nu}(x) - \frac{1}{4} B_{\mu\nu}(x) B^{\mu\nu}(x)$$

$$\begin{aligned}
 & +\iota \sum_{l=e,\mu,\tau} \bar{L}_l(x) \gamma^\mu D_\mu L_l(x) + \iota \sum_{l=e,\mu,\tau} \bar{R}_l(x) \gamma^\mu D_\mu R_l(x) \\
 & +\iota \sum_{q=(u,d),(c,s),(t,b)} \bar{L}_q(x) \gamma^\mu D_\mu L_q(x) + \iota \sum_{q^\dagger=u,c,t} \bar{R}_{q^\dagger}(x) \gamma^\mu D_\mu R_{q^\dagger}(x) \\
 & +\iota \sum_{q^\dagger=d,s,b} \bar{R}_{q^\dagger}(x) \gamma^\mu D_\mu R_{q^\dagger}(x)
 \end{aligned} \tag{30}$$

The corresponding Yang-Mills strength tensors are

$$\vec{F}_{\mu\nu}(x) = \partial_\mu \vec{W}_\nu(x) - \partial_\nu \vec{W}_\mu(x) + g_2 \vec{W}_\mu(x) \wedge \vec{W}_\nu(x) \tag{31}$$

$$B_{\mu\nu}(x) = \partial_\mu B_\nu(x) - \partial_\nu B_\mu(x) \tag{32}$$

These expressions clearly show that whenever the gauge symmetry is related to a non-abelian group, there are self interactions of gauge fields. This is an important difference with QED because the photon is the gauge field of the abelian symmetry  $U(1)$ .

Here we summarize the covariant derivatives. First for LEPTONS:

$$(for L_l(x) :) D_\mu \equiv \partial_\mu + \iota g_2 \frac{\vec{\tau}}{2} \cdot \vec{W}_\mu(x) - \iota \frac{1}{2} g_1 B_\mu(x) \tag{33}$$

$$(for R_l(x) :) D_\mu \equiv \partial_\mu - \iota g_1 B_\mu(x) \tag{34}$$

and then for QUARKS

$$(for L_{q^\dagger}(x) :) D_\mu \equiv \partial_\mu + \iota g_2 \frac{\vec{\tau}}{2} \cdot \vec{W}_\mu(x) + \iota \frac{1}{6} g_1 B_\mu(x) \tag{35}$$

$$(for R_{q^\dagger}(x) :) D_\mu \equiv \partial_\mu + \iota \frac{2}{3} g_1 B_\mu(x) \tag{36}$$

$$(for R_{q^\dagger}(x) :) D_\mu \equiv \partial_\mu - \iota \frac{1}{3} g_1 B_\mu(x) \tag{37}$$

It remains the problem of assigning masses to the massive vector bosons to be identified with the mediators of weak interaction and also to give masses to the massive fermions. Remember that for the first ones, gauge symmetry does not allow for a mass term of gauge fields, and for fermions, a typical Dirac mass term

$$\mathcal{L}_{mass} = -m \bar{\Psi} \Psi = -m (\bar{\Psi}_L \Psi_R + \bar{\Psi}_R \Psi_L)$$

is unacceptable because  $\Psi_L$  is a  $SU(2)_L$  doublet while  $\Psi_R$  is a singlet.

Certainly, the solution is provided by latent symmetry and the Higgs mechanism. In this way one gets massive gauge fields. But we have to remember that and the end of the play, one massless photon have to remain. In other words, the spontaneous breaking of the symmetry should go from  $SU(2)_L \otimes U(1)_Y$  to a  $U(1)_{e.m.}$  realized a la Wigner-Weyl, because the photon is strictly massless. The mechanism must guarantee that we go from 4 massless bosons to 3 massive and 1 massless.

The photon, the massless one, is represented by the field  $A^\mu$  that results as the neutral combination of  $B^\mu$  and the third component of the triplet  $\vec{W}^\mu$ ,  $W_3^\mu$ . Correspondingly, the orthogonal combination to  $A^\mu$  is the weak neutral boson  $Z^\mu$ . These combinations are measured by a mixing angle  $\theta_W$ , called Weinberg angle. The good combinations reads

$$A_\mu = \sin \theta_W W_{\mu,3} + \cos \theta_W B_\mu \tag{38}$$

$$Z_\mu = \cos \theta_W W_{\mu,3} - \sin \theta_W B_\mu \tag{39}$$

On the other hand, the charged weak bosons are the combinations

$$W_\pm^\mu = \frac{1}{\sqrt{2}} (W_1^\mu \mp \iota W_2^\mu) \tag{40}$$

Now we can explicitly write the weak and electromagnetic interactions contained in the Lagrangian (30)

$$\begin{aligned}\mathcal{L} = & \frac{g_2}{2\sqrt{2}} [J_-^\mu W_{\mu,+} + J_+^\mu W_{\mu,-}] \\ & + [(g_2 \cos \theta_W + g_1 \sin \theta_W) J_3^\mu - g_1 \sin \theta_W J_{em}^\mu] Z_\mu \\ & + [(g_1 \cos \theta_W J_{em}^\mu + (g_1 \cos \theta_W - g_2 \sin \theta_W) J_3^\mu] A_\mu\end{aligned}$$

where  $J_\pm^\mu = 2(J_1^\mu \mp iJ_2^\mu)$  and  $J_{em}^\mu = J_3^\mu + J_Y^\mu$  while  $J_1$  and  $J_2$  the currents containing  $\tau_1$  and  $\tau_2$  respectively.

Now, in order to reobtain the standard electromagnetic interaction  $\mathcal{L}_{int}^{em} = e J_{em}^\mu A_\mu$ , the identification

$$e = g_1 \cos \theta_W = g_2 \sin \theta_W \quad (41)$$

is required. Consequently, the interaction Lagrangian above gets the form

$$\begin{aligned}\mathcal{L} = & \frac{e}{2\sqrt{2} \sin \theta_W} [J_-^\mu W_{\mu,+} + J_+^\mu W_{\mu,-}] \\ & + \frac{e}{2 \cos \theta_W \sin \theta_W} Z^\mu J_{\mu,NC} + e A^\mu J_{\mu,em}\end{aligned} \quad (42)$$

where one defined the neutral current  $J_{NC}^\mu = 2(J_3^\mu - \sin^2 \theta_W J_{em}^\mu)$ .

To summarize, we remark that the electromagnetic interaction is measured by  $e$  as it should be, the charged weak interactions coupling is  $e/(2\sqrt{2} \sin \theta_W)$  while the neutral weak interaction one is  $e/(2 \cos \theta_W \sin \theta_W)$ . Now the Feynman rules can easily be obtained and from them, any perturbative calculation as cross sections, decay rates, etc., follow.

It remains to discuss the problem of masses of both, the weak gauge bosons and the massive fermions. As it was stated above, the Higgs mechanism is in order. The idea is to put the  $2_1$ -symmetry in a latent version but being careful to keep a  $U(1)$ , corresponding to electromagnetism, realized à la Wigner-Weyl. It is necessary to introduce scalar fields being non-trivial under  $SU(2)_L \otimes U(1)_Y$ . The simplest election is a  $SU(2)_L$  doublet

$$\Phi = \begin{pmatrix} \phi^0 \\ \phi^- \end{pmatrix} \quad (43)$$

with both  $\phi^0$  and  $\phi^-$  complex fields. Clearly, the field  $\Phi$  is a  $SU(2)_L$  doublet and one chooses  $Y_\Phi = -1/2$ . The latent character of the symmetry is obtained by the inclusion of the interaction

$$V(\Phi^\dagger \Phi) = \frac{\lambda}{4} (\Phi^\dagger \Phi - v^2)^2 \quad (44)$$

that implies that

$$\langle 0|\Phi|0\rangle = \begin{pmatrix} v \\ 0 \end{pmatrix}$$

and as we show below, this election guarantees the  $U(1)_{em}$ . In fact, one can perform a test of coherence of the spontaneous symmetry breaking just proposed. It is easy to compute the effect of the generators of  $SU(2)_L$  and the one of  $U(1)_Y$  on the vacuum expectation value of  $\Phi$ , namely

$$\begin{aligned}G_1 \langle 0|\Phi|0\rangle &= \frac{1}{2} \begin{pmatrix} 0 \\ v \end{pmatrix} \neq 0 \\ G_2 \langle 0|\Phi|0\rangle &= \frac{1}{2} \begin{pmatrix} 0 \\ iv \end{pmatrix} \neq 0 \\ G_3 \langle 0|\Phi|0\rangle &= \frac{1}{2} \begin{pmatrix} v \\ 0 \end{pmatrix} \neq 0\end{aligned}$$



$$G_Y \langle 0|\Phi|0\rangle = -\frac{1}{2} \begin{pmatrix} v \\ 0 \end{pmatrix} \neq 0$$

that at first sight seems to leave the complete symmetry latent. However, the electromagnetic generator of the symmetry  $U(1)$ , that is  $Q = I_3 + Y$  certainly verifies

$$Q \langle 0|\Phi|0\rangle = 0$$

that characterizes a symmetry realized à la Wigner-Weyl. The zero mass for the photon is guaranteed.

There are three would be Goldstone bosons in (43) that decouple from the theory and provide the longitudinal polarizations that the massive  $W^+$ ,  $W^-$  and  $Z$  require.

The  $\Phi$  original Lagrangian reads

$$\mathcal{L} = -\frac{1}{2} (D_\mu \Phi)^\dagger D^\mu \Phi - V(\Phi^\dagger \Phi)$$

with

$$D_\mu \Phi = \left( \partial_\mu + i \frac{g_1}{2} Y_\mu - i g_2 \frac{\vec{\tau}}{2} \cdot \vec{W}_\mu \right)$$

now one performs the standard shift  $\Phi(x) = H(x) + v$  with  $\langle 0|H(x)|0\rangle = 0$ . The  $H(x)$  is the Higgs boson field with a mass, that can be obtained from the Lagrangian after the shift, being

$$m_H^2 = 2 \lambda v^2$$

The covariant derivative of the Higgs field gives rise to the gauge fields-Higgs interactions, that after the shift allows one to read the effective mass terms of  $W^\pm$  and  $Z$ . In fact

$$\begin{aligned} \mathcal{L}_{GaugeMass} &= -\frac{1}{4} (g_2 v)^2 W_+^\mu W_{\mu,-} \\ &\quad -\frac{1}{8} \left( \frac{g_2 v}{\cos \theta_W} \right)^2 Z^\mu Z_\mu, \end{aligned}$$

and the masses of the three massive weak gauge bosons are

$$M_{W^\pm}^2 = \frac{g_2^2 v^2}{4} \tag{45}$$

and

$$M_Z^2 = \frac{g_2^2 v^2}{4 \cos^2 \theta_W} = \frac{M_W^2}{\cos^2 \theta_W} \tag{46}$$

Notice that the experimental determination of  $M_W$  allows to fix, once the coupling is determined, the value of  $v \approx 246 \text{ GeV}$ , the scale of breaking of the symmetry. This, together with the measurement of  $M_Z$ , fixes

$$\sin^2 \theta_W = 1 - \frac{M_W^2}{M_Z^2}$$

that results  $\sin^2 \theta_W = 0.2234$  in the so called on-shell scheme.

The usual way of obtaining the value of the gauge couplings is by comparison with the low energy weak interaction phenomenology. In this regime, corresponding to  $q^2 \ll M_W^2$ , the Fermi model works satisfactorily. Moreover, under this condition, the matrix elements contributing to a weak process, say the  $\beta$ -decay, are identified with the Fermi prediction, as soon as

$$\frac{g_2^2}{8 M_W^2} \equiv \frac{G_F}{\sqrt{2}} \tag{47}$$

Consequently

$$v = \frac{2M_W}{g_2} = \left(\sqrt{2}G_F\right)^{-1/2} \simeq 246 \text{ GeV} \quad (48)$$

It remains to discuss the mechanism to provide mass to the leptons and quarks. One takes profit of the Higgs mechanism and starts introducing a Yukawa type of coupling with the scalar doublet. Let us begin with leptons, remembering that the Standard Model works with massless neutrinos. The Yukawa lagrangian for leptons coupled via  $h_l$  to  $\Phi(x)$  reads

$$\mathcal{L}_Y(x) = - \sum_{l=(\nu_e, e), (\nu_\mu, \mu), (\nu_\tau, \tau)} h_l \bar{R}_l(x) \Phi^\dagger(x) L_l(x) + h.c. \quad (49)$$

Notice that this expression is correct since it implies the product of a doublet times an adjoint doublet times a singlet. Now, once the shift from  $\Phi(x)$  to  $H(x)$  is performed after the spontaneous breaking of the symmetry, it appears a mass like coupling

$$\mathcal{L}_Y(x) = - \sum_{l=e, \mu, \tau} h_l v \bar{\Psi}_l(x) \Psi_l(x) \quad (50)$$

that allows to identify the mass of each lepton

$$m_l = h_l v \quad (51)$$

It should be notice that the couplings  $h_l$  are not fixed by any principle in 2!. They are external parameters.

In the case of quarks, as both members of the doublets are massive, it is necessary to consider two isoscalars

$$\mathcal{L}_Y(x) = - \sum_{i,j} h_{i,j} \bar{L}_{q_i}(x) \Phi(x) R_{q_j}(x) - \tilde{h}_{i,j} \bar{L}_{q_i}(x) \tilde{\Phi}(x) R_{q_j}(x) + h.c. \quad (52)$$

where

$$\tilde{\Phi}(x) = i \tau_2 \Phi^*(x) = \begin{pmatrix} \phi^0(x) \\ -\phi^+(x) \end{pmatrix}^*$$

Consequently, after the spontaneous breaking of the symmetry one has

$$\mathcal{L}_{masses}(x) = -\bar{q}_L^\dagger(x) \mathcal{M}(2/3) q_R^\dagger(x) - \bar{q}_L^\dagger(x) \mathcal{M}(-1/3) q_R^\dagger(x) + h.c.$$

where

$$q_{L,R}^\dagger(x) \equiv \frac{1}{2} (1 \mp \gamma_5) \begin{pmatrix} u(x) \\ c(x) \\ t(x) \end{pmatrix}$$

and

$$q_{L,R}^\dagger(x) \equiv \frac{1}{2} (1 \mp \gamma_5) \begin{pmatrix} d(x) \\ s(x) \\ b(x) \end{pmatrix}$$

Consequently, the  $3 \times 3$  mass matrices are

$$\mathcal{M}(2/3) \equiv \tilde{h}_{ij} v ; i, j = u, c, t$$

and

$$\mathcal{M}(-1/3) \equiv h_{ij} v ; i, j = d, s, b$$

These matrices can be diagonalized via

$$U_L(Q) \mathcal{M}(Q) U_R(Q) = \hat{M}(Q) (\text{diagonal}) ; (Q = \frac{2}{3}, -\frac{1}{3})$$

In this way, the mass Lagrangian results

$$\mathcal{L}_{masses}(x) = \bar{\psi}_L^\dagger(x) \hat{M}(2/3) \psi_R^\dagger(x) + \bar{\psi}_L^\dagger(x) \hat{M}(-1/3) \psi_R^\dagger(x)$$

where the mass eigenstates are

$$\begin{aligned} \psi_{L,R}^\dagger(x) &= U_{L,R}(2/3) q_{L,R}^\dagger(x) \\ \psi_{L,R}^\dagger(x) &= U_{L,R}(-1/3) q_{L,R}^\dagger(x) \end{aligned}$$

The process of diagonalizing the mass matrices has as a consequence that in the charged current interaction sector there is flavor mixing. In fact, the term

$$\bar{q}_L^\dagger(x) \gamma^\mu q_L^\dagger(x) W_\mu^\dagger(x) + h.c.$$

gives rise, after the "rotation" in the process, to

$$\bar{\psi}_L^\dagger(x) \gamma^\mu U_L(2/3) U_L^\dagger(-1/3) \psi_L^\dagger(x) W_\mu^\dagger(x) + h.c.$$

The resulting unitary matrix

$$V \equiv U_L(2/3) U_L^\dagger(-1/3) \tag{53}$$

is called Cabibbo-Kobayashi-Maskawa mixing matrix. When writing the charged weak interaction with  $W$ , that matrix acts on flavor degrees of freedom of charged  $-1/3$  quarks ( $\downarrow$ ). It is a generalization of the Cabibbo matrix that was introduced when only four flavors were known and in order to keep the universality of weak interactions.

This last discussion ends the introduction of 2!, the Standard Model of electroweak interactions. It has an impressive success and a strong predictive power. We recommend the reader to consult, for example,

<http://www-pdg.lbl.gov/2016/reviews/rpp2016-rev-standard-model.pdf>

to have a clear idea of this success of the model and also of the constraints it implies to eventual new physics.

## 12 Finale

This lectures that were presented at the 2017 CERN - Latin-American School of High Energy Physics have to be considered as a mere "opening of a door". A door that should allows us to enter in an almost unbounded territory: the land of Contemporary Physics. Certainly, we have left aside, for time restrictions, many interesting and important topics that has to be studied and understood by the reader, consulting the vast list of references included in the Bibliography.

### Acknowledgments

I would like to thank Kate Ross, Maria Elena Tejada-Yeomans, Nick Ellis and Martijn Mulders, for organizing and achieving an enjoyable and fruitful School. The continuous interest of the participants made me glad all the time. My special thanks to Alejandro Szykman for the very careful reading of these notes and his precise comments that improved them in all aspects.

### Bibliography

- T.P. Cheng and L.F. Li: "Gauge Theory of Elementary Particle Physics", Oxford University Press 1991
- F. Halzen and A.D. Martin: " Quarks and Leptons: An Introductory Course in Modern Particle Physics" John Wiley 1984

- B. Kayser, F. Gibrat-Debu and F. Perrier: "The Physics of Massive Neutrinos", World Scientific 1989
- W.N. Cottingham and D.A. Greenwood: "An Introduction to the Standard Model of Particle Physics", Cambridge University Press 2005
- J.F. Donoghue, E. Golowich and B.R. Holstein: "Dynamics of the Standard Model", Cambridge University Press 1992
- E. Leader and E. Predazzi: "Gauge Theories and the New Physics", Cambridge University Press 1982
- M.E. Peskin and D.V. Schroeder: "An Introduction to Quantum Field Theory", Addison-Wesley 1995
- D.H. Perkins: "Introduction to High Energy Physics", Addison-Wesley 1987
- D.H. Perkins: "Particle Astrophysics", Oxford University Press 2005
- F.J. Yndurain: "The Theory of Quarks and Gluons Interactions" Springer-Verlag 1999
- T. Muta: "Foundations of Quantum Chromodynamics" World Scientific 1987
- A. Pich: "Quantum Chromodynamics" hep-ph/9505231
- M. Herrero: "The Standard Model" hep-ph/9812242
- A. Stocchi: "Current Status of the CKM Matrix and the CP Violation" hep-ph/0405038
- J. D. Lykken: "The Standard Model: Alchemy and Astrology" hep-ph/0609274
- G. Ecker; "Quantum Chromodynamics" hep-ph/0604165
- S. Bethke: "Experimental Tests of Asymptotic Freedom" hep-ex/0606035
- L.J. Dixon: "Hard QCD Processes at Colliders" arXiv:0712.3064 [hep-ph]
- M.H. Seymour: "Quantum Chromodynamics" hep-ph/0505192
- G.F. Giudice: "Theories for de Fermi Scale" arXiv:0710.3294
- S. Willenbrock: "Symmetries of the Standard Model" hep-ph/0410370
- A. Pich: "The Standard Model of Electroweak Interactions" arXiv:0705.4264
- T.G. Rizzo: "Pedagogical Introduction to Extra Dimensions" hep-ph/0409309
- R.D. Peccei: "Neutrino Physics" hep-ph/9906509
- W. Buchmuller, C. Ludeling: "Field Theory and Standard Model" hep-ph/0609174
- C.A. Garcia Canal: "The Private Life of Three Factorial" AFA, 1993
- D. Rainwater: "Searching for the Higgs Boson" hep-ph/0702124
- Nature Insight: "The Large Hadron Collider, Vol 448, 7151, 19 July 2007"
- M.L. Mangano: "The future of Particle Physics" hep-ph/0608198
- H. Murayama: "Outlook: The Next Twenty Years" hep-ph/0312096
- N.E. Mavromatos: "LHC Physics and Cosmology" arXiv:0708.0134
- J.B. Dainton et al.: "Deep Inelastic Electron-Nucleon Scattering at the LHC" hep-ex/0603016
- P. Higgs: "Evading the Goldstone Theorem" Nobel Lecture, 8 December 2013  
[https://www.nobelprize.org/nobel\\_prizes/physics/laureates/2013/higgs-lecture.pdf](https://www.nobelprize.org/nobel_prizes/physics/laureates/2013/higgs-lecture.pdf)
- C. Quigg: "Spontaneous Symmetry Breaking as a Basis of Particle Mass" arXiv:0704.2232
- W. Kilian, P.M. Zerwas: "ILC: Physics Scenarios" hep-ph/0601217
- R. Barbieri: "Ten Lectures on the ElectroWeak Interactions" arXiv:0706.0684
- J. Ellis: "Physics at LHC" hep-ph/0611237
- N. Borghini, U.A. Wiedemann: "Predictions for the LHC Heavy Ion Programme" arXiv:0707.0564
- J. Erler: "Electroweak Physics at LHC" hep-ph/0607323
- J. Grosse-Knetter: "Early Standard Model Physics and Early Discovery Strategy in ATLAS" ATL-





# QCD AND THE PHYSICS OF HADRONIC COLLISIONS

*M. L. Mangano*

Theoretical Physics Department, CERN, Geneva, Switzerland

## Abstract

This series of lectures reviews the basic principles underlying the use of quantum chromodynamics in understanding the structure of high- $Q^2$  processes in high-energy hadronic collisions. Several applications of relevance to the LHC are illustrated.

## Keywords

Lectures; QCD; hadronic collisions; gluon emission; hadronization; jets

## 1 Introduction

The initial state of any LHC collision is formed by a pair of protons. Whether the hard process we are interested in is of electroweak origin (e.g.  $W$  production) or a strong-interaction process (e.g. the production of jets, or the production of a pair of gluinos in Supersymmetry), its description requires the understanding of the structure of the proton. Quantum Chromodynamics (QCD) is the theory that describes the structure of the proton, and is therefore the starting point of any study of LHC physics. QCD is formulated in terms of elementary fields (quarks and gluons), whose interactions obey the principles of a relativistic quantum field theory, with a non-Abelian gauge invariance  $SU(3)$ . To review the emergence of QCD as a theory of strong interactions, analysing the various experimental data and the theoretical ideas available in the years 1960–1973 (see, for example, Refs. [1, 2]), would require more time than I have available. I shall therefore assume that you all know more or less what QCD is: that hadrons are made of quarks, that quarks are spin-1/2, colour-triplet fermions, interacting via the exchange of an octet of spin-1 gluons; I assume you know the concept of running couplings, asymptotic freedom and of confinement. I shall finally assume that you have some familiarity with the fundamental ideas and formalism of quantum field theory (Feynman rules, renormalization, gauge invariance), even though I shall make only very limited use of them.

In these lectures I shall focus on some elementary applications of QCD in high-energy phenomena. The material covered in these lectures includes the following:

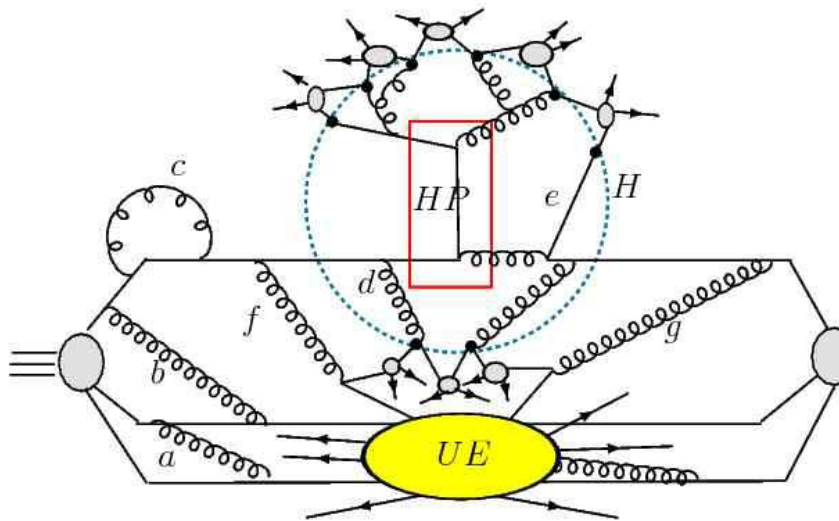
1. The structure of the proton
2. The evolution of final states: from quarks and gluons to hadrons
3. Some key hard processes in hadron–hadron collisions: formalism,  $W/Z$  production, jet production

The treatment will be very elementary, and the emphasis will be on basic and intuitive physics concepts. Given the large number of papers that contributed to the development of the field, it is impossible to provide a complete and fair bibliography. I therefore limit my bibliography to some excellent review books, and to references to some of the key results discussed here. For an excellent description of the early ideas about quarks, the classic reference is Feynman’s book [3]. For a general, but rather formal, introduction to QCD, see, for example, Ref. [4]. For a more modern and pedagogical introduction, in the context of introductory course to field theory, use the excellent book by Peskin [5]. For a general introduction to collider physics, see Ref. [6]. For QCD applications to LEP, Tevatron and LHC, see Ref. [7] and, specifically for the LHC, see Ref. [8]. Explicit calculations, including the nitty-gritty details of next-to-leading-order (NLO) calculations and renormalization, are given in great detail for several concrete cases of interest in Ref. [12]. Many of the ideas used in my lectures are inspired by the very physical perspective presented in Ref. [13]. Papers on specific items can easily be found by consulting the standard hep-th and hep-ph preprint archives.

## 2 QCD and the proton structure at large $Q^2$

The understanding of the structure of the proton at short distances is one of the key ingredients to be able to predict cross-sections for processes involving hadrons in the initial state. All processes in hadronic collisions, even those intrinsically of electroweak nature such as the production of  $W/Z$  bosons or photons, are in fact induced by the quarks and gluons contained inside the hadron. In this lecture I shall introduce some important concepts, such as the notion of partonic densities of the proton, and of parton evolution. These are the essential tools used by theorists to predict production rates for hadronic reactions.

We shall limit ourselves to processes where a proton–(anti)proton pair collides at large centre-of-mass energy ( $\sqrt{S}$ , typically larger than several hundred GeV) and undergoes a very inelastic interaction, with momentum transfers between the participants in excess of several GeV. The outcome of this hard interaction could be the simple scattering at large angle of some of the hadron’s elementary constituents, their annihilation into new massive resonances, or a combination of the two. In all cases the final state consists of a large multiplicity of particles, associated to the evolution of the fragments of the initial hadrons, as well as of the new states produced. As discussed below, the fundamental physical concept that makes the theoretical description of these phenomena possible is ‘factorization’, namely the ability to isolate separate independent phases of the overall collision. These phases are dominated by different dynamics, and the most appropriate techniques can be applied to describe each of them separately. In particular, factorization allows one to decouple the complexity of the proton structure and of the final-state hadron formation from the elementary nature of the perturbative hard interaction among the partonic constituents.



**Fig. 1:** General structure of a hard proton–proton collision

Figure 1 illustrates how this works. As the left proton travels freely before coming into contact with the hadron coming in from the right, its constituent quarks are held together by the constant exchange of virtual gluons (e.g. gluons  $a$  and  $b$  in the picture). These gluons are mostly soft, because any hard exchange would cause the constituent quarks to fly apart, and a second hard exchange would be necessary to reestablish the balance of momentum and keep the proton together. Gluons of high virtuality (gluon  $c$  in the picture) prefer therefore to be reabsorbed by the same quark, within a time inversely proportional to their virtuality, as prescribed by the uncertainty principle. The state of the quark is, however, left unchanged by this process. Altogether this suggests that the global state of the proton, although defined by a complex set of gluon exchanges between quarks, is nevertheless determined by interactions which have a time scale of the order of  $1/m_p$ . When seen in the laboratory frame where the proton is



moving with energy  $\sqrt{S}/2$ , this time is furthermore Lorentz dilated by a factor  $\gamma = \sqrt{S}/2m_p$ . If we disturb a quark with a probe of virtuality  $Q \gg m_p$ , the time frame for this interaction is so short ( $1/Q$ ) that the interactions of the quark with the rest of the proton can be neglected. The struck quark cannot negotiate with its partners a coherent response to the external perturbation: it simply does not have the time to communicate to them that it is being kicked away. On this time scale, only gluons with energy of the order of  $Q$  can be emitted, something which, to happen coherently over the whole proton, is suppressed by powers of  $m_p/Q$  (this suppression characterizes the ‘elastic form factor’ of the proton). In this figure, the hard process is represented by the rectangle labelled HP. In this example a head-on collision with a gluon from the opposite hadron, leads to a  $gg \rightarrow gg$  scattering with a momentum exchange of the order of  $Q$ . This and other possible processes can be calculated from first principles in perturbative QCD.

When the constituent is suddenly deflected, the partons that it had recently radiated cannot be reabsorbed (as happened to gluon  $c$  earlier) because the constituent is no longer there waiting for the partons to come back. This is the case, for example, of the gluon  $d$  emitted by the quark, and of the quark  $e$  from the opposite hadron; the emitted gluon got engaged in the hard interaction. The number of ‘liberated’ partons will depend on the hard scale  $Q$ : the larger the value of  $Q$ , the more sudden the deflection of the struck parton, and the fewer the partons that can reconnect before its departure (typically only partons with virtuality larger than  $Q$ ).

After the hard process, the partons liberated during the evolution prior to the collision and the partons created by the hard collision will themselves emit radiation. The radiation process, governed by perturbative QCD, continues until a low virtuality scale is reached (the boundary region labelled with a dotted line, H, in our figure). To describe this perturbative evolution phase, proper care has to be taken to incorporate quantum coherence effects, which in principle connect the probabilities of radiation off different partons in the event. Once the low virtuality scale is reached, the memory of the hard-process phase has been lost, once again as a result of different time scales in the problem, and the final phase of hadronization takes over. Because of the decoupling from the hard-process phase, the hadronization is assumed to be independent of the initial hard process, and its parametrization, tuned to the observables of some reference process, can then be used in other hard interactions (universality of hadronization). Nearby partons merge into colour-singlet clusters (the grey blobs in fig. 1), which are then decayed phenomenologically into physical hadrons. To complete the picture, we need to understand the evolution of the fragments of the initial hadrons. As shown in the figure, this evolution cannot be entirely independent of what happens in the hard event, because at least colour quantum numbers must be exchanged to guarantee the overall neutrality and conservation of baryon number. In our example, the gluons  $f$  and  $g$ , emitted early on in the perturbative evolution of the initial state, split into  $q\bar{q}$  pairs which are shared between the hadron fragments (whose overall interaction is represented by the oval labelled UE, for Underlying Event) and the clusters resulting from the evolution of the initial state.

The above ideas are embodied in the following factorization formula, which represents the starting point of any theoretical analysis of cross-sections and observables in hadronic collisions:

$$\frac{d\sigma}{dX} = \sum_{j,k} \int_{\hat{X}} f_j(x_1, Q_i) f_k(x_2, Q_i) \frac{d\hat{\sigma}_{jk}(Q_i, Q_f)}{d\hat{X}} F(\hat{X} \rightarrow X; Q_i, Q_f), \quad (1)$$

where:

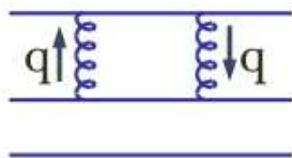
- $X$  is some hadronic observable (e.g. the transverse momentum of a pion, the energy of a jet, the invariant mass of a combination of particles, etc.);
- the sum over  $j$  and  $k$  extends over the partons types inside the colliding hadrons;
- the function  $f_j(x, Q)$  (known as parton distribution function, PDF) represents the number density of parton type  $j$  with momentum fraction  $x$  in a proton probed at a scale  $Q_i$  (more later on the meaning of this scale);

- $\hat{X}$  is a parton-level kinematical variable (e.g. the transverse momentum of a parton from the hard scattering);
- $\hat{\sigma}_{jk}$  is the parton-level cross-section, differential in the variable  $\hat{X}$ ;
- $F(\hat{X} \rightarrow X; Q_i, Q_f)$  is a transition function, weighting the probability that the partonic state defining  $\hat{X}$  gives rise, after hadronization, to the hadronic observable  $X$ ;
- the scales  $Q_i$  and  $Q_f$  correspond to the scales at which we separate the hard, perturbative, process from the initial and final-state evolutions, respectively.

In the rest of this Section I shall cover the above ideas in some more detail. While I shall not provide you with a rigorous proof of the legitimacy of this approach, I shall try to justify it qualitatively to make it sound at least plausible.

## 2.1 The parton densities and their evolution

As mentioned above, the binding forces responsible for the quark confinement are due to the exchange of rather soft gluons. If a quark were to exchange just a single a hard virtual gluon with another quark, the recoil would tend to break the proton apart. It is easy to verify that the exchange of gluons with virtuality larger than  $Q$  is then proportional to some large power of  $m_p/Q$ ,  $m_p$  being the proton mass. Since the gluon coupling constant gets smaller at large  $Q$ , exchange of hard gluons is significantly suppressed<sup>1</sup>. Consider in fact the picture in Fig. 2. The exchange of two gluons is required to ensure that the mo-



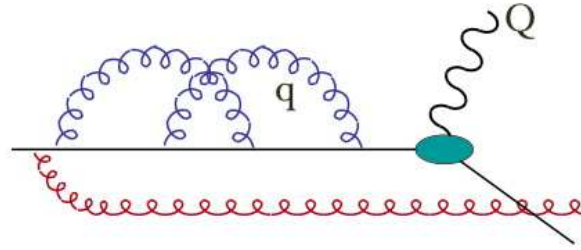
**Fig. 2:** Gluon exchange inside the proton

mentum exchanged after the first gluon emission is returned to the quark, and the proton maintains its structure. The contributions of hard gluons to this process can be approximated by integrating the loop over large momenta:

$$\int_Q \frac{d^4q}{q^6} \sim \frac{1}{Q^2}. \quad (2)$$

At large  $Q$  this contribution is suppressed by powers of  $(m_p/Q)^2$ , where the proton mass  $m_p$  is included as being the only dimensionful quantity available (one could use here the fundamental scale of QCD,  $\Lambda_{QCD}$ , but numerically this is anyway of the order of a GeV). The interactions keeping the proton together are therefore dominated by soft exchanges, with virtuality  $Q$  of the order of  $m_p$ . Owing to Heisenberg's uncertainty principle, the typical time scale of these exchanges is of the order of  $1/m_p$ : this is the time during which fluctuations with virtuality of the order of  $m_p$  can survive. In the laboratory system, where the proton travels with energy  $E$ , this time is Lorentz dilated to  $\tau \sim \gamma/m_p = E/m_p^2$ . If we probe the proton with an off-shell photon, the interaction takes place during the limited lifetime of the virtual photon, which, once more from the uncertainty principle, is given by the inverse of its virtuality. Assuming the virtuality  $Q \gg m_p$ , once the photon gets 'inside' the proton and meets a quark, the struck quark has no time to negotiate a coherent response with the other quarks, because the time scale for it to 'talk' to its partners is too long compared with the duration of the interaction with the photon itself. As a result, the struck quark has no option but to interact with the photon as if it were a free particle. Let us look in more detail at what happens during such a process. In Fig. 3 we see a

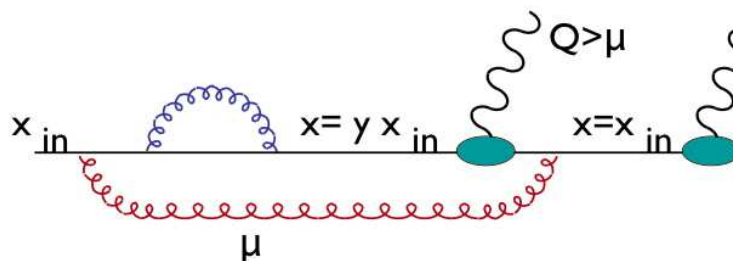
<sup>1</sup>The fact that the coupling decreases at large  $Q$  plays a fundamental role in this argument. Were this not true, the parton picture could not be used!.



**Fig. 3:** Gluon emission at different scales during the approach to a hard collision.

proton as it approaches a hard collision with a photon of virtuality  $Q$ . Gluons emitted at a scale  $q > Q$  have the time to be reabsorbed, since their lifetime is very short. Their contribution to the process can be calculated in perturbative QCD, since the scale is large and in the domain where perturbative calculations are meaningful. Since after being reabsorbed the state of the quark remains the same, their only effect is an overall renormalization of the wave function, and they do not affect the quark density. A gluon emitted at a scale  $q < Q$ , however, has a lifetime longer than the time it takes for the quark to interact with the photon, and by the time it tries to reconnect to its parent quark, the quark has been kicked away by the photon, and is no longer there. Since the gluon has taken away some of the quark momentum, the momentum fraction  $x$  of the quark as it enters the interaction with the photon is different than the momentum it had before, and therefore its density  $f(x)$  is affected. Furthermore, when the scale  $q$  is of the order of 1 GeV the state of the quark is not calculable in perturbative QCD. This state depends on the internal wave function of the proton, which perturbative QCD cannot easily predict. We can, however, say that the wave function of the proton, and therefore the state of the ‘free’ quark, are determined by the dynamics of the soft-gluon exchanges inside the proton itself. Since the time scale of this dynamics is long relative to the time scale of the photon–quark interaction, we can safely argue that the photon sees to good approximation a static snapshot of the proton’s inner guts. In other words, the state of the quark had been prepared long before the photon arrived. This also suggests that the state of the quark will not depend on the precise nature of the external probe, provided the time scale of the hard interaction is very short compared to the time it would take for the quark to readjust itself. As a result, if we could perform some measurement of the quark state using, say, a virtual-photon probe, we could then use this knowledge on the state of the quark to perform predictions for the interaction of the proton with any other probe (e.g. a virtual  $W$  or even a gluon from an opposite beam of hadrons). This is the essence of the universality of the parton distributions.

The above picture leads to an important observation. It appears in fact that the distinction between which gluons are reabsorbed and which ones are not depends on the scale  $Q$  of the hard probe. As a result, the parton density  $f(x)$  appears to depend on  $Q$ . This is illustrated in Fig. 4. The gluon emitted



**Fig. 4:** Scale dependence of the gluon emission during a hard collision

at a scale  $\mu$  has a lifetime short enough to be reabsorbed before a collision with a photon of virtuality

$Q < \mu$ , but too long for a photon of virtuality  $Q > \mu$ . When going from  $\mu$  to  $Q$ , therefore, the partonic density  $f(x)$  changes. We can easily describe this variation as follows:

$$f(x, Q) = f(x, \mu) + \int_x^1 dx_{in} f(x_{in}, \mu) \int_\mu^Q dq^2 \int_0^1 dy \mathcal{P}(y, q^2) \delta(x - yx_{in}), \quad (3)$$

Here we obtain the density at the scale  $Q$  by adding to  $f(x)$  at the scale  $\mu$  (which we label as  $f(x, \mu)$ ) all the quarks with momentum  $x_{in} > x$  that retain a proton-momentum fraction  $x = y/x_{in}$  by emitting a gluon. The function  $\mathcal{P}(y, Q^2)$  describes the ‘probability’ that the quark emits a gluon at a scale  $Q$ , keeping a fraction  $y$  of its momentum. This function does not depend on the details of the hard process, it simply describes the radiation of a free quark subject to an interaction with virtuality  $Q$ . Since  $f(x, Q)$  does not depend upon  $\mu$  ( $\mu$  is just used as a reference scale to construct our argument), the total derivative of the right-hand side w.r.t.  $\mu$  should vanish, leading to the following equation:

$$\frac{df(x, Q)}{d\mu^2} = 0 \quad \Rightarrow \quad \frac{df(x, \mu)}{d\mu^2} = \int_x^1 \frac{dy}{y} f(y, \mu) \mathcal{P}(x/y, \mu^2). \quad (4)$$

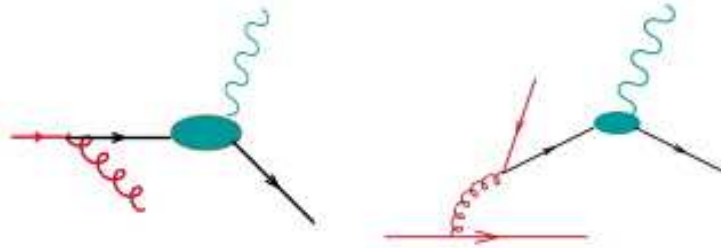
With additional considerations and explicit calculations, one can prove that

$$\mathcal{P}(x, Q^2) = \frac{\alpha_s}{2\pi} \frac{1}{Q^2} P(x) \quad (5)$$

from which the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) equation follows [14–16]:

$$\frac{df(x, \mu)}{d \log \mu^2} = \frac{\alpha_s}{2\pi} \int_x^1 \frac{dy}{y} f(y, \mu) P_{qq}(x/y). \quad (6)$$

The so-called *splitting function*  $P_{qq}(x)$  can be calculated in perturbative QCD. The subscript  $qq$  is a labelling convention indicating that  $x$  refers to the momentum fraction retained by a quark after emission of a gluon.



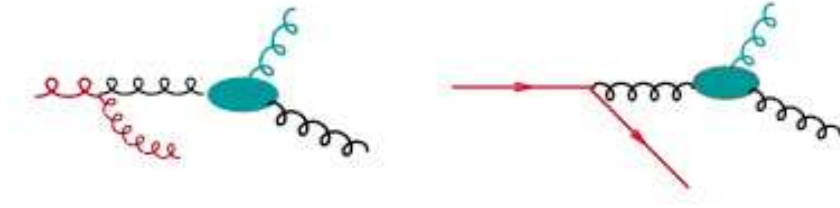
**Fig. 5:** The processes leading to the evolution of the quark density

More generally, one should consider additional processes. For example, one should include cases in which the quark interacting with the photon comes from the splitting of a gluon. This is shown in Fig. 5: the left diagram is the one we considered above; the right diagram corresponds to processes where an emitted gluon has the time to split into a  $q\bar{q}$  pair, and it is one of these quarks which interacts with the photon. The overall evolution equation, including the effect of gluon splitting, is given by

$$\frac{dq(x, Q)}{dt} = \frac{\alpha_s}{2\pi} \int_x^1 \frac{dy}{y} \left[ q(y, Q) P_{qq}\left(\frac{x}{y}\right) + g(y, Q) P_{gq}\left(\frac{x}{y}\right) \right], \quad (7)$$

where  $t = \log Q^2$ . For external probes that couple to gluons (for example an external gluon, coming e.g. from an incoming proton), we have a similar evolution of the gluon density (see Fig. 6):

$$\frac{dg(x, Q)}{dt} = \frac{\alpha_s}{2\pi} \int_x^1 \frac{dy}{y} \left[ g(y, Q) P_{gg}\left(\frac{x}{y}\right) + \sum_{q, \bar{q}} q(y, Q) P_{gq}\left(\frac{x}{y}\right) \right]. \quad (8)$$



**Fig. 6:** The processes leading to the evolution of the gluon density

The explicit calculation of the splitting functions  $P_{ij}(x)$  (see, for example, Ref. [12]) gives then the following expressions<sup>2</sup>:

$$P_{qq}(x) = P_{gq}(1-x) = C_F \frac{1+x^2}{1-x} \quad (9)$$

$$P_{qg}(x) = \frac{1}{2} [x^2 + (1-x)^2] \quad (10)$$

$$P_{gg}(x) = 2C_A \left[ \frac{1-x}{x} + \frac{x}{1-x} + x(1-x) \right], \quad (11)$$

where  $C_F = (N_C^2 - 1)/2N_C$  and  $C_A = 2N_C$  are the Casimir invariants of the fundamental and adjoint representation of  $SU(N_C)$  ( $N_C = 3$  for QCD). In the following we shall derive some general properties of the PDF evolution, and give a few concrete examples.

## 2.2 General properties of parton density evolution

Defining the moments of an arbitrary function  $g(x)$  as follows:

$$g_n = \int_0^1 \frac{dx}{x} x^n g(x)$$

it is easy to prove that the evolution equations for the moments turn into ordinary linear differential equations:

$$\frac{df_i^{(n)}}{dt} = \frac{\alpha_s}{2\pi} [P_{qq}^{(n)} f_i^{(n)} + P_{qg}^{(n)} f_g^{(n)}], \quad (12)$$

$$\frac{df_g^{(n)}}{dt} = \frac{\alpha_s}{2\pi} [P_{gg}^{(n)} f_g^{(n)} + P_{gq}^{(n)} f_i^{(n)}]. \quad (13)$$

It is convenient to introduce the concepts of *valence* ( $V(x, t)$ ) and *singlet* ( $\Sigma(x, t)$ ) densities:

$$V(x) = \sum_i f_i(x) - \sum_{\bar{i}} f_{\bar{i}}(x), \quad (14)$$

$$\Sigma(x) = \sum_i f_i(x) + \sum_{\bar{i}} f_{\bar{i}}(x), \quad (15)$$

where the index  $\bar{i}$  refers to the antiquark flavours. The evolution equations then become

$$\frac{dV^{(n)}}{dt} = \frac{\alpha_s}{2\pi} P_{qq}^{(n)} V^{(n)}, \quad (16)$$

<sup>2</sup>The expressions given here are strictly valid only for  $x \neq 1$ . The slight modifications required to extend them to  $x = 1$  will be justified and introduced in the next section.

$$\frac{d\Sigma^{(n)}}{dt} = \frac{\alpha_s}{2\pi} \left[ P_{qq}^{(n)} \Sigma^{(n)} + 2n_f P_{qg}^{(n)} f_g^{(n)} \right], \quad (17)$$

$$\frac{df_g^{(n)}}{dt} = \frac{\alpha_s}{2\pi} \left[ P_{gq}^{(n)} \Sigma^{(n)} + P_{gg}^{(n)} f_g^{(n)} \right]. \quad (18)$$

Note that the equation for the valence density decouples from the evolution of the gluon and singlet densities, which are coupled among themselves. This is physically very reasonable, since in perturbation theory the contribution to the quark and the antiquark densities coming from the evolution of gluons (via their splitting into  $q\bar{q}$  pairs) is the same, and will cancel out in the definition of the valence. The valence therefore only evolves because of gluon emission. On the contrary, gluons and  $q\bar{q}$  pairs in the proton *sea* evolve into one another.

The first moment of  $V(x)$ ,  $V^{(1)} = \int_0^1 dx V(x)$ , counts the number of valence quarks. We therefore expect it to be independent of  $Q^2$ :

$$\frac{dV^{(1)}}{dt} \equiv 0 = \frac{\alpha_s}{2\pi} P_{qq}^{(1)} V^{(1)} = 0. \quad (19)$$

Since  $V^{(1)}$  itself is different from 0, we obtain a constraint on the first moment of the splitting function:  $P_{qq}^{(1)} = 0$ . This constraint is satisfied by including the effect of the virtual corrections, which generate a contribution to  $P_{qq}(z)$  proportional to  $\delta(1-z)$ . This correction is incorporated in  $P_{qq}(z)$  via the redefinition

$$P_{qq}(z) \rightarrow \left( \frac{1+z^2}{1-z} \right)_+ \equiv \frac{1+z^2}{1-z} - \delta(1-z) \int_0^1 dy \left( \frac{1+y^2}{1-y} \right) \quad (20)$$

where the + sign turns  $P_{qq}(z)$  into a distribution. In this way,  $\int_0^1 dz P_{qq}(z) = 0$  and the valence sum-rule is obeyed at all  $Q^2$ .

Another sum rule which does not depend on  $Q^2$  is the momentum sum-rule, which imposes the constraint that all of the momentum of the proton be carried by its constituents (valence plus sea plus gluons):

$$\int_0^1 dx x \left[ \sum_{i,\bar{i}} f_i(x) + f_g(x) \right] \equiv \Sigma^{(2)} + f_g^{(2)} = 1. \quad (21)$$

Once more this relation should hold for all  $Q^2$  values, and you can prove by using the evolution equations that this implies

$$P_{qq}^{(2)} + P_{gq}^{(2)} = 0, \quad (22)$$

$$P_{gg}^{(2)} + 2n_f P_{qg}^{(2)} = 0. \quad (23)$$

You can check using the definition of second moment, and the explicit expressions of the  $P_{qq}$  and  $P_{gq}$  splitting functions, that the first condition is automatically satisfied. The second condition is satisfied by including the virtual effects in the gluon propagator, which contribute a term proportional to  $\delta(1-z)$ . It is a simple exercise to verify that the final form of the  $P_{gg}(z)$  splitting function, satisfying Eq. (23), is

$$P_{gg} \rightarrow 2C_A \left\{ \frac{x}{(1-x)_+} + \frac{1-x}{x} + x(1-x) \right\} + \delta(1-x) \left[ \frac{11C_A - 2n_f}{6} \right]. \quad (24)$$

### 2.3 Solution of the evolution equations

The evolution equations formulated in the previous section can be solved analytically in moment space. The boundary conditions are given by the moments of the parton densities at a given scale  $\mu$ , where, in principle, they can be obtained from a direct measurement. The solution at different values of the scale  $Q$  can then be obtained by inverting numerically the expression for the moments back to  $x$  space. The

resulting evolved densities can then be used to calculate cross-sections for an arbitrary process involving hadrons, at an arbitrary scale  $Q$ . We shall limit ourselves here to studying some properties of the analytic solutions, and will present and comment on some plots obtained from numerical studies available in the literature.

As an exercise, you can show that the solution of the evolution equation for the valence density is the following:

$$V^{(n)}(Q^2) = V^{(n)}(\mu^2) \left[ \frac{\log Q^2/\Lambda^2}{\log \mu^2/\Lambda^2} \right]^{P_{qq}^{(n)}/2\pi b_0} = V^{(n)}(\mu^2) \left[ \frac{\alpha_s(\mu^2)}{\alpha_s(Q^2)} \right]^{P_{qq}^{(n)}/2\pi b_0} \quad (25)$$

where the running of  $\alpha_s(\mu^2)$  has to be taken into account to get the right result. Since all moments  $P^{(n)}$  are negative, the evolution to larger values of  $Q$  makes the valence distribution softer and softer. This is physically reasonable, since the only thing that the valence quarks can do is to lose energy because of gluon emission.

The solutions for the gluon and singlet distributions  $f_g$  and  $\Sigma$  can be obtained by diagonalizing the  $2 \times 2$  system in Eqs. (17) and (18). We study the case of the second moments, which correspond to the momentum fractions carried by quarks and gluons separately. In the asymptotic limit,  $\Sigma^{(2)}$  goes to a constant, and  $\frac{d\Sigma^{(2)}}{dt} = 0$ . Then, using the momentum sum rule:

$$P_{qq}^{(2)} \Sigma^{(2)} + 2n_f P_{qg}^{(2)} f_g^{(2)} = 0, \quad (26)$$

$$\Sigma^{(2)} + f_g^{(2)} = 1. \quad (27)$$

The solution of this system is

$$\Sigma^{(2)} = \frac{1}{1 + \frac{4C_F}{n_f}} \quad (= 15/31 \text{ for } n_f = 5) \quad (28)$$

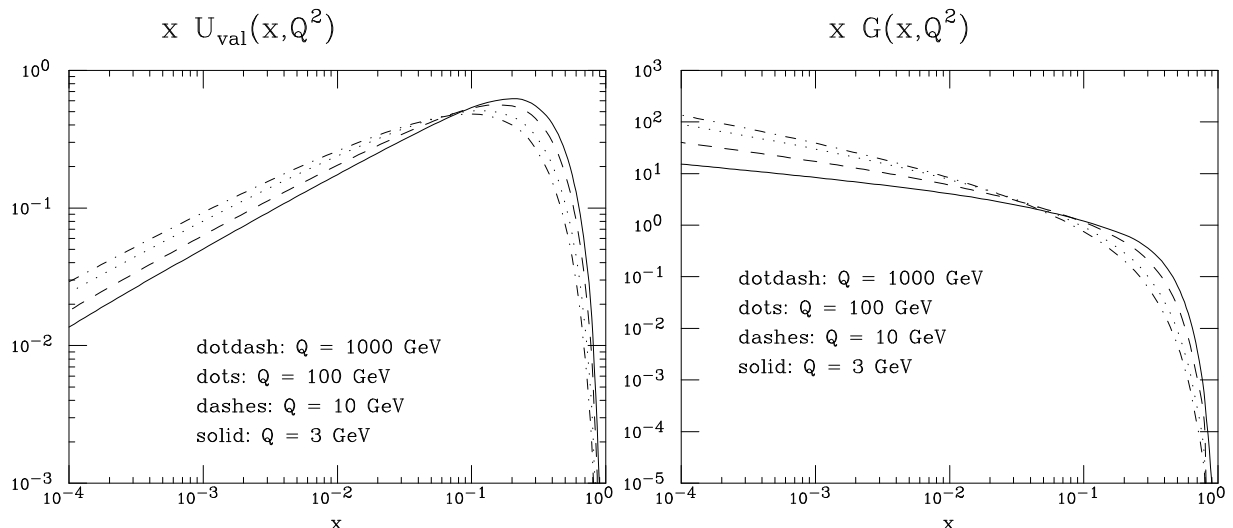
$$f_g^{(2)} = \frac{4C_F}{4C_F + n_f} \quad (= 16/31 \text{ for } n_f = 5). \quad (29)$$

As a result, the fraction of momentum carried by gluons is asymptotically approximately 50% of the total proton momentum. It is interesting to note that, experimentally, this asymptotic value is actually reached already at rather low values of  $Q^2$ . It was indeed observed already in the first deep-inelastic  $ep$  experiments, which exposed the possible presence of quarks in the proton, that only approximately 50% of the proton momentum was carried by charged constituents. This was one of the early pieces of evidence for the existence of gluons.

## 2.4 Example: quantitative evolution of parton densities

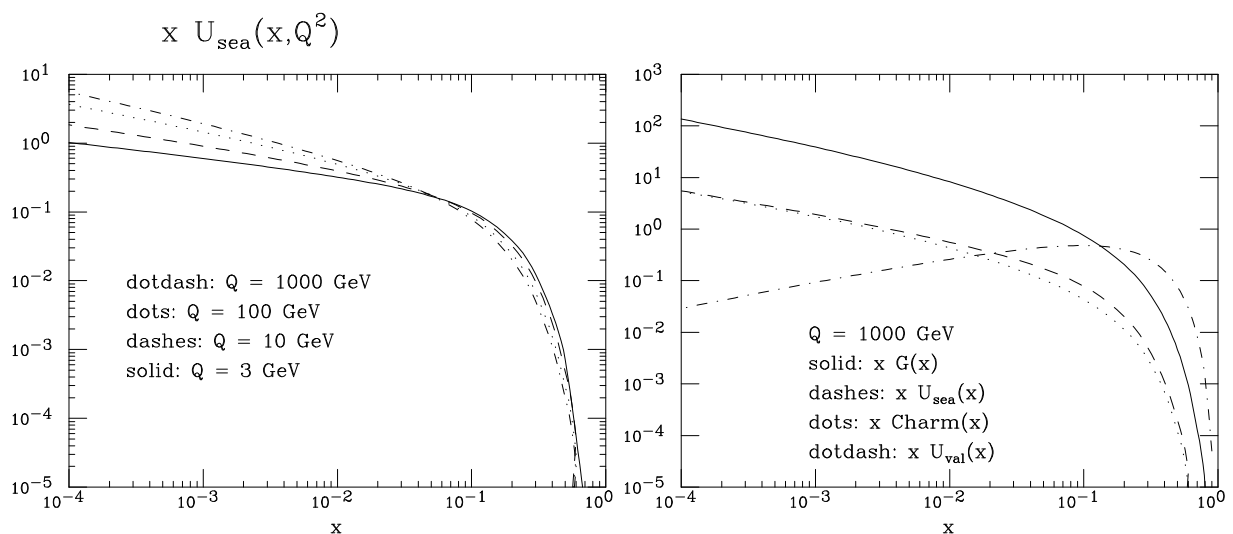
As I mentioned earlier, a complete solution for the evolved parton densities in  $x$  space can only be obtained from a numerical analysis. This work has been done in the past by several groups (see, for example, the chapter on PDFs contained in Ref. [8]), and is continuously being updated by including the most up-to-date experimental results used for the determination of the input densities at a fixed scale. The left side of fig. 7 shows the up-quark valence momentum density at various scales  $Q$ . Note the softening at larger scales, and the clear  $\log Q^2$  evolution. As  $Q^2$  grows, the valence quarks emit more and more radiation, since they change direction over a shorter amount of time (larger acceleration). They therefore lose more momentum to the emitted gluons, and their spectrum becomes softer. The most likely momentum fraction carried by a valence up quark in the proton goes from  $x \sim 20\%$  at  $Q = 3$  GeV, to  $x \lesssim 10\%$  at  $Q = 1000$  GeV. Notice finally that the density vanishes at small  $x$ .

The right-hand side of fig. 7 shows the gluon momentum density. This grows at small  $x$ , with an approximate  $g(x) \sim 1/x^{1+\delta}$  behaviour, and  $\delta > 0$  slowly increasing at large  $Q^2$ . This low- $x$  growth



**Fig. 7:** Left: Valence up-quark momentum-density distribution, for different scales  $Q$ . Right: Gluon momentum-density distribution.

is due to the  $1/x$  emission probability for the radiation of gluons, which was discussed in the previous lecture and which is represented by the  $1/x$  factors in the  $P_{gq}(x)$  and  $P_{gg}(x)$  splitting functions. As  $Q^2$  grows we find an increasing number of gluons at small  $x$ , as a result of the increased radiation off quarks, as well as off the harder gluons.



**Fig. 8:** Left: Sea up-quark momentum-density distribution, for different scales  $Q$ . Right: Momentum-density distribution for several parton species, at  $Q = 1000$  GeV.

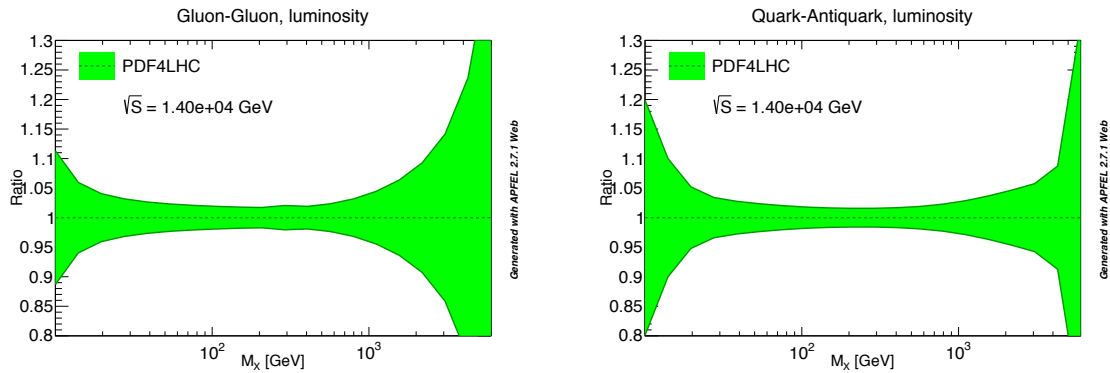
The left-hand side of Fig. 8 shows the evolution of the up-quark *sea* momentum density. Shape and evolution match those of the gluon density, a consequence of the fact that sea quarks come from the splitting of gluons. Since the gluon-splitting probability is proportional to  $\alpha_s$ , the approximate ratio  $sea/gluon \sim 0.1$  which can be obtained by comparing Figs. 7 and 8 is perfectly justified.

Finally, the momentum densities for gluons, up-sea, charm, and up-valence distributions are shown, for  $Q = 1000$  GeV, on the right side of Fig.8. Note here that  $u_{sea}$  and charm are approximately the same at very large  $Q$  and small  $x$ , as will be discussed in more detail in the next subsection.



The proton momentum is mostly carried by valence quarks and by gluons. The contribution of sea quarks is negligible.

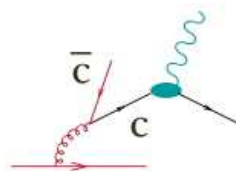
Parton densities are extracted from experimental data. Their determination is therefore subject to the statistical and systematic uncertainties of the experiments and of the theoretical analysis (e.g. the treatment of non-perturbative effects, the impact of missing higher-order perturbative corrections). Techniques have been introduced recently to take into account these uncertainties, and to evaluate their impact on concrete observables. A summary of such an analysis, for the LHC, is given in Fig. 9. What is plotted is the uncertainty bands for partonic luminosities<sup>3</sup> corresponding to the  $gg$  and  $q\bar{q}$  initial-state channels. The partonic flux is given as a function of  $\hat{s}$ , the partonic CM invariant mass. Obvious features include the growth of uncertainty of the  $gg$  density at large mass, corresponding to the lack of data covering the large- $x$  region of the gluon density. Notice that the  $gg$  luminosity drives the production of both  $t\bar{t}$  and Higgs production, at mass values around 350 and 125 GeV, respectively. In this mass range, the PDF uncertainty is today at the level of 2-3%. For the  $q\bar{q}$ , which drives the production of DY pairs (or of new vector gauge bosons), the uncertainty is likewise in the few percent range up to masses of about 1 TeV, and remains below 20% even up to 4 TeV.



**Fig. 9:** Uncertainty in the parton luminosity functions at the LHC, using the PDF4LHC set of PDFs [9]. The plot was obtained using the Apfel web resource [10, 11].

### 2.5 Example: the charm content of the proton

If the virtuality of the external probe is large enough, the time scale of the hard interaction is so short that gluon fluctuations into virtual heavy quark states can be directly exposed, and the virtual heavy quarks (charm quarks in our example) can be brought on-shell via the interaction with the photon (see fig. 10). To the external photon, it will therefore appear as if the proton contained some charm. While in the



**Fig. 10:** Gluon evolution leading to a charm quark content of the proton

case of the gluons and of light quarks the boundary condition for the DGLAP evolution at small  $Q$  is non-perturbative and cannot be derived from first principles, in the case of a heavy quark  $Q$  the boundary condition  $f_Q(x, Q_0) = 0$  holds at a scale  $Q_0 \sim m_Q$  that is large enough for perturbation theory to apply.

<sup>3</sup>For the definition of parton luminosity see Section 4.1.

The charm density can be calculated assuming that the heavy quark density itself is 0 at  $Q \sim m_c$ , and builds up according to the DGLAP evolution equation:

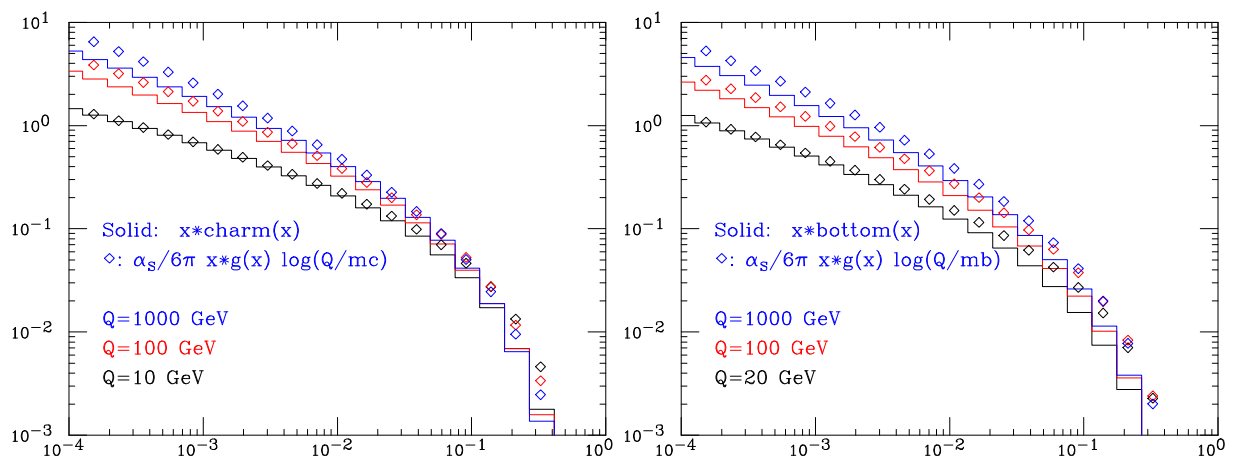
$$\frac{dc(x, Q)}{dt} = \frac{\alpha_s}{2\pi} \int_x^1 \frac{dy}{y} g(y, Q) P_{qg}\left(\frac{x}{y}\right). \quad (30)$$

Assuming a gluon density behaving like  $g(x, Q) \sim A/x$ , which is a first approximation to a bremsstrahlung spectrum, we can easily calculate

$$\begin{aligned} \frac{dc(x, Q)}{dt} &= \frac{\alpha_s}{2\pi} \int_x^1 \frac{dy}{y} g(x/y, Q) P_{qg}(y) = \frac{\alpha_s}{2\pi} \int_x^1 dy \frac{A}{x} \frac{1}{2} [y^2 + (1-y)^2] \\ &= \frac{\alpha_s}{6\pi} \frac{A}{x} c(x, Q) \sim \frac{\alpha_s}{6\pi} \log\left(\frac{Q^2}{m_c^2}\right) g(x, Q). \end{aligned} \quad (31)$$

The charm density is therefore proportional to the gluon density, up to an overall factor proportional to  $\alpha_s$ . When  $Q$  becomes very large, the effect of the quark mass becomes subleading, and we expect all sea quarks to reach asymptotically the same density.

While this is a simplified approach to the estimate of the heavy quark density of the proton, the approximation is rather good. This is shown by the plots in fig. 11, which compare the charm and bottom PDF as given by Eq. 31 with the result extracted from a full set of PDFs. The solid histograms in these plots represent the exact result, for three values of the evolution scale  $Q$ . The diamonds give the approximate results. Notice that the agreement is very good at small  $x$  and at the smaller values of  $Q$ . At larger  $x$  the approximation deteriorates, since in that case the assumption that  $g(x) \sim 1/x$  is no longer valid. At higher scales  $Q$  the exact result becomes smaller than the approximate one, since the latter neglects the momentum loss due to the higher-order gluon radiation (namely the contributions to the evolution equation proportional to  $P_{qq}(y) \times Q(x/y)$ ). Of course, any accurate calculation of cross-sections involving initial-state heavy quarks will make use of the exact results, but it is interesting to see that even in such a complex process it is possible to identify useful analytic approximations that can give us good order-of-magnitude estimates!



**Fig. 11:** Charm and bottom quark PDFs, as obtained from the exact and approximate evolutions.

### 3 The evolution of quarks and gluons

We discussed in the previous section the initial-state evolution of quarks and gluons as the proton approaches the hard collision. We study here how quarks and gluons evolve, and finally transform into hadrons, neutralizing their colours. We start by considering the simplest case:  $e^+e^-$  collisions, which

provide the cleanest environment in which to study applications of QCD at high energy. This is the place where theoretical calculations have today reached their best accuracy, and where experimental data are the most precise, especially thanks to the huge statistics accumulated by LEP, LEP2 and SLC. The key process is the annihilation of the  $e^+e^-$  pair into a virtual photon or  $Z^0$  boson, which will subsequently decay to a  $q\bar{q}$  pair.  $e^+e^-$  collisions have therefore the big advantage of providing an almost point-like source of quark pairs, so that, in contrast to the case of interactions involving hadrons in the initial state, we at least know very precisely the state of the quarks at the beginning of the interaction process.

Nevertheless, it is by no means obvious that this information is sufficient to predict the properties of the hadronic final state. We know that this final state is clearly not simply a  $q\bar{q}$  pair, but some high-multiplicity set of hadrons. For example, as shown in Fig. 12, the average multiplicity of charged hadrons in the decay of a  $Z^0$  is approximately 20. It is therefore not obvious that a calculation done using the

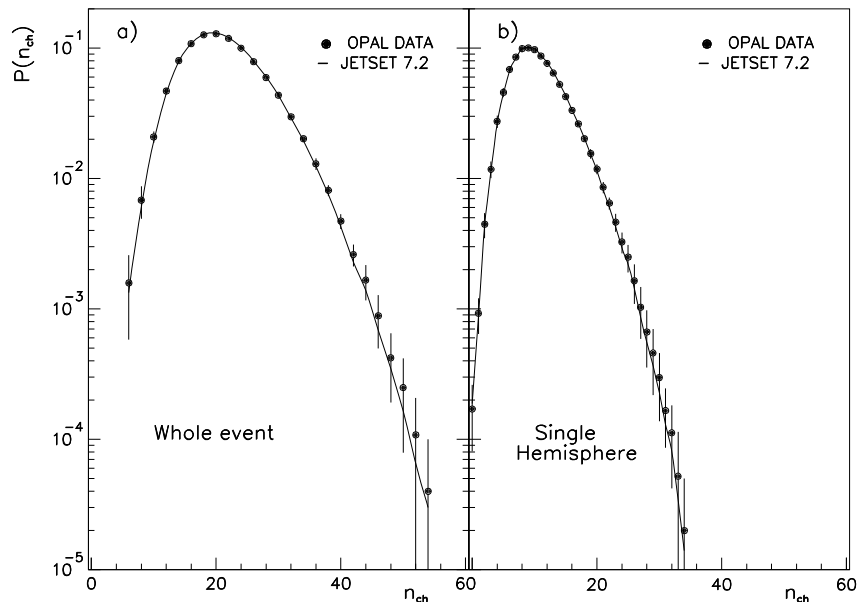


Fig. 12: Charged particle multiplicity distribution in  $Z^0$  decays

simple picture  $e^+e^- \rightarrow q\bar{q}$  (see Fig. 13) has anything to do with reality. For example, one may wonder

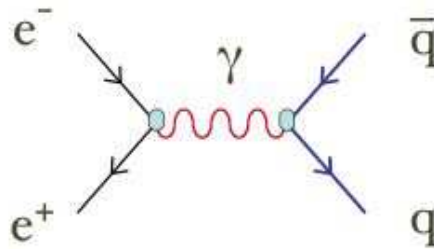


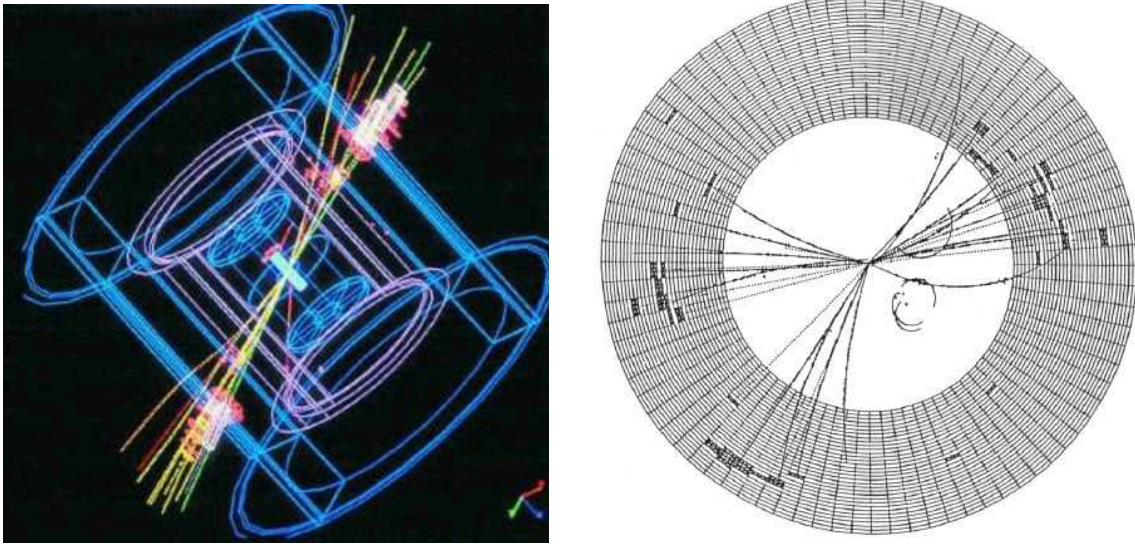
Fig. 13: Tree level production of a  $q\bar{q}$  pair in  $e^+e^-$  collisions

why we do not need to calculate  $\sigma(e^+e^- \rightarrow q\bar{q}g \dots g \dots)$  for all possible gluon multiplicities to get an accurate estimate of  $\sigma(e^+e^- \rightarrow \text{hadrons})$ . And since in any case the final state is not made of  $q$ 's and  $g$ 's, but of  $\pi$ 's,  $K$ 's,  $\rho$ 's, etc., why would  $\sigma(e^+e^- \rightarrow q\bar{q}g \dots g)$  be enough?

The solution to this puzzle lies both in a question of time and energy scales, and in the dynamics of QCD. When the  $q\bar{q}$  pair is produced, the force binding  $q$  and  $\bar{q}$  is proportional to  $\alpha_s(s)$  ( $\sqrt{s}$  being

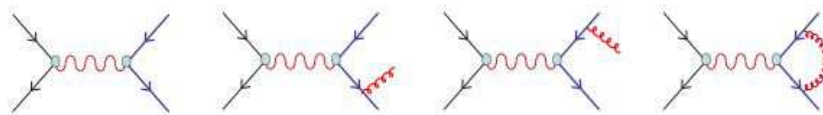
the  $e^+e^-$  centre-of-mass energy). Therefore it is weak, and  $q$  and  $\bar{q}$  behave to good approximation like free particles. The radiation emitted in the first instants after the pair creation is also perturbative, and it will stay so until a time after creation of the order of  $(1 \text{ GeV})^{-1}$ , when radiation with wavelengths  $\gtrsim (1 \text{ GeV})^{-1}$  starts being emitted. At this scale the coupling constant is large, and non-perturbative phenomena and hadronization start playing a rôle. However, as we shall show, colour emission during the perturbative evolution organizes itself in such a way as to form colour-neutral, low-mass, parton clusters highly localized in phase-space. As a result, the complete colour-neutralization (i.e., the hadronization) does not involve long-range interactions between partons far away in phase-space. This is very important, because the forces acting among coloured objects at this time scale would be huge. If the perturbative evolution were to separate far apart colour-singlet  $q\bar{q}$  pairs, the final-state interactions taking place during the hadronization phase would totally upset the structure of the final state.

In this picture, the identification of the perturbative cross-section  $\sigma(e^+e^- \rightarrow q\bar{q})$  with observable, high-multiplicity hadronic final states is realised by jets, namely collimated streams of hadrons that are the final result of the perturbative and non-perturbative evolution of each quark. The large multiplicity



**Fig. 14:** Experimental pictures of 2- and 3-jet final states from  $e^+e^-$  collisions

of the final states, shown in fig. 12, corresponds to the many particles that emerge from the collinear emissions of many gluons from each quark. The dynamics of these emissions leads these particles to grossly follow the direction of the primary quark, and the emergent bundle, the jet, inherits the kinematics of the initial quark. This is shown in the left image of fig. 14. Three-jet events, shown in the right image of the figure, arise from the  $O(\alpha_s)$  corrections to the tree-level process, namely to diagrams such as those shown in fig. 15.



**Fig. 15:**  $O(\alpha_s)$  corrections to the tree-level  $e^+e^- \rightarrow q\bar{q}$  process

An important additional result of this ‘pre-confining’ evolution, is that the memory of where the local colour-neutral clusters came from is totally lost. So we expect the properties of hadronization to be

universal: a model that describes hadronization at a given energy will work equally well at some other energy. Furthermore, so much time has passed since the original  $q\bar{q}$  creation that the hadronization phase cannot significantly affect the total hadron production rate. Perturbative corrections due to the emission of the first hard partons should be calculable in PT, providing a finite, meaningful cross-section.

The nature of non-perturbative corrections to this picture can be explored. One can prove, for example, that the leading correction to the total rate  $R_{e^+e^-}$  is of order  $F/s^2$ , where  $F \propto \langle 0 | \alpha_s F_{\mu\nu}^a F^{\mu\nu a} | 0 \rangle$  is the so-called gluon condensate. Since  $F \sim \mathcal{O}(1 \text{ GeV}^4)$ , these NP corrections are usually very small. For example, they are of  $\mathcal{O}(10^{-8})$  at the  $Z^0$  peak. Corrections scaling like  $\Lambda^2/s$  or  $\Lambda/\sqrt{s}$  can nevertheless appear in other less inclusive quantities, such as event shapes or fragmentation functions.

We now come back to the perturbative evolution, and shall devote the first part of this lecture to justifying the picture given above.

### 3.1 Soft gluon emission

Emission of soft gluons plays a fundamental rôle in the evolution of the final state [7, 13]. Soft gluons are emitted with large probability, since the emission spectrum behaves like  $dE/E$ , typical of bremsstrahlung as familiar in QED. They provide the seed for the bulk of the final-state multiplicity of hadrons. The study of soft-gluon emission is simplified by the simplicity of their couplings. Being soft (i.e., long wavelength) they are insensitive to the details of the very-short-distance dynamics: they cannot distinguish features of the interactions which take place on time scales shorter than their wavelength. They are also insensitive to the spin of the partons: the only feature they are sensitive to is the colour charge. To prove this let us consider soft-gluon emission in the  $q\bar{q}$  decay of an off-shell photon:

$$(32)$$

$$\begin{aligned} A_{soft} &= \bar{u}(p)\epsilon(k)(ig) \frac{-i}{\not{p} + \not{k}} \Gamma^\mu v(\bar{p}) \lambda_{ij}^a + \bar{u}(p) \Gamma^\mu \frac{i}{\not{p} + \not{k}} (ig)\epsilon(k) v(\bar{p}) \lambda_{ij}^a \\ &= \left[ \frac{g}{2p \cdot k} \bar{u}(p)\epsilon(k) (\not{p} + \not{k}) \Gamma^\mu v(\bar{p}) - \frac{g}{2\bar{p} \cdot k} \bar{u}(p) \Gamma^\mu (\not{p} + \not{k}) \epsilon(k) v(\bar{p}) \right] \lambda_{ij}^a . \end{aligned}$$

I used the generic symbol  $\Gamma_\mu$  to describe the interaction vertex with the photon to stress the fact that the following manipulations are independent of the specific form of  $\Gamma_\mu$ . In particular,  $\Gamma_\mu$  can represent an arbitrarily complicated vertex form factor. Neglecting the factors of  $\not{k}$  in the numerators (since  $k \ll p, \bar{p}$ , by definition of soft) and using the Dirac equations, we get:

$$A_{soft} = g\lambda_{ij}^a \left( \frac{p \cdot \epsilon}{p \cdot k} - \frac{\bar{p}\epsilon}{\bar{p} \cdot k} \right) A_{Born} . \quad (33)$$

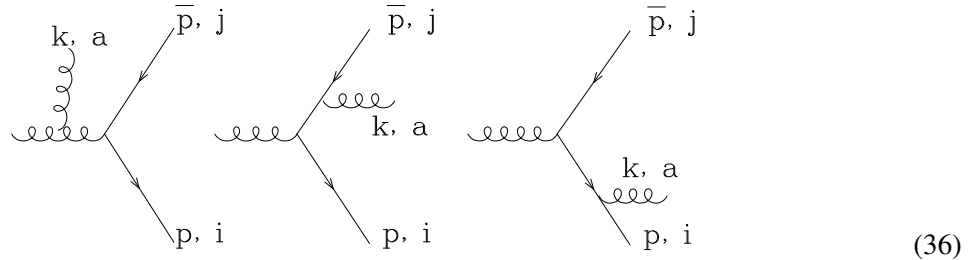
We then conclude that soft-gluon emission factorizes into the product of an emission factor, times the Born-level amplitude. From this exercise, one can extract general Feynman rules for soft-gluon emission:

$$p, j \xrightarrow{\text{gluon}} p, i \quad \text{a, } \mu = g \lambda_{ij}^a 2p^\mu . \quad (34)$$

**Exercise:** Derive the  $g \rightarrow gg$  soft-emission rules:

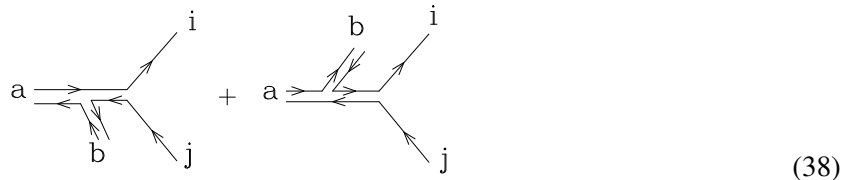
$$c, \nu \xrightarrow{\text{gluon}} b, \rho \quad \text{a, } \mu = ig f^{abc} 2p^\mu g^{\nu\rho} . \quad (35)$$

**Example:** Consider the ‘decay’ of a virtual gluon into a quark pair. One more diagram should be added to those considered in the case of the electroweak decay. The fact that the quark pair is no longer in a colour-singlet state makes things a bit more interesting:

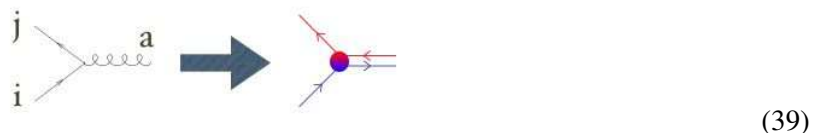


$$\begin{aligned} &\stackrel{k \rightarrow 0}{=} \left[ ig f^{abc} \lambda_{ij}^c \left( \frac{Q\epsilon}{Qk} \right) + g (\lambda^b \lambda^a)_{ij} \left( \frac{p\epsilon}{pk} \right) - g (\lambda^a \lambda^b)_{ij} \left( \frac{\bar{p}\epsilon}{pk} \right) \right] A_{Born} \\ &= g (\lambda^a \lambda^b)_{ij} \left[ \frac{Q\epsilon}{Qk} - \frac{\bar{p}\epsilon}{pk} \right] + g (\lambda^b \lambda^a)_{ij} \left[ \frac{p\epsilon}{pk} - \frac{Q\epsilon}{Qk} \right] . \end{aligned} \quad (37)$$

The two factors correspond to the two possible ways colour can flow in this process:



The basis for this representation of the colour flow is the following diagram which makes explicit the relation between the colours of the quark, antiquark, and gluon entering a QCD vertex:



We can therefore represent the gluon as a double line, one line carrying the colour inherited from the quark, the other carrying the anticolour inherited from the antiquark. In the first diagram in (38) the antiquark (colour label  $j$ ) is colour connected to the soft gluon (colour label  $b$ ), and the quark (colour label  $i$ ) is connected to the decaying gluon (colour label  $a$ ). In the second case, the order is reversed. The two emission factors correspond to the emission of the soft gluon from the antiquark, and from the quark line, respectively. When squaring the total amplitude, and summing over initial and final-state colours, the interference between the two pieces is suppressed by  $1/N^2$  relative to the individual squares:

$$\sum_{a,b,i,j} |(\lambda^a \lambda^b)_{ij}|^2 = \sum_{a,b} \text{tr} \left( \lambda^a \lambda^b \lambda^b \lambda^a \right) = \frac{N^2 - 1}{2} C_F = \mathcal{O}(N^3). \quad (40)$$

$$\sum_{a,b,i,j} (\lambda^a \lambda^b)_{ij} [(\lambda^b \lambda^a)_{ij}]^* = \sum_{a,b} \text{tr}(\lambda^a \lambda^b \lambda^a \lambda^b) = \frac{N^2 - 1}{2} \underbrace{\left( C_F - \frac{C_A}{2} \right)}_{-\frac{1}{2N}} = \mathcal{O}(N). \quad (41)$$

As a result, the emission of a soft gluon can be described, to the leading order in  $1/N^2$ , as the incoherent sum of the emission from the two colour currents.

### 3.2 Angular ordering for soft-gluon emission

The results presented above have important consequences for the perturbative evolution of the quarks. A key property of the soft-gluon emission is the so-called *angular ordering*. This phenomenon consists in the continuous reduction of the opening angle at which successive soft gluons are emitted by the evolving quark. As a result, this radiation is confined within smaller and smaller cones around the quark direction, and the final state will look like a collimated jet of partons. In addition, the structure of the colour flow during the jet evolution forces the  $q\bar{q}$  pairs which are in a colour-singlet state to be close in phase-space, thereby achieving the pre-confinement of colour-singlet clusters alluded to at the beginning of this section.

Let us start first by proving the property of colour ordering. Consider the  $q\bar{q}$  pair produced by the decay of a rapidly moving virtual photon. The amplitude for the emission of a soft gluon was given in Eq. (33). Squaring, summing over colours, and including the gluon phase-space we get the following result:

$$\begin{aligned} d\sigma_g &= \sum |A_{soft}|^2 \frac{d^3k}{(2\pi)^3 2k^0} \sum |A_0|^2 \frac{-2p^\mu \bar{p}^\nu}{(pk)(\bar{p}k)} g^2 \sum \epsilon_\mu \epsilon_\nu^* \frac{d^3k}{(2\pi)^3 2k^0} \\ &= d\sigma_0 \frac{2(p\bar{p})}{(pk)(\bar{p}k)} g^2 C_f \left( \frac{d\phi}{2\pi} \right) \frac{k^0 dk^0}{8\pi^2} d \cos \theta \\ &= d\sigma_0 \frac{\alpha_s C_F}{\pi} \frac{dk^0}{k^0} \frac{d\phi}{2\pi} \frac{1 - \cos \theta_{ij}}{(1 - \cos \theta_{ik})(1 - \cos \theta_{jk})} d \cos \theta \end{aligned} \quad (42)$$

where  $\theta_{\alpha\beta} = \theta_\alpha - \theta_\beta$ , and  $i, j, k$  refer to the  $q, \bar{q}$  and gluon directions, respectively. We can write the following identity:

$$\begin{aligned} \frac{1 - \cos \theta_{ij}}{(1 - \cos \theta_{ik})(1 - \cos \theta_{jk})} &= \\ \frac{1}{2} \left[ \frac{\cos \theta_{jk} - \cos \theta_{ij}}{(1 - \cos \theta_{ik})(1 - \cos \theta_{jk})} + \frac{1}{1 - \cos \theta_{ik}} \right] &+ \frac{1}{2} [i \leftrightarrow j] \equiv W_{(i)} + W_{(j)}. \end{aligned} \quad (43)$$

We would like to interpret the two functions  $W_{(i)}$  and  $W_{(j)}$  as radiation probabilities from the quark and antiquark lines. Each of them is in fact only singular in the limit of gluon emission parallel to the respective quark:

$$W_{(i)} \rightarrow \text{finite if } k \parallel j \text{ (} \cos \theta_{jk} \rightarrow 1 \text{)}. \quad (44)$$



$$W_{(j)} \rightarrow \text{finite if } k \parallel i \text{ (} \cos \theta_{ik} \rightarrow 1 \text{)}. \quad (45)$$

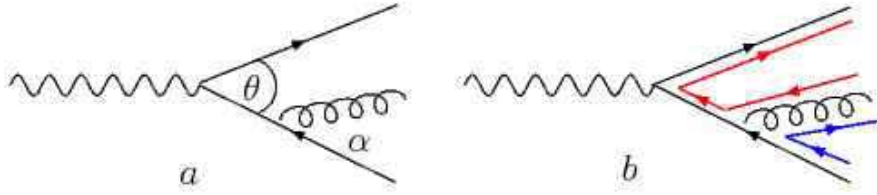
The interpretation as probabilities is however limited by the fact that neither  $W_{(i)}$  nor  $W_{(j)}$  are positive definite. However, you can easily prove that

$$\int \frac{d\phi}{2\pi} W_{(i)} = \begin{cases} \frac{1}{1-\cos\theta_{ik}} & \text{if } \theta_{ik} < \theta_{ij} \\ 0 & \text{otherwise} \end{cases} \quad (46)$$

where the integral is the azimuthal average around the  $q$  direction. A similar result holds for  $W_{(j)}$ :

$$\int \frac{d\phi}{2\pi} W_{(j)} = \begin{cases} \frac{1}{1-\cos\theta_{jk}} & \text{if } \theta_{jk} < \theta_{ij} \\ 0 & \text{otherwise} \end{cases} \quad (47)$$

As a result, the emission of soft gluons outside the two cones obtained by rotating the antiquark direction around the quark's, and vice versa, averages to 0. Inside the two cones, one can consider the radiation from the emitters as being uncorrelated. In other words, the two colour lines defined by the quark and antiquark currents act as independent emitters, and the quantum coherence (i.e. the effects of interference between the two graphs contributing to the gluon-emission amplitude) is accounted for by constraining the emission to take place within those fixed cones.



**Fig. 16:** Radiation off  $q\bar{q}$  pair produced by an off-shell photon

A simple derivation of angular ordering, which more directly exhibits its physical origin, can be obtained as follows. Consider Fig. 16(a), which shows a Feynman diagram for the emission of a gluon from a quark line. The quark momentum is denoted by  $l$  and the gluon momentum by  $k$ ,  $\theta$  is the opening angle between the quark and antiquark, and  $\alpha$  is the angle between the nearest quark and the emitted gluon. We shall work in the double-log enhanced soft  $k^0 \ll l^0$  and collinear  $\alpha \ll 1$  region. The internal quark propagator  $p = (l + k)$  is off-shell, setting the time scale for the gluon emission:

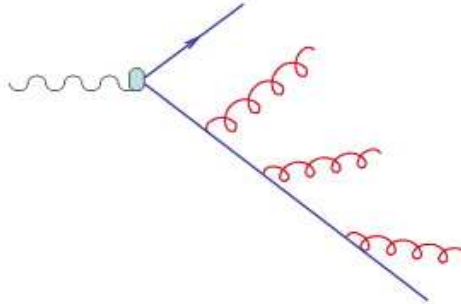
$$\Delta t \simeq \frac{1}{\Delta E} = \frac{l^0}{(k+l)^2} \rightarrow \Delta t \simeq \frac{1}{k^0 \alpha^2}. \quad (48)$$

In order to resolve the quarks, the transverse wavelength of the gluon  $\lambda_{\perp} = 1/E_{\perp}$  must be smaller than the separation between the quarks  $b(t) \simeq \theta \Delta t$ , giving the constraint  $1/(\alpha k^0) < \theta \Delta t$ . Using the results of Eq. 48 for  $\Delta t$ , we arrive at the angular ordering constraint  $\alpha < \theta$ . Gluon emissions at an angle smaller than  $\theta$  can resolve the two individual colour quarks and are allowed; emissions at greater angles do not see the colour charge and are therefore suppressed. In processes involving more partons, the angle  $\theta$  is defined not by the nearest parton, but by the colour connected parton (e.g. the parton that forms a colour singlet with the emitting parton). Figure 16(b) shows the colour connections for the  $q\bar{q}$  event after the gluon is emitted. Colour lines begin on quarks and end on antiquarks. Because gluons are colour octets, they contain the beginning of one line and the end of another, as we showed in (38).

If one repeats now the exercise for emission of one additional gluon, one will find the same angular constraint, but this time applied to the colour lines defined by the previously established *antenna*. As

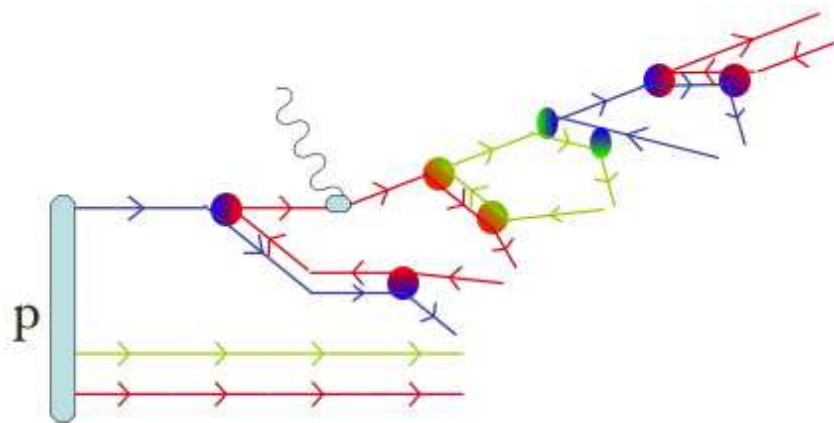


shown in the previous subsection, the  $q\bar{q}g$  state can be decomposed at the leading order in  $1/N$  into two independent emitters, one given by the colour line flowing from the gluon to the quark, the other given by the colour line flowing from the antiquark to the gluon. So the emission of the additional gluon will be constrained to take place either within the cone formed by the quark and the gluon, or within the cone formed by the gluon and the antiquark. Either way, the emission angle will be smaller than the angle of the first gluon emission. This leads to the concept of angular ordering, with successive emission of soft gluons taking place within cones which get smaller and smaller, as in Fig. 17



**Fig. 17:** Collimation of soft gluon emission during the jet evolution

The fact that colour always flows directly from the emitting parton to the emitted one, the collimation of the jet, and the softening of the radiation emitted at later stages, ensure that partons forming a colour-singlet cluster are close in phase-space. As a result, hadronization (the non-perturbative process that will bind together colour-singlet parton pairs) takes place locally inside the jet and is not a long-distance phenomenon connecting partons far away in the evolution tree: only pairs of nearby partons are involved. In particular, there is no direct link between the precise nature of the hard process and the hadronization. These two phases are totally decoupled and, as in the case of the partonic densities, one can infer that hadronization factorizes from the hard process and can be described in a universal (i.e. hard-process independent) fashion. The inclusive properties of jets (e.g. the particle multiplicity, jet mass, jet broadening) are independent of the hadronization model, up to corrections of order  $(\Lambda/\sqrt{s})^n$  (for some integer power  $n$ , which depends on the observable), with  $\Lambda \lesssim 1$  GeV.



**Fig. 18:** The colour flow diagram for a DIS event

The final picture, in the case of a DIS event, appears therefore as in Fig. 18. After being deflected by the photon, the struck quark emits the first gluon that takes away the quark colour and passes on its own anticolour to the escaping quark. This gluon is therefore colour-connected with the last gluon emitted

before the hard interaction. As the final-state quark continues its evolution, more and more gluons are emitted, each time leaving their colour behind and transmitting their anticolour to the emerging quark. Angular ordering forces all these gluons to be close in phase-space, until the evolution is stopped once the virtuality of the quark becomes of the order of the strong-interaction scale. The colour of the quark is left behind, and when hadronization takes over it is only the nearby colour-connected gluons which are transformed, with a phenomenological model, in hadrons. This mechanism for the transfer of colour across subsequent gluon emissions is similar to what happens when we place a charge near the surface of a dielectric medium. This will become polarized, and a charge will appear on the medium opposite end. The appearance of the charge is the result of a sequence of local charge shifts, whereby neighbouring atoms get polarized, as in Fig. 19.



Fig. 19: Charge transfer in a dielectric medium, via a sequence of local polarizations

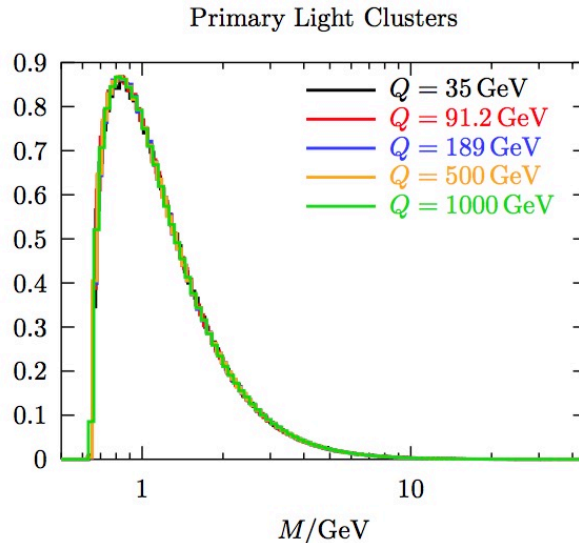
### 3.3 Hadronization

The application of perturbation theory to the evolution of a jet, with the sequential emission of partons, governed by QCD splitting probabilities and angular ordering to enforce quantum mechanical coherence, will stop once the scale of the emissions reaches values in the range of 1 GeV. This is called the *infrared cutoff*. There are two reasons why we need to stop the emission of gluons at this scale. To start with, we cannot control with perturbation theory the domain below this scale, where the strong coupling constant  $\alpha_s$  becomes very large. Furthermore, we know that the number of physical particles that can be produced inside a jet must be finite, since the lightest object we can produce is a pion, and energy conservation sets a limit to how many pions can be created. This is different from what happens in a QED cascade, where the evolution of an accelerated charge can lead to the emission of an arbitrary number of photons. This is possible because the photon is massless, and can have arbitrarily small energy. The gluons of a QCD cascade, on the contrary, must have enough energy to create pions.

When the perturbative evolution of the jet terminates, we are left with some number of gluons. As shown in the previous subsection, and displayed in fig. 18, these gluons are pairwise colour connected. As two colour-connected gluons travel away from each other, a constant force pulls them together. Phenomenological models (see Ref. [7] for a more complete review) are then used to describe how this force determines the evolution of the system from this point on. What I shall describe here is the so-called *cluster model* [17], but the main qualitative features are shared by other alternatives, such as the Lund string approach [18].

Most of the hadrons emerging from the evolution of a jet are known to be made of quarks; glueballs, i.e. hadrons made of bound gluons, are expected to exist, but their production is greatly suppressed compared to that of quark-made particles. For this reason, the first step in the description of hadronization is to assume that the force among gluons will rip them apart into a  $q\bar{q}$  pair, and that these quarks will act as seeds for the hadron production. The break-up into quarks is not parametrized using the DGLAP  $g \rightarrow q\bar{q}$  splitting function, since we are dealing here with a non-perturbative transition. One typically employs therefore a pure phase-space ‘decay’ of the gluon into the  $q\bar{q}$  pair, introducing as phenomenological parameters the relative probabilities of selecting the various active flavours (up, down, strange, etc.). The quark  $q_i$  from one gluon ( $i$  representing the flavour) then forms a colour-singlet pair with the

antiquark  $\bar{q}_j$  emerging from the break-up of the neighboring gluon. This colour-singlet  $q_i\bar{q}_j$  pair cannot, however, directly form a hadron, since in general the quarks will still be moving apart, and the invariant mass of the pair will not coincide with the mass of an existing physical state. As they separate subject to



**Fig. 20:** Invariant mass distribution of clusters of colour-singlet quarks after non-perturbative gluon splitting. The spectra for final states corresponding to different centre-of-mass energies are normalized to the same area, displaying the energy independence of the shapes

a constant force, however, their kinetic energy turns into a linearly-rising potential energy. The potential energy accumulated in the system will be able to convert into a new quark–antiquark pair,  $q_k\bar{q}_k$  once its value exceeds the relevant mass threshold. We are now left with two colour-singlet pairs,  $q_i\bar{q}_k$  and  $q_k\bar{q}_j$ . One can force the kinematics of the  $q_k\bar{q}_k$  pair to allow for both  $q_i\bar{q}_k$  and  $q_k\bar{q}_j$  invariant masses to coincide with some resonance with the proper flavour. The residual energy of the system is then assumed to be entirely kinetic, and the two resonances fly away free. Once again, one can associate phenomenological parameters to the probabilities of selecting flavours  $k$  of a given type. Since the pair of flavour indices  $ik$  does not specify uniquely a hadron (e.g. a  $u\bar{d}$  system could be a  $\pi^+$ , a  $\rho^+$ , as well as many other objects), the model has a further set of rules and/or parameters to select the precise flavour type. For example, a phenomenologically successful description of the  $\pi/\rho$  ratio is obtained by simply assuming a production rate proportional to the number of spin states (one for the scalar pion, three for the vector rho) and to a Boltzmann factor  $\exp(-M/T)$ , where  $M$  is the resonance mass and  $T$  is a universal parameter, to be fit from data. Furthermore, one can introduce the possibility of converting the potential energy into a diquark–antidiquark pair, namely  $(q_kq_\ell)(\bar{q}_k\bar{q}_\ell)$ . The resulting hadrons,  $q_iq_kq_\ell$  and  $\bar{q}_j\bar{q}_k\bar{q}_\ell$  will be a baryon–antibaryon pair.

The measurement of hadron multiplicities from  $Z^0$  decays is used to tune the few phenomenological parameters of the model, and these parameters can be used to describe hadronization at different energies and in different high-energy hadron-production processes. The internal consistency of this assumption is supported by fig. 20, which shows the invariant mass distribution of clusters of colour-singlet quarks, after the non-perturbative gluon splitting, for  $e^+e^-$  collisions at different center-of-mass energies. All curves are normalized to 1, and they all overlap very accurately. This confirms the validity of the implementation of factorization in the Monte Carlo: higher initial energies provide more room for the perturbative evolution, leading to more splitting and more emitted radiation; but the structure and distribution of colour-singlet clusters at the end of evolution is independent of the initial energy, and the same model of hadronization can be applied.

**Table 1:** Average particle multiplicities per event in  $e^+e^-$  collisions at 91.2 GeV. Experimental data were measured by the following collaborations at LEP and at SLC: ALEPH(A), DELPHI(D), L3(L), OPAL(O), MARK2(M) and SLD(S). The theoretical predictions in the last three columns, taken from Ref. [19], correspond to various implementations of the cluster hadronization model (see Ref. [19] for details). The \* indicates a prediction that differs from the measured value by more than three standard deviations.

Particle	Experiment	Measured	Old Model	Herwig++	Fortran
All Charged	M,A,D,L,O	$20.924 \pm 0.117$	20.22*	20.814	20.532*
$\gamma$	A,O	$21.27 \pm 0.6$	23.03	22.67	20.74
$\pi^0$	A,D,L,O	$9.59 \pm 0.33$	10.27	10.08	9.88
$\rho(770)^0$	A,D	$1.295 \pm 0.125$	1.235	1.316	1.07
$\pi^\pm$	A,O	$17.04 \pm 0.25$	16.30	16.95	16.74
$\rho(770)^\pm$	O	$2.4 \pm 0.43$	1.99	2.14	2.06
$\eta$	A,L,O	$0.956 \pm 0.049$	0.886	0.893	0.669*
$\omega(782)$	A,L,O	$1.083 \pm 0.088$	0.859	0.916	1.044
$\eta'(958)$	A,L,O	$0.152 \pm 0.03$	0.13	0.136	0.106
$K^0$	S,A,D,L,O	$2.027 \pm 0.025$	2.121*	2.062	2.026
$K^*(892)^0$	A,D,O	$0.761 \pm 0.032$	0.667	0.681	0.583*
$K^*(1430)^0$	D,O	$0.106 \pm 0.06$	0.065	0.079	0.072
$K^\pm$	A,D,O	$2.319 \pm 0.079$	2.335	2.286	2.250
$K^*(892)^\pm$	A,D,O	$0.731 \pm 0.058$	0.637	0.657	0.578
$\phi(1020)$	A,D,O	$0.097 \pm 0.007$	0.107	0.114	0.134*
$p$	A,D,O	$0.991 \pm 0.054$	0.981	0.947	1.027
$\Delta^{++}$	D,O	$0.088 \pm 0.034$	0.185	0.092	0.209*
$\Sigma^-$	O	$0.083 \pm 0.011$	0.063	0.071	0.071
$\Lambda$	A,D,L,O	$0.373 \pm 0.008$	0.325*	0.384	0.347*
$\Sigma^0$	A,D,O	$0.074 \pm 0.009$	0.078	0.091	0.063
$\Sigma^+$	O	$0.099 \pm 0.015$	0.067	0.077	0.088
$\Sigma(1385)^\pm$	A,D,O	$0.0471 \pm 0.0046$	0.057	0.0312*	0.061*
$\Xi^-$	A,D,O	$0.0262 \pm 0.001$	0.024	0.0286	0.029
$\Xi(1530)^0$	A,D,O	$0.0058 \pm 0.001$	0.026*	0.0288*	0.009*
$\Omega^-$	A,D,O	$0.00125 \pm 0.00024$	0.001	0.00144	0.0009
$f_2(1270)$	D,L,O	$0.168 \pm 0.021$	0.113	0.150	0.173
$f_2'(1525)$	D	$0.02 \pm 0.008$	0.003	0.012	0.012
$D^\pm$	A,D,O	$0.184 \pm 0.018$	0.322*	0.319*	0.283*
$D^*(2010)^\pm$	A,D,O	$0.182 \pm 0.009$	0.168	0.180	0.151*
$D^0$	A,D,O	$0.473 \pm 0.026$	0.625*	0.570*	0.501
$D_s^\pm$	A,O	$0.129 \pm 0.013$	0.218*	0.195*	0.127
$D_s^{*\pm}$	O	$0.096 \pm 0.046$	0.082	0.066	0.043
$J/\Psi$	A,D,L,O	$0.00544 \pm 0.00029$	0.006	0.00361*	0.002*
$\Lambda_c^+$	D,O	$0.077 \pm 0.016$	0.006*	0.023*	0.001*
$\Psi'(3685)$	D,L,O	$0.00229 \pm 0.00041$	0.001*	0.00178	0.0008*

An example of the quality of the fits to  $Z^0$ -decay data is given in table 1, which is taken from [19]. There more details are given on the possible variants of cluster hadronization model and on the choice of parameters used in the fits. Overall, the agreement is excellent!

#### 4 Applications to hadronic collisions

In hadronic collisions all phenomena are QCD related. The dynamics is more complex than in  $e^+e^-$  or DIS, since both beam and target have a non-trivial partonic structure. As a result, calculations (and experimental analyses) are more complicated. QCD phenomenology is however much richer, and the higher energies available in hadronic collisions allow one to probe the structure of the proton and of its

constituents at the smallest scales attainable in a laboratory, in addition to probing the large scales where new physics phenomena may appear.

Recent remarkable progress in theoretical calculations, furthermore, has allowed to achieve next-to-leading-order (NLO) accuracy for all observables of interest, and a large majority of these can now be calculated also to next-to-next-to-leading order (NNLO). The overall precision in the theoretical predictions is therefore reaching now the level of accuracy that until recently was only imaginable for  $e^+e^-$  or  $ep$  collisions.

The key ingredients for the calculation of production rates and distributions in hadronic collisions are

- the matrix elements for the hard, partonic process (e.g.,  $gg \rightarrow gg, gg \rightarrow b\bar{b}, q\bar{q}' \rightarrow W, \dots$ ),
- the hadronic parton densities, discussed in the previous lecture.

Then the production rate for a given final state  $H$  is given by a factorization formula similar to the one used to describe DIS:

$$d\sigma(p\bar{p} \rightarrow H + X) = \int dx_1 dx_2 \sum_{i,j} f_i(x_1, Q) f_j(x_2, Q) d\hat{\sigma}(ij \rightarrow H). \quad (49)$$

where the parton density  $f_i$ 's are evaluated at a scale  $Q$  typical of the hard process under consideration. For example  $Q \simeq M_{DY}$  for production of a Drell–Yan pair,  $Q \simeq E_T$  for high transverse-energy ( $E_T$ ) jets,  $Q^2 \simeq p_T^2 + m_Q^2$  for high- $p_T$  heavy quarks, etc.

In this lecture we shall briefly explore two of the QCD phenomena currently studied in hadronic collisions: Drell–Yan, and inclusive jet production. More details can be found in Refs. [7, 8].

#### 4.1 QCD aspects of inclusive vector boson production

The main feature of inclusive gauge boson production in hadronic collisions is that the leading-order (LO) amplitude, describing the elementary process  $q\bar{q}' \rightarrow V$  ( $V = W, Z$ ) is purely EW. The dynamics of strong interactions, at this order, only enters indirectly through the parton distribution functions (PDFs), which parameterize in a phenomenological way the quark and gluon content of the proton. At the large momentum scales typical of gauge boson production ( $Q \sim M_V$ ), higher-order perturbative QCD corrections to the inclusive production are proportional to  $\alpha_s(Q)$  and are typically small, in the range of 10–20%. They are known [20, 21] today to next-to-next-to-leading order (NNLO), including the description of the differential distributions of the boson and of its decay leptons [22–25], leaving theoretical uncertainties from higher-order QCD effects at the percent level. These results have been incorporated in full Monte Carlo calculations including the shower evolution, to give a complete description of the physical final states [26–28]. Next-to-leading-order (NLO) EW corrections are also known [29–32], and play an important role both for precision measurements, and in the production rate of dilepton pairs at large  $p_T$  or with large mass, above the TeV, where they can be larger than 10%. Finally, progress towards a complete calculation of the mixed  $\mathcal{O}(\alpha_s\alpha)$  corrections has been recently reported in Ref. [33].

When considering the first and second generation quarks that dominate the production of  $W$  and  $Z$  bosons, their weak couplings, including the CKM mixing parameters, are known experimentally with a precision better than a percent. This exceeds the accuracy of possible measurements in hadronic collisions, indicating that such measurements could not be possibly affected, at this level of precision, by the presence of new physics phenomena. They therefore provide an excellent ground to probe to percent precision the effects of higher-order QCD corrections and of PDFs [34]. To be more explicit, consider the leading-order (LO) cross section given by eq. 49. In the case of  $W$  production (a similar result holds for the  $Z$ ), the LO partonic cross section is given by:

$$\hat{\sigma}(q_i\bar{q}_j \rightarrow W) = \pi \frac{\sqrt{2}}{3} |V_{ij}|^2 G_F M_W^2 \delta(\hat{s} - M_W^2) = A_{ij} M_W^2 \delta(\hat{s} - M_W^2) \quad (50)$$

Here  $\hat{s} = x_1 x_2 S$  is the partonic centre-of-mass energy squared, and  $V_{ij}$  is the element of the Cabibbo–Kobayashi–Maskawa (CKM) matrix.

Written in terms of  $\tau = x_1 x_2$  and of the rapidity  $y = \log[(E_W + p_W^z)/(E_W - p_W^z)]^{1/2} \equiv \log(x_1/x_2)^{1/2}$ , the differential and total cross sections are given by:

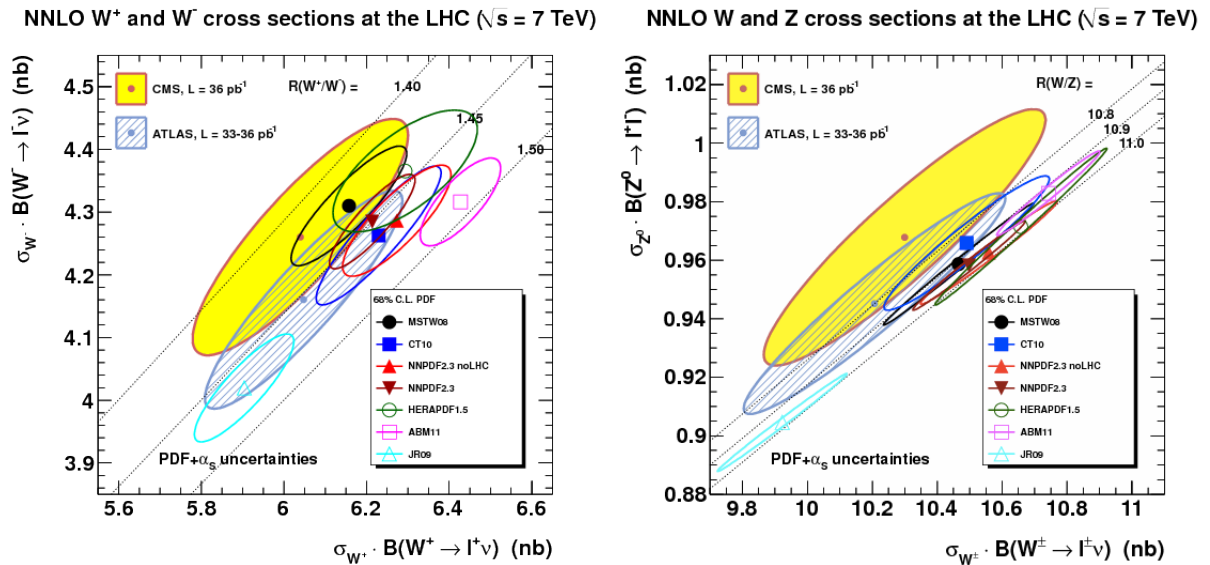
$$\frac{d\sigma_W}{dy} = \sum_{i,j} \frac{\pi A_{ij}}{M_W^2} \tau f_i(x_1) f_j(x_2), \quad x_{1,2} = \sqrt{\tau} e^{\pm y} \quad (51)$$

$$\sigma_W = \sum_{i,j} \frac{\pi A_{ij}}{M_W^2} \tau \int_{\tau}^1 \frac{dx}{x} f_i(x) f_j\left(\frac{\tau}{x}\right) \equiv \sum_{i,j} \frac{\pi A_{ij}}{M_W^2} \tau L_{ij}(\tau) \quad (52)$$

where the function  $L_{ij}(\tau)$  is usually called *partonic luminosity*. In the case of  $u\bar{d}$  collisions,  $\frac{\pi A_{ij}}{M_W^2} \sim 6.5 \text{ nb}$ . It is interesting to study the partonic luminosity as a function of the hadronic centre-of-mass energy. This can be done by taking a simple approximation for the parton densities. Using the approximate behaviour  $f_i(x) \sim 1/x^{1+\delta}$ , with  $\delta < 1$ , one easily obtains:

$$L(\tau) = \frac{1}{\tau^{1+\delta}} \log\left(\frac{1}{\tau}\right) \quad \text{and} \quad \sigma_W \propto \left(\frac{S}{M_W^2}\right)^{\delta} \log\left(\frac{S}{M_W^2}\right). \quad (53)$$

The gauge boson production cross section grows therefore at least logarithmically with the hadronic centre-of-mass energy.



**Fig. 21:**  $W$  and  $Z$  boson cross sections in  $pp$  collisions at  $\sqrt{S} = 7 \text{ TeV}$ : ATLAS [35] and CMS [36] data, compared to NNLO predictions for various PDF sets [34].

#### 4.1.1 Rapidity spectrum of $W$ and $Z$ bosons and $W$ charge asymmetries

The features of the momentum distribution of vector bosons along the beam direction ( $z$ ) are mostly controlled by properties of the parton PDFs. In particular, in the case of  $W$  bosons, the differences between the PDFs of up- and down-type quarks and antiquarks lead to interesting production asymmetries. Since the measurement of asymmetries is typically very accurate, due to the cancellation of many experimental and theoretical uncertainties, these play a fundamental role in the precision determination of quark and antiquark PDFs. Furthermore, the production asymmetries are modulated by the parity violation of the

vector boson couplings, leading to further handles for the discrimination of quark and antiquark densities, and inducing a sensitivity to the weak mixing angle  $\sin^2 \theta_{\text{eff}}^{\text{lept}}$ , which controls the vector and axial components of  $Z$  boson interactions.

For  $p\bar{p}$  collisions, and assuming for simplicity the dominance of  $u$  and  $d$  quarks, we have:

$$\frac{d\sigma_{W^+}}{dy} \propto f_u^p(x_1) f_{\bar{d}}^{\bar{p}}(x_2) + f_d^p(x_1) f_{\bar{u}}^{\bar{p}}(x_2) \quad (54)$$

$$\frac{d\sigma_{W^-}}{dy} \propto f_{\bar{u}}^p(x_1) f_d^{\bar{p}}(x_2) + f_d^p(x_1) f_{\bar{u}}^{\bar{p}}(x_2) \quad (55)$$

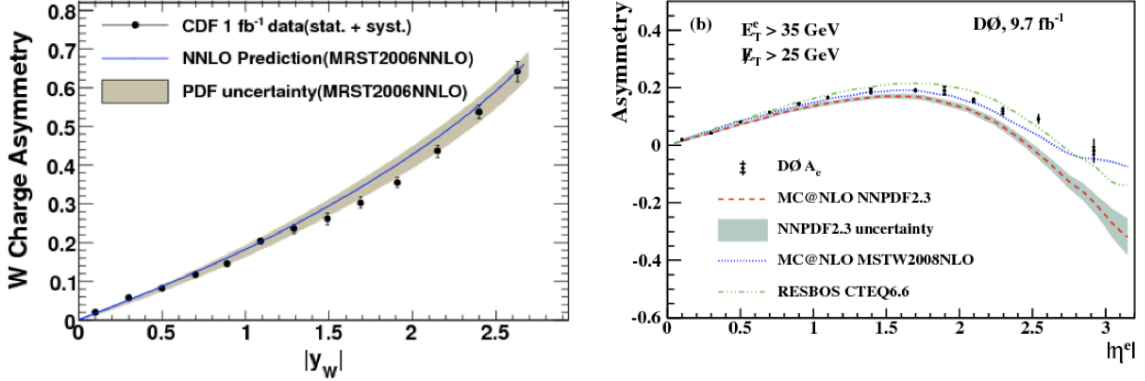
We can then construct the following charge asymmetry (using  $f_{\bar{q}}^{\bar{p}} = f_q^p$  and assuming the dominance of the quark densities over the antiquark ones, which is valid in the kinematical region of interest for  $W$  production at the Tevatron):

$$A(y) = -A(-y) = \frac{\frac{d\sigma_{W^+}}{dy} - \frac{d\sigma_{W^-}}{dy}}{\frac{d\sigma_{W^+}}{dy} + \frac{d\sigma_{W^-}}{dy}} \sim \frac{f_u^p(x_1) f_d^p(x_2) - f_d^p(x_1) f_u^p(x_2)}{f_u^p(x_1) f_d^p(x_2) + f_d^p(x_1) f_u^p(x_2)} \quad (56)$$

Setting  $f_u(x) = f_d(x)R(x)$  we then get:

$$A(y) \sim \frac{R(x_1) - R(x_2)}{R(x_1) + R(x_2)}, \quad (57)$$

which gives an explicit relation between asymmetry and the functional dependence of the  $u(x)/d(x)$  ratio. This ratio is close to 1 at small  $x$ , where the quark distributions arise mostly from sea quarks, and it increases at larger  $x$ , where the valence contribution dominates. At positive  $y$ , where  $x_1 > x_2$ , we therefore expect a positive asymmetry. This is confirmed in the left plot of Fig. 22, showing the asymmetry measured at the Tevatron by the CDF experiment [38], and compared to the NNLO QCD prediction [24, 32, 39, 40] and an estimate of the PDF uncertainty. When measuring the charged lepton



**Fig. 22:** Production [38] (left) and leptonic [41] (right) charge asymmetries of  $W$  bosons in  $p\bar{p}$  collisions at the Tevatron,  $\sqrt{S} = 1.96$  TeV.

from  $W$  decay, the  $W$  production asymmetry is however modulated by the  $W$  decay asymmetry caused by parity violation. The squared amplitude for the process  $f_1 \bar{f}_2 \rightarrow W \rightarrow f_3 \bar{f}_4$  is proportional to  $(p_1 \cdot p_4)(p_2 \cdot p_3)$ , where  $f_{1,3}$  are fermions and  $f_{2,4}$  antifermions, of momenta  $p_{1,\dots,4}$ . In the rest frame of this process, this is proportional to  $(1 + \cos \theta)^2$ , where  $\theta$  is the scattering angle between final- and initial-state fermions. The momentum of the final-state fermion, therefore, points preferentially in the direction of the initial-state fermion's momentum,  $\cos \theta \rightarrow 1$ . For  $d\bar{u} \rightarrow W^- \rightarrow \ell^- \bar{\nu}$  the charged lepton (a fermion) is more likely to move in the direction of  $\ell$  of the  $d$  quark, while for  $u\bar{d} \rightarrow W^+ \rightarrow \ell^+ \nu$  the charged lepton (an antifermion) is more likely to move backward. The rapidity distribution of charged

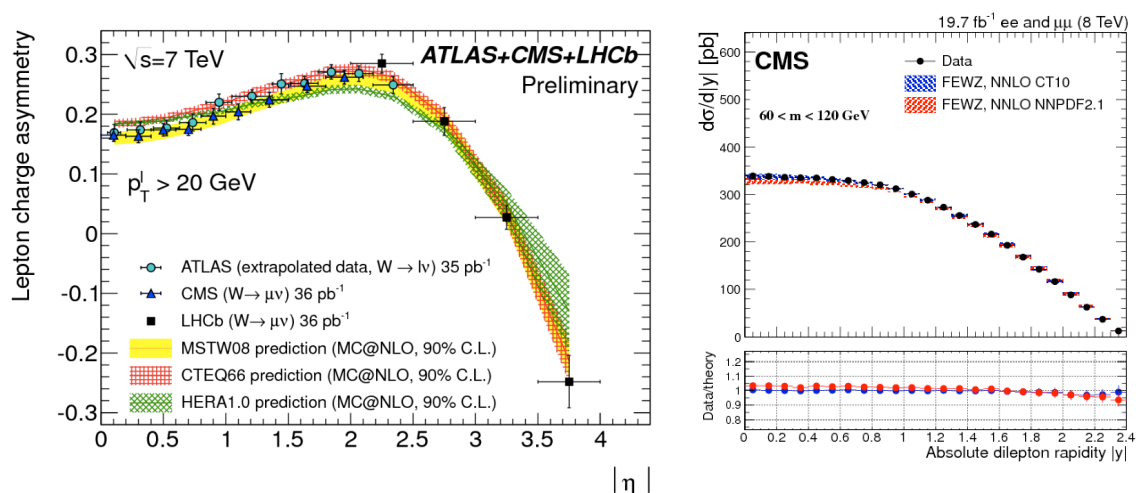


leptons is therefore subject to a tension between the  $W$  production asymmetry, which at positive rapidity favours  $W^+$  over  $W^-$ , and the decay asymmetry, which at positive rapidity favours  $\ell^-$  over  $\ell^+$ . The net result is a distribution that changes sign, becoming negative at large lepton rapidity. This is seen explicitly in the right plot of Fig. 22, from the D0 experiment [41], which also shows the great sensitivity of this quantity to different PDF parameterizations, and the potential to improve their determination.

In  $pp$  collisions, assuming again the dominance of the first generation of quarks and  $f_q^p(x) \gg f_{\bar{q}}^p(x)$  ( $q = u, d$ ) at large  $x$ , the  $W$  charge asymmetry takes the form:<sup>4</sup>

$$A(y) = A(-y) \sim \frac{R(x_{max}) - r(x_{min})}{R(x_{max}) + r(x_{min})}, \quad (58)$$

where  $x_{max(min)} = max(min)(x_1, x_2)$  and  $f_u^p(x) = r(x)f_d^p(x)$ . The extended rapidity coverage of-



**Fig. 23:** Left: leptonic charge asymmetries in  $W$  production at the LHC ( $\sqrt{S} = 7$  TeV), extracted from the measurements of the ATLAS [42], CMS [43] and LHCb [37] experiments. Right:  $Z$  boson rapidity spectrum from CMS [44], compared with NNLO predictions [32].

ferred by the combination of the ATLAS, CMS and LHCb detectors at the LHC, allows to fully exploit the potential of asymmetry measurements as a probe of the proton structure. This is highlighted in the left plot of Fig. 23, which summarizes the LHC experimental results for the lepton charge asymmetry, obtained at  $\sqrt{S} = 7$  TeV, compared to the theoretical predictions based on several sets of PDFs. In particular, notice the large spread of predictions in the largest rapidity regions, spread to be reduced once these data are included as new constraints in global PDF fits (see for example Refs. [45–48]). The PDF sensitivity can be further enhanced by considering the  $W$  asymmetry at large rapidity in events produced in association with a high- $p_T$  jet, as discussed in Ref. [49].

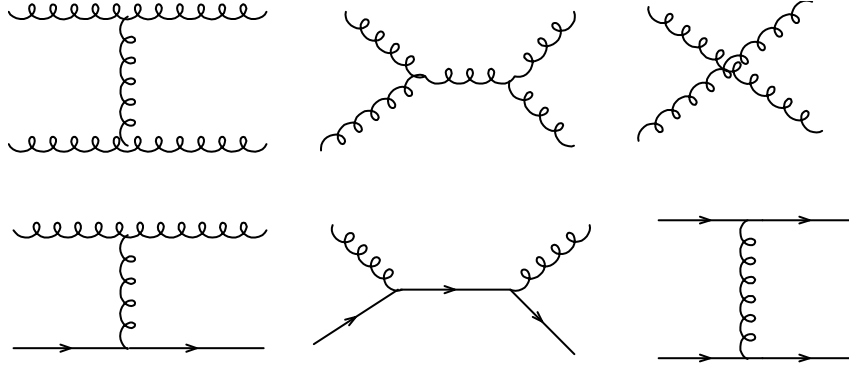
## 4.2 Jet production

Jet production is the hard process with the largest rate in hadronic collisions. For example, the cross-section for producing jets of transverse energy  $E_T \gtrsim 100$  GeV in  $pp$  collisions at the LHC ( $\sqrt{S_{had}} = 14$  TeV) is of the order of a  $\mu\text{b}$ . This means  $\sim 10^4$  events/s at the luminosities available at the LHC. The data collected during the 8 TeV LHC run extend all the way up to  $E_T \gtrsim 2$  TeV, and they will reach

<sup>4</sup>It goes without saying that in actual analyses the contributions of all quark and antiquark flavours are taken into account. At the LHC, in particular, the contribution of strange and charm quarks is significant for the  $W^\pm$  production rate, at the level of  $\sim 30\%$ .



$\sim 4 - 5$  TeV by the end of its operations. These events are generated by collisions among partons that carry over 60% of the available  $pp$  energy, and allow one to probe the shortest distances ever reached. The leading mechanisms for jet production are shown in fig. 24.



**Fig. 24:** Representative diagrams for the production of jet pairs in hadronic collisions

The 2-jet inclusive cross-section can be obtained from the formula

$$d\sigma = \sum_{ijkl} dx_1 dx_2 f_i^{(H_1)}(x_1, \mu) f_j^{(H_2)}(x_2, \mu) \frac{d\hat{\sigma}_{ij \rightarrow k+l}}{d\Phi_2} d\Phi_2 \quad (59)$$

that has to be expressed in terms of the rapidity and transverse momentum of the quarks (or jets), in order to make contact with physical reality. The two-particle phase space is given by

$$d\Phi_2 = \frac{d^3k}{2k^0(2\pi)^3} 2\pi \delta((p_1 + p_2 - k)^2), \quad (60)$$

and, in the centre-of-mass of the colliding partons, we get

$$d\Phi_2 = \frac{1}{2(2\pi)^2} d^2k_T dy 2\delta(\hat{s} - 4(k^0)^2), \quad (61)$$

where  $k_T$  is the transverse momentum of the final-state partons. Here  $y$  is the rapidity of the produced parton in the parton centre-of-mass frame. It is given by

$$y = \frac{y_1 - y_2}{2} \quad (62)$$

where  $y_1$  and  $y_2$  are the rapidities of the produced partons in the laboratory frame (in fact, in any frame). One also introduces

$$y_0 = \frac{y_1 + y_2}{2} = \frac{1}{2} \log \frac{x_1}{x_2}, \quad \tau = \frac{\hat{s}}{S_{had}} = x_1 x_2. \quad (63)$$

We have

$$dx_1 dx_2 = dy_0 d\tau. \quad (64)$$

We obtain

$$d\sigma = \sum_{ijkl} dy_0 \frac{1}{S_{had}} f_i^{(H_1)}(x_1, \mu) f_j^{(H_2)}(x_2, \mu) \frac{d\hat{\sigma}_{ij \rightarrow k+l}}{d\Phi_2} \frac{1}{2(2\pi)^2} 2 dy d^2k_T \quad (65)$$

which can also be written as

$$\frac{d\sigma}{dy_1 dy_2 d^2k_T} = \frac{1}{S_{had} 2(2\pi)^2} \sum_{ijkl} f_i^{(H_1)}(x_1, \mu) f_j^{(H_2)}(x_2, \mu) \frac{d\hat{\sigma}_{ij \rightarrow k+l}}{d\Phi_2}. \quad (66)$$

The variables  $x_1, x_2$  can be obtained from  $y_1, y_2$  and  $k_T$  from the equations

$$y_0 = \frac{y_1 + y_2}{2} \quad (67)$$

$$y = \frac{y_1 - y_2}{2} \quad (68)$$

$$x_T = \frac{2k_T}{\sqrt{S_{had}}} \quad (69)$$

$$x_1 = x_T e^{y_0} \cosh y \quad (70)$$

$$x_2 = x_T e^{-y_0} \cosh y. \quad (71)$$

For the partonic variables, we need  $\hat{s}$  and the scattering angle in the parton centre-of-mass frame  $\theta$ , since

$$t = -\frac{\hat{s}}{2} (1 - \cos \theta), \quad u = -\frac{\hat{s}}{2} (1 + \cos \theta). \quad (72)$$

Neglecting the parton masses, you can show that the rapidity can also be written as:

$$y = -\log \tan \frac{\theta}{2} \equiv \eta, \quad (73)$$

with  $\eta$  usually being referred to as pseudorapidity.

The leading-order Born cross-sections for parton-parton scattering are reported in Table 2.

**Table 2:** Cross-sections for light parton scattering. The notation is  $p_1 p_2 \rightarrow kl$ ,  $\hat{s} = (p_1 + p_2)^2$ ,  $\hat{t} = (p_1 - k)^2$ ,  $\hat{u} = (p_1 - l)^2$ .

Process	$\frac{d\hat{\sigma}}{d\Phi_2}$
$qq' \rightarrow qq'$	$\frac{1}{2\hat{s}} \frac{4}{9} \frac{\hat{s}^2 + \hat{u}^2}{\hat{t}^2}$
$qq \rightarrow qq$	$\frac{1}{2} \frac{1}{2\hat{s}} \left[ \frac{4}{9} \left( \frac{\hat{s}^2 + \hat{u}^2}{\hat{t}^2} + \frac{\hat{s}^2 + \hat{t}^2}{\hat{u}^2} \right) - \frac{8}{27} \frac{\hat{s}^2}{\hat{u}\hat{t}} \right]$
$q\bar{q} \rightarrow q'\bar{q}'$	$\frac{1}{2\hat{s}} \frac{4}{9} \frac{\hat{t}^2 + \hat{u}^2}{\hat{s}^2}$
$q\bar{q} \rightarrow q\bar{q}$	$\frac{1}{2\hat{s}} \left[ \frac{4}{9} \left( \frac{\hat{s}^2 + \hat{u}^2}{\hat{t}^2} + \frac{\hat{t}^2 + \hat{u}^2}{\hat{s}^2} \right) - \frac{8}{27} \frac{\hat{u}^2}{\hat{s}\hat{t}} \right]$
$q\bar{q} \rightarrow gg$	$\frac{1}{2} \frac{1}{2\hat{s}} \left[ \frac{32}{27} \frac{\hat{t}^2 + \hat{u}^2}{\hat{t}\hat{u}} - \frac{8}{3} \frac{\hat{t}^2 + \hat{u}^2}{\hat{s}^2} \right]$
$gg \rightarrow q\bar{q}$	$\frac{1}{2\hat{s}} \left[ \frac{1}{6} \frac{\hat{t}^2 + \hat{u}^2}{\hat{t}\hat{u}} - \frac{3}{8} \frac{\hat{t}^2 + \hat{u}^2}{\hat{s}^2} \right]$
$gg \rightarrow qq$	$\frac{1}{2\hat{s}} \left[ -\frac{4}{9} \frac{\hat{s}^2 + \hat{u}^2}{\hat{s}\hat{u}} + \frac{\hat{u}^2 + \hat{s}^2}{\hat{t}^2} \right]$
$gg \rightarrow gg$	$\frac{1}{2} \frac{1}{2\hat{s}} \frac{9}{2} \left( 3 - \frac{\hat{t}\hat{u}}{\hat{s}^2} - \frac{\hat{s}\hat{u}}{\hat{t}^2} - \frac{\hat{s}\hat{t}}{\hat{u}^2} \right)$

It is interesting to note that a good approximation to the exact results can easily be obtained by using the soft-gluon techniques introduced in the third lecture. Based on the fact that even at  $90^\circ$   $\min(|t|, |u|)$  does not exceed  $s/2$ , and that therefore everything else being equal a propagator in the  $t$  or  $u$  channel contributes to the square of an amplitude 4 times more than a propagator in the  $s$  channel, it is reasonable to assume that the amplitudes are dominated by the diagrams with a gluon exchanged in the  $t$  (or  $u$ ) channel. It is easy to calculate the amplitudes in this limit using the soft-gluon approximation. For example, the amplitude for the exchange of a soft gluon among a  $qq'$  pair is given by:

$$(\lambda_{ij}^a) (\lambda_{kl}^a) 2p_\mu \frac{1}{t} 2p'_\mu = \lambda_{ij}^a \lambda_{kl}^a \frac{4p \cdot p'}{t} = \frac{2s}{t} \lambda_{ij}^a \lambda_{kl}^a. \quad (74)$$

The  $p_\mu$  and  $p'_\mu$  factors represent the coupling of the exchanged gluon to the  $q$  and  $q'$  quark lines, respectively (see Eq. (34)). Squaring, and summing and averaging over spins and colours, gives

$$\overline{\sum_{\text{colours,spin}} |M_{qq'}|^2} = \frac{1}{N^2} \left( \frac{N^2 - 1}{4} \right) \frac{4s^2}{t^2} = \frac{8}{9} \frac{s^2}{t^2}. \quad (75)$$

Since for this process the diagram with a  $t$ -channel gluon exchange is symmetric for  $s \leftrightarrow u$  exchange, and since  $u \rightarrow -s$  in the  $t \rightarrow 0$  limit, the above result can be rewritten in an explicitly  $(s, u)$  symmetric way as

$$\frac{4}{9} \frac{s^2 + u^2}{t^2}, \quad (76)$$

which indeed exactly agrees with the result of the exact calculation, as given in Table 2. The corrections which appear from  $s$  or  $u$  gluon exchange when the quark flavours are the same or when we study a  $q\bar{q}$  process are small, as can be seen by comparing the above result to the expressions in the table.

As another example we consider the case of  $qg \rightarrow qg$  scattering. The amplitude will be exactly the same as in the  $qq' \rightarrow qq'$  case, up to the different colour factors. A simple calculation then gives:

$$\overline{\sum_{\text{colours,spin}} |M_{qg}|^2} = \frac{9}{4} \overline{\sum |M_{qq'}|^2} = \frac{s^2 + u^2}{t^2}.$$

The exact result is

$$\frac{u^2 + s^2}{t^2} - \frac{4}{9} \frac{u^2 + s^2}{us}, \quad (77)$$

which even at  $90^\circ$ , the point where the  $t$ -channel exchange approximation is worse, only differs from this latter by no more than 25%.

As a final example we consider the case of  $gg \rightarrow gg$  scattering, which in our approximation gives:

$$\overline{\sum |M_{gg}|^2} = \frac{9}{2} \frac{s^2}{t^2}. \quad (78)$$

By  $u \leftrightarrow t$  symmetry we should expect the simple improvement:

$$\overline{\sum |M_{gg}|^2} \sim \frac{9}{2} \left( \frac{s^2}{t^2} + \frac{s^2}{u^2} \right). \quad (79)$$

This only differs by 20% from the exact result at  $90^\circ$ .

Notice that at small  $t$  the following relation holds:

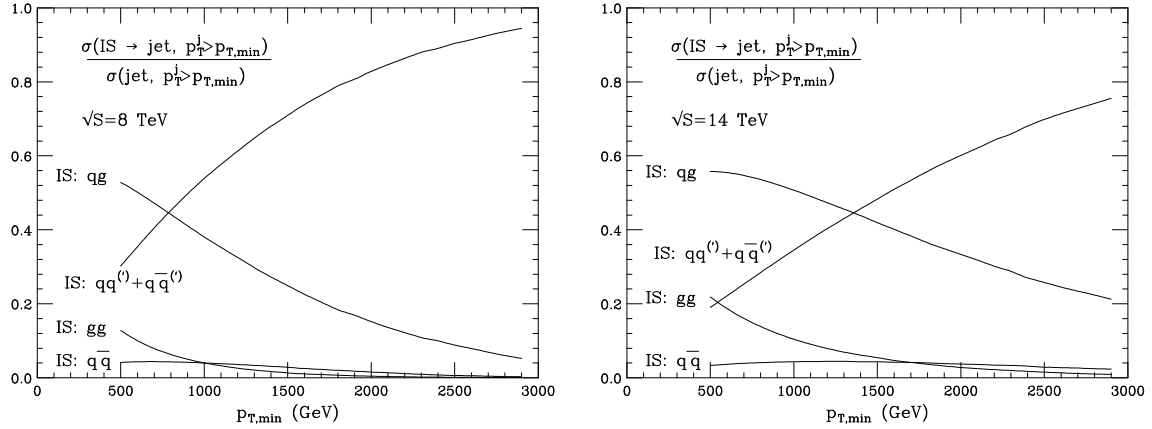
$$\hat{\sigma}_{gg} : \hat{\sigma}_{qg} : \hat{\sigma}_{q\bar{q}} = \left( \frac{9}{4} \right) : 1 : \left( \frac{4}{9} \right). \quad (80)$$

The  $9/4$  factors are simply the ratios of the colour factors for the coupling to gluons of a gluon ( $C_A$ ) and of a quark ( $T_F$ ), after including the respective colour-average factors ( $1/(N^2 - 1)$  for the gluon, and  $1/N$  for the quark). Using Eq. (80), we can then write:

$$\begin{aligned} d\sigma_{hadr} &= \int dx_1 dx_2 \sum_{i,j} f_i(x_1) f_j(x_2) d\hat{\sigma}_{ij} \\ &= \int dx_1 dx_2 F(x_1) F(x_2) d\hat{\sigma}_{gg}(gg \rightarrow \text{jets}) \end{aligned} \quad (81)$$

where the object

$$F(x) = f_g(x) + \frac{4}{9} \sum_f [q_f(x) + \bar{q}_f(x)] \quad (82)$$

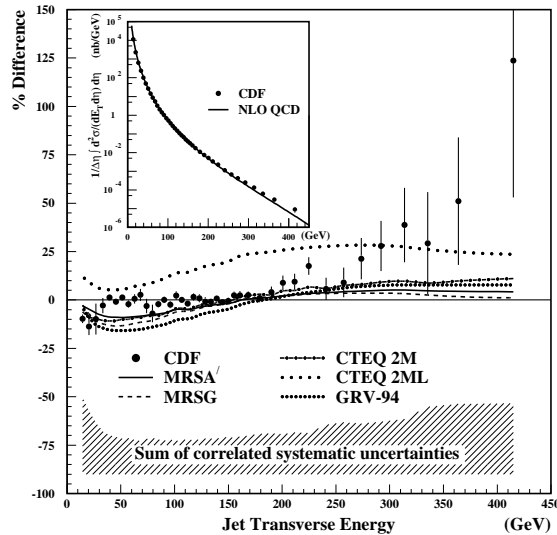


**Fig. 25:** Relative contribution to the jet- $E_T$  rates from the different production channels, in  $pp$  collisions at 8 and 14 TeV.

is usually called the *effective structure function*. This result indicates that the measurement of the inclusive jet cross-section does not allow in principle to disentangle the independent contribution of the various partonic components of the proton, unless of course one is considering a kinematical region where the production is dominated by a single process. The relative contributions of the different channels, calculated using current fits of parton densities, are shown for different center of mass energies at the LHC in fig. 25.

### 4.3 Jet $E_T$ spectra: comparison of theory and experimental data

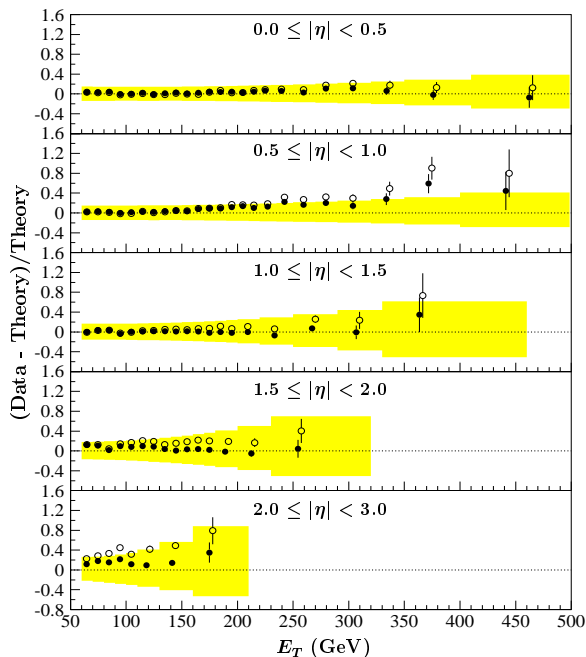
Predictions for jet production at colliders are available today at the next-to-leading order in QCD (see the review in Ref. [8]). One of the preferred observables is the inclusive  $E_T$  spectrum. An accurate



**Fig. 26:** Comparison of the data vs theory in an early measurement of the jet cross-section at the Tevatron, by the CDF experiment [51]

comparison of data and theory, should it exhibit discrepancies at the largest values of  $E_T$ , could pro-

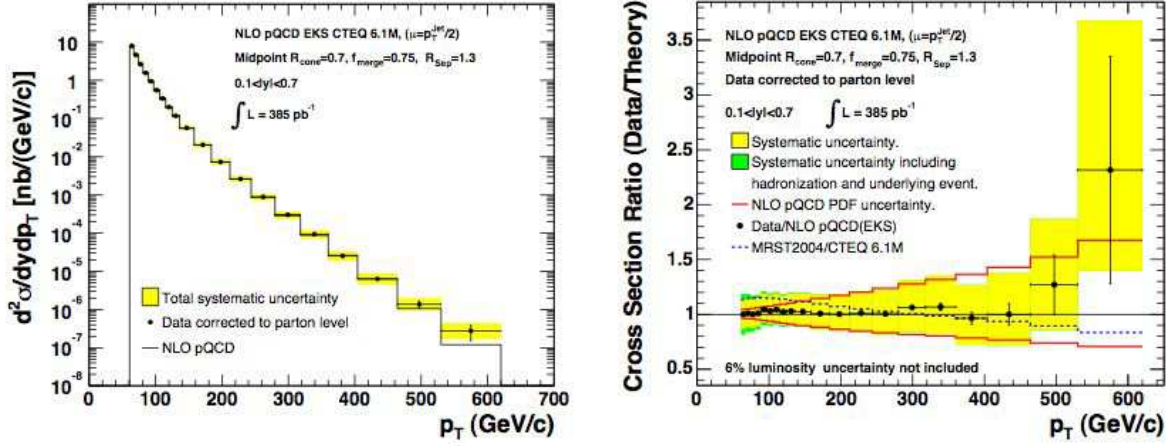
vide evidence for new phenomena, such as the existence of a quark substructure. For years it has been known [50] that an underlying quark compositeness would increase the rate of the highest- $E_T$  jets. The real question, therefore, is how do we convince ourselves that the prediction is, indeed, *accurate*. This question became particularly relevant in 1995, when CDF measured a jet cross-section that appeared to deviate from theory in precisely the way predicted by an underlying quark compositeness (see fig. 26). How do we know that this is not due to poorly known quark or gluon densities at large  $x$ ? In principle one could incorporate the CDF jet data into a global fit to the partonic PDFs, and verify whether it is possible to modify them so as to maintain agreement with the other data, and at the same time to also fit satisfactorily the jet data themselves. On the other hand, doing this would prevent us from using the jet spectrum as a probe of new physics. In other words, we might be hiding away a possible signal of new physics by ascribing it to the PDFs. Is it possible to have a complementary determination of



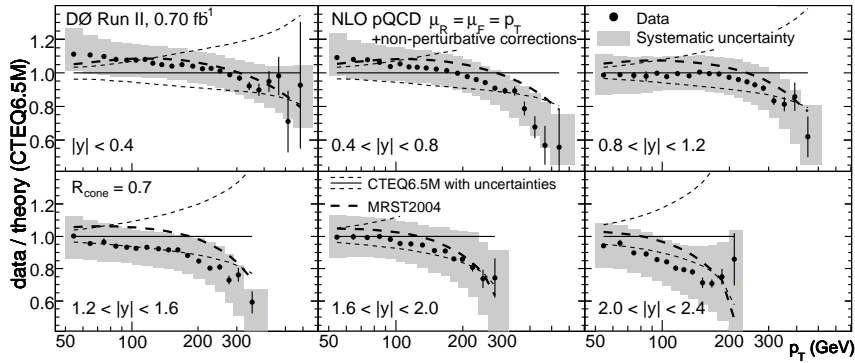
**Fig. 27:** Inclusive  $E_T$  spectra for jets in different rapidity regions, as measured at the run 1 of the Tevatron by the D0 Collaboration [52].

the PDF at high- $x$ , that could constrain the possible PDF systematics of the jet cross-section and at the same time leave the high- $E_T$  tail as an independent and usable observable? This is indeed possible, by fully exploiting the kinematics of dijet production and the wide rapidity coverage of the collider detectors. One could in fact consider final states where the dijet system is highly boosted in the forward or backward region. For example, one could consider cases where  $x_1 \rightarrow 1$  and  $x_2 \ll 1$ . In this case, the invariant mass of the dijet system would be small (since  $M_{jj}^2 = x_1 x_2 S \ll S$ ), and we know from lower-energy measurements that at this scale jets must behave like pointlike particles, following exactly the QCD-predicted rate. These final states are characterized by having jets at large positive rapidity. One can therefore perform a measurement with forward jets, and use these data to fit the  $x_1 \rightarrow 1$  behaviour of the quark and gluon PDFs without the risk of washing away possible new-physics effects. At that point, the large- $x$  PDFs thus constrained can be safely applied to the kinematical configurations where both  $x_1$  and  $x_2$  are large, namely the highest- $E_T$  final states, and, if any residual discrepancy between data and theory is observed, infer the possible presence of new physics.

In the case of the Tevatron data, the study of the forward-jet configurations was performed by D0 [52]. Figure 27, from their work, shows the comparison between data and theory for different jet-rapidity intervals. Two different PDF sets are used, CTEQ4M, and CTEQ4HJ [53], the latter having



**Fig. 28:** Inclusive  $E_T$  spectra for central jets, as measured by the CDF experiment at the Tevatron [54], compared to NLO QCD calculations



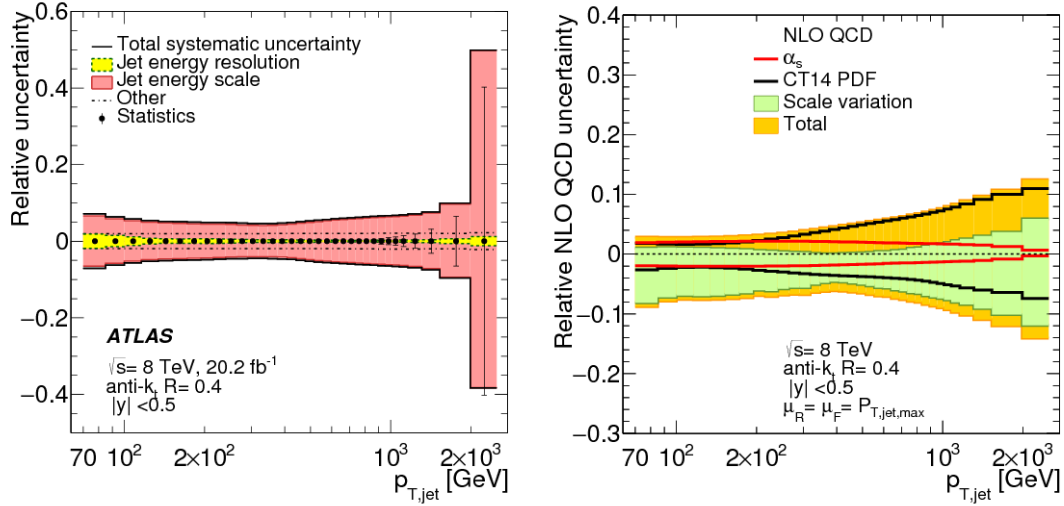
**Fig. 29:** Comparison of run 2 inclusive jets cross-sections at D0 [55] with QCD calculations. The dashed lines represent the systematics band due to PDF uncertainties

been tuned to describe the CDF high- $E_T$  jet tail. Notice the good overall agreement of this prediction for the whole set of rapidities. After this tuning, the residual discrepancy between the CDF high- $E_T$  data and QCD is within the theoretical and experimental systematic uncertainties, confirming that jets behave as expected in the Standard Model. This conclusion has been strengthened by the analysis of the run 2 data [54, 55], at  $\sqrt{S} = 1.96 \text{ TeV}$ , as shown in Figs. 28 and 29.

#### 4.4 Jets at the LHC

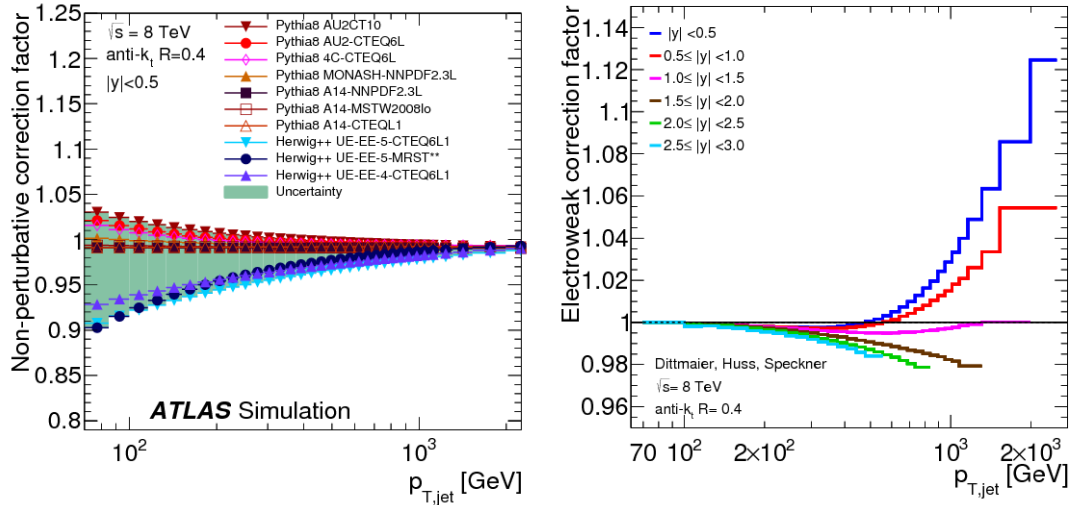
The huge statistics and energy lever arm available at the LHC is pushing even further the precision and the sensitivity of jet data. This progress is proceeding hand in hand with the improved theoretical calculations, which have now reached the level of NNLO precision [56].

The experimental systematics, shown in Fig. 30 for the recent final analysis of the ATLAS 8 TeV data [57], are today reduced to less than 10% for the whole range of jet transverse energies up to  $\sim 2 \text{ TeV}$ , and are dominated by the knowledge of the absolute jet energy scale. The higher statistics available in run 2 of the LHC, and beyond, will allow a further reduction of this systematics, using several experimental handles such as the energy balancing between prompt photons and jets. The theoretical uncertainties include several components. Those of purely perturbative nature are shown in the right plot of Fig. 30, and include the scale,  $\alpha_s$  and PDF uncertainties. The latter dominate at the largest  $E_T$  values and, once



**Fig. 30:** Experimental (left) and theoretical NLO (right) systematics in the measurement and prediction of the inclusive jet cross section at  $\sqrt{S} = 8$  TeV [57].

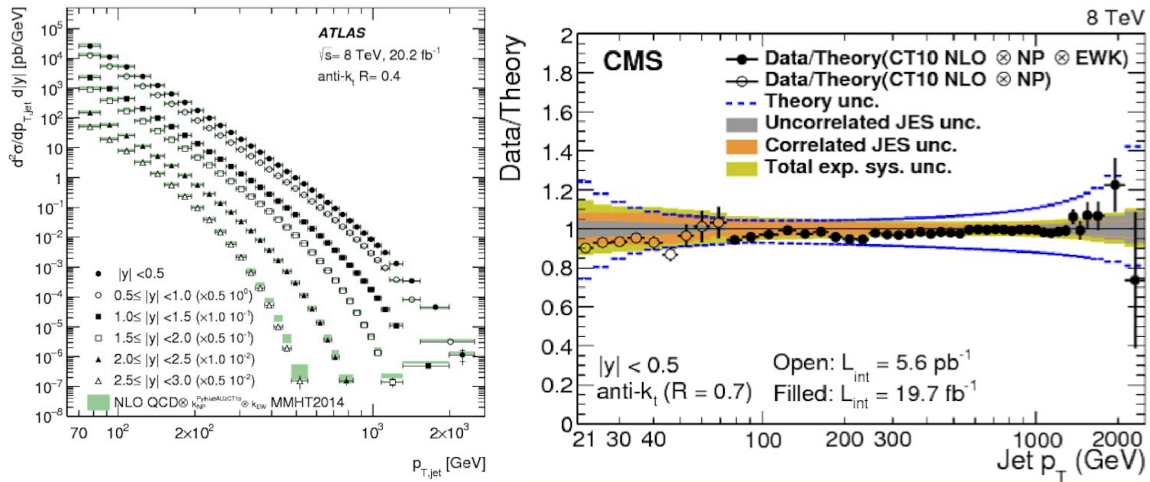
again, will likely be reduced with future PDF studies at higher luminosity, along the lines discussed in the previous section in the context of the Tevatron studies.



**Fig. 31:** Hadronization systematics (left) and EW NLO (right) corrections in the theoretical prediction of the inclusive jet cross section at  $\sqrt{S} = 8$  TeV [57].

Further theoretical systematics are related to the non-perturbative corrections needed to translate the jet energy, defined at the perturbative level by the partons, to the actual energy after partons shower and evolve into hadrons. These effects, shown in the left plot of Fig. 31, are enhanced at the lowest energies, and quickly vanish in the interesting multi-TeV domain. At the highest energies, finally, electroweak (EW) corrections become important, and must be included in the calculations. Their size can reach the 10% value above the TeV, as shown in the right plot of Fig. 31.

A comparison of data and the NLO theory (corrected for hadronization effects) is shown for the ATLAS measurement [57] in the left plot of Fig. 32. The data extend to  $E_T$  above 2 TeV, covering 10



**Fig. 32:** ATLAS jet cross section data at  $\sqrt{S} = 8$  TeV [57], compared to theoretical calculations (left). Ratio of data and theory for the 8 TeV measurement by CMS [58].

orders of magnitude in rate. The level of agreement is highlighted in the right plot, using the CMS results of the 8 TeV data sample [58]. The agreement is well within the  $\pm 10\%$  overall systematics.

### Acknowledgements

It is a pleasure to thank the organizers of this School for the successful efforts made to bring together top-quality students, in a fantastic environment.



## References

- [1] D.J. Gross, hep-th/9809060.
- [2] G. 't Hooft, hep-th/9808154.
- [3] R.P. Feynman, Photon–Hadron Interactions (W.A. Benjamin, New York, 1972).
- [4] T. Muta, Foundations of QCD (World Scientific, Singapore, 1998).
- [5] M.E. Peskin and D.V. Schroeder, An Introduction to Quantum Field Theory (Addison-Wesley, Reading, MA, 1995).
- [6] V. Barger and R.J.N. Phillips, Collider Physics (Addison-Wesley, Redwood City, CA, 1997).
- [7] R.K. Ellis, W.J. Stirling and B.R. Webber, QCD and Collider Physics (Cambridge University Press, 1996).
- [8] J. M. Campbell, J. W. Huston and W. J. Stirling, Rep. Prog. Phys. **70** (2007) 89 [arXiv:hep-ph/0611148].
- [9] J. Butterworth *et al.*, J. Phys. G **43** (2016) 023001 doi:10.1088/0954-3899/43/2/023001 [arXiv:1510.03865 [hep-ph]].
- [10] V. Bertone, S. Carrazza and J. Rojo, Comput. Phys. Commun. **185** (2014) 1647 doi:10.1016/j.cpc.2014.03.007 [arXiv:1310.1394 [hep-ph]].
- [11] S. Carrazza, A. Ferrara, D. Palazzo and J. Rojo, J. Phys. G **42** (2015) no.5, 057001 doi:10.1088/0954-3899/42/5/057001 [arXiv:1410.5456 [hep-ph]].
- [12] R. Field, Applications of Perturbative QCD (Addison-Wesley, Redwood City, CA, 1989).
- [13] Yu.L. Dokshitzer, V.A. Khoze, A.H. Mueller and S.I. Troyan, Basics of Perturbative QCD (Editions Frontières, Gif-sur-Yvette, 1991).
- [14] V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15** (1972) 438 [Yad. Fiz. **15** (1972) 781].
- [15] G. Altarelli and G. Parisi, Nucl. Phys. B **126** (1977) 298.
- [16] Y. L. Dokshitzer, Sov. Phys. JETP **46** (1977) 641 [Zh. Eksp. Teor. Fiz. **73** (1977) 1216].
- [17] B. R. Webber, Nucl. Phys. B **238** (1984) 492.
- [18] B. Andersson, G. Gustafson, G. Ingelman and T. Sjostrand, Phys. Rep. **97** (1983) 31.
- [19] S. Gieseke, A. Ribon, M. H. Seymour, P. Stephens and B. Webber, JHEP **0402** (2004) 005 [arXiv:hep-ph/0311208].
- [20] R. Hamberg, W. L. van Neerven and T. Matsuura, Nucl. Phys. B **359** (1991) 343 [Nucl. Phys. B **644** (2002) 403].
- [21] R. V. Harlander and W. B. Kilgore, Phys. Rev. Lett. **88** (2002) 201801 [hep-ph/0201206].
- [22] C. Anastasiou, L. J. Dixon, K. Melnikov and F. Petriello, Phys. Rev. D **69** (2004) 094008 [hep-ph/0312266].
- [23] K. Melnikov and F. Petriello, Phys. Rev. D **74** (2006) 114017 [hep-ph/0609070].
- [24] S. Catani, L. Cieri, G. Ferrera, D. de Florian and M. Grazzini, Phys. Rev. Lett. **103** (2009) 082001 [arXiv:0903.2120 [hep-ph]].
- [25] R. Gavin, Y. Li, F. Petriello and S. Quackenbush, Comput. Phys. Commun. **182** (2011) 2388 [arXiv:1011.3540 [hep-ph]].
- [26] A. Karlberg, E. Re and G. Zanderighi, JHEP **1409** (2014) 134 [arXiv:1407.2940 [hep-ph]].
- [27] S. Höche, Y. Li and S. Prestel, Phys. Rev. D **91** (2015) 7, 074015 [arXiv:1405.3607 [hep-ph]].
- [28] S. Alioli, C. W. Bauer, C. Berggren, F. J. Tackmann and J. R. Walsh, Phys. Rev. D **92** (2015) 9, 094020 [arXiv:1508.01475 [hep-ph]].
- [29] S. Dittmaier and M. Kramer, Phys. Rev. D **65** (2002) 073007 [hep-ph/0109062].
- [30] U. Baur and D. Wackerth, Phys. Rev. D **70** (2004) 073015 [hep-ph/0405191].
- [31] U. Baur, O. Brein, W. Hollik, C. Schappacher and D. Wackerth, Phys. Rev. D **65** (2002) 033007

- [hep-ph/0108274].
- [32] Y. Li and F. Petriello, Phys. Rev. D **86** (2012) 094034 [arXiv:1208.5967 [hep-ph]].
- [33] S. Dittmaier, A. Huss and C. Schwinn, arXiv:1511.08016 [hep-ph].
- [34] S. Forte and G. Watt, Ann. Rev. Nucl. Part. Sci. **63** (2013) 291 [arXiv:1301.6754 [hep-ph]].
- [35] G. Aad *et al.* [ATLAS Collaboration], Phys. Rev. D **85** (2012) 072004 [arXiv:1109.5141 [hep-ex]].
- [36] S. Chatrchyan *et al.* [CMS Collaboration], JHEP **1110** (2011) 132 [arXiv:1107.4789 [hep-ex]].
- [37] R. Aaij *et al.* [LHCb Collaboration], JHEP **1412** (2014) 079 [arXiv:1408.4354 [hep-ex]].
- [38] T. Aaltonen *et al.* [CDF Collaboration], Phys. Rev. Lett. **102** (2009) 181801 [arXiv:0901.2169 [hep-ex]].
- [39] S. Catani, G. Ferrera and M. Grazzini, JHEP **1005** (2010) 006 [arXiv:1002.3115 [hep-ph]].
- [40] R. Gavin, Y. Li, F. Petriello and S. Quackenbush, Comput. Phys. Commun. **184** (2013) 208 [arXiv:1201.5896 [hep-ph]].
- [41] V. M. Abazov *et al.* [D0 Collaboration], Phys. Rev. D **91** (2015) 3, 032007 [Phys. Rev. D **91** (2015) 7, 079901] [arXiv:1412.2862 [hep-ex]].
- [42] G. Aad *et al.* [ATLAS Collaboration], Phys. Lett. B **701** (2011) 31 [arXiv:1103.2929 [hep-ex]].
- [43] S. Chatrchyan *et al.* [CMS Collaboration], Phys. Rev. D **90** (2014) 3, 032004 [arXiv:1312.6283 [hep-ex]].
- [44] V. Khachatryan *et al.* [CMS Collaboration], Eur. Phys. J. C **75** (2015) 4, 147 [arXiv:1412.1115 [hep-ex]].
- [45] S. Alekhin, J. Bluemlein and S. Moch, Phys. Rev. D **89** (2014) 5, 054028 [arXiv:1310.3059 [hep-ph]].
- [46] R. D. Ball *et al.* [NNPDF Collaboration], JHEP **1504** (2015) 040 [arXiv:1410.8849 [hep-ph]].
- [47] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne, Eur. Phys. J. C **75** (2015) 5, 204 [arXiv:1412.3989 [hep-ph]].
- [48] S. Dulat *et al.*, arXiv:1506.07443 [hep-ph].
- [49] S. Farry and R. Gauld, arXiv:1505.01399 [hep-ph].
- [50] E. Eichten, K. D. Lane and M. E. Peskin, Phys. Rev. Lett. **50** (1983) 811.
- [51] F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **77** (1996) 438 [arXiv:hep-ex/9601008].
- [52] B. Abbott *et al.* [D0 Collaboration], Phys. Rev. Lett. **86** (2001) 1707 [arXiv:hep-ex/0011036].
- [53] H. L. Lai *et al.*, Phys. Rev. D **55** (1997) 1280 [arXiv:hep-ph/9606399].
- [54] A. Abulencia *et al.* [CDF Collaboration], Phys. Rev. D **74** (2006) 071103 [arXiv:hep-ex/0512020].
- [55] V. M. Abazov *et al.* [D0 Collaboration], Phys. Rev. Lett. **101** (2008) 062001 [arXiv:0802.2400 [hep-ex]].
- [56] J. Currie, A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, A. Huss and J. Pires, Phys. Rev. Lett. **119** (2017) no.15, 152001 doi:10.1103/PhysRevLett.119.152001 [arXiv:1705.10271 [hep-ph]].
- [57] M. Aaboud *et al.* [ATLAS Collaboration], JHEP **1709** (2017) 020 doi:10.1007/JHEP09(2017)020 [arXiv:1706.03192 [hep-ex]].
- [58] V. Khachatryan *et al.* [CMS Collaboration], JHEP **1703** (2017) 156 doi:10.1007/JHEP03(2017)156 [arXiv:1609.05331 [hep-ex]].

# Flavour Dynamics and Violations of the CP Symmetry

Antonio Pich

IFIC, University of València – CSIC, E-46071 València, Spain

## Abstract

An overview of flavour physics and CP-violating phenomena is presented. The Standard Model quark-mixing mechanism is discussed in detail and its many successful experimental tests are summarized. Flavour-changing transitions put very stringent constraints on new-physics scenarios beyond the Standard Model framework. Special attention is given to the empirical evidences of  $\mathcal{CP}$  violation and their important role in our understanding of flavour dynamics. The current status of the so-called flavour anomalies is also reviewed.

## Keywords

Flavour physics; quark mixing; CP violation; electroweak interactions.

## 1 Fermion families

We have learnt experimentally that there are six different quark flavours  $u, d, s, c, b, t$ , three different charged leptons  $e, \mu, \tau$  and their corresponding neutrinos  $\nu_e, \nu_\mu, \nu_\tau$ . We can include all these particles into the  $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$  Standard Model (SM) framework [1–3], by organizing them into three families of quarks and leptons:

$$\begin{bmatrix} \nu_e & u \\ e^- & d' \end{bmatrix}, \quad \begin{bmatrix} \nu_\mu & c \\ \mu^- & s' \end{bmatrix}, \quad \begin{bmatrix} \nu_\tau & t \\ \tau^- & b' \end{bmatrix}, \quad (1)$$

where (each quark appears in three different colours)

$$\begin{bmatrix} \nu_i & u_i \\ \ell_i^- & d'_i \end{bmatrix} \equiv \begin{pmatrix} \nu_i \\ \ell_i^- \end{pmatrix}_L, \quad \begin{pmatrix} u_i \\ d'_i \end{pmatrix}_L, \quad \ell_{iR}^-, \quad u_{iR}, \quad d'_{iR}, \quad (2)$$

plus the corresponding antiparticles. Thus, the left-handed fields are  $SU(2)_L$  doublets, while their right-handed partners transform as  $SU(2)_L$  singlets. The three fermionic families appear to have identical properties (gauge interactions); they differ only by their mass and their flavour quantum numbers.

The fermionic couplings of the photon and the  $Z$  boson are flavour conserving, i.e., the neutral gauge bosons couple to a fermion and its corresponding antifermion. In contrast, the  $W^\pm$  bosons couple any up-type quark with all down-type quarks because the weak doublet partner of  $u_i$  turns out to be a quantum superposition of down-type mass eigenstates:  $d'_i = \sum_j \mathbf{V}_{ij} d_j$ . This flavour mixing generates a rich variety of observable phenomena, including  $\mathcal{CP}$ -violation effects, which can be described in a very successful way within the SM [4, 5].

In spite of its enormous phenomenological success, The SM does not provide any real understanding of flavour. We do not know yet why fermions are replicated in three (and only three) nearly identical copies. Why the pattern of masses and mixings is what it is? Are the masses the only difference among the three families? What is the origin of the SM flavour structure? Which dynamics is responsible for the observed  $\mathcal{CP}$  violation? The fermionic flavour is the main source of arbitrary free parameters in the SM: 9 fermion masses, 3 mixing angles and 1 complex phase, for massless neutrinos. 7 (9) additional parameters arise with non-zero Dirac (Majorana) neutrino masses: 3 masses, 3 mixing angles and 1 (3) phases. The problem of fermion mass generation is deeply related with the mechanism responsible for the electroweak Spontaneous Symmetry Breaking (SSB). Thus, the origin of these parameters lies in the



**Fig. 1:** Flavour-changing transitions through the charged-current couplings of the  $W^\pm$  bosons.

most obscure part of the SM Lagrangian: the scalar sector. Clearly, the dynamics of flavour appears to be “terra incognita” which deserves a careful investigation.

The following sections contain a short overview of the quark flavour sector and its present phenomenological status. The most relevant experimental tests are briefly described. A more pedagogic introduction to the SM can be found in Ref. [4].

## 2 Flavour structure of the Standard Model

In the SM flavour-changing transitions occur only in the charged-current sector (Fig. 1):

$$\mathcal{L}_{\text{CC}} = -\frac{g}{2\sqrt{2}} \left\{ W_\mu^\dagger \left[ \sum_{ij} \bar{u}_i \gamma^\mu (1 - \gamma_5) \mathbf{V}_{ij} d_j + \sum_\ell \bar{\nu}_\ell \gamma^\mu (1 - \gamma_5) \ell \right] + \text{h.c.} \right\}. \quad (3)$$

The so-called Cabibbo–Kobayashi–Maskawa (CKM) matrix  $\mathbf{V}$  [6, 7] is generated by the same Yukawa couplings giving rise to the quark masses. Before SSB, there is no mixing among the different quarks, i.e.,  $\mathbf{V} = \mathbf{I}$ . In order to understand the origin of the matrix  $\mathbf{V}$ , let us consider the general case of  $N_G$  generations of fermions, and denote  $\nu'_j, \ell'_j, u'_j, d'_j$  the members of the weak family  $j$  ( $j = 1, \dots, N_G$ ), with definite transformation properties under the gauge group. Owing to the fermion replication, a large variety of fermion-scalar couplings are allowed by the gauge symmetry. The most general Yukawa Lagrangian has the form

$$\begin{aligned} \mathcal{L}_Y = & - \sum_{jk} \left\{ (\bar{u}'_j, \bar{d}'_j)_L \left[ c_{jk}^{(d)} \begin{pmatrix} \phi^{(+)} \\ \phi^{(0)} \end{pmatrix} d'_{kR} + c_{jk}^{(u)} \begin{pmatrix} \phi^{(0)*} \\ -\phi^{(-)} \end{pmatrix} u'_{kR} \right] \right. \\ & \left. + (\bar{\nu}'_j, \bar{\ell}'_j)_L c_{jk}^{(\ell)} \begin{pmatrix} \phi^{(+)} \\ \phi^{(0)} \end{pmatrix} \ell'_{kR} \right\} + \text{h.c.}, \end{aligned} \quad (4)$$

where  $\phi^T(x) \equiv (\phi^{(+)}, \phi^{(0)})$  is the SM scalar doublet and  $c_{jk}^{(d)}$ ,  $c_{jk}^{(u)}$  and  $c_{jk}^{(\ell)}$  are arbitrary coupling constants. The second term involves the  $\mathcal{C}$ -conjugate scalar field  $\phi^c(x) \equiv i\sigma_2 \phi^*(x)$ .

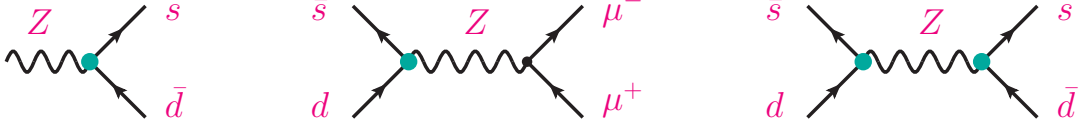
In the unitary gauge  $\phi^T(x) \equiv \frac{1}{\sqrt{2}} (0, v + H)$ , where  $v$  is the electroweak vacuum expectation value and  $H(x)$  the Higgs field. The Yukawa Lagrangian can then be written as

$$\mathcal{L}_Y = - \left( 1 + \frac{H}{v} \right) \{ \bar{\mathbf{d}}'_L \mathbf{M}'_d \mathbf{d}'_R + \bar{\mathbf{u}}'_L \mathbf{M}'_u \mathbf{u}'_R + \bar{\mathbf{\ell}}'_L \mathbf{M}'_\ell \mathbf{\ell}'_R + \text{h.c.} \}. \quad (5)$$

Here,  $\mathbf{d}'$ ,  $\mathbf{u}'$  and  $\mathbf{\ell}'$  denote vectors in the  $N_G$ -dimensional flavour space, with components  $d'_j, u'_j$  and  $\ell'_j$ , respectively, and the corresponding mass matrices are given by

$$(\mathbf{M}'_d)_{ij} \equiv c_{ij}^{(d)} \frac{v}{\sqrt{2}}, \quad (\mathbf{M}'_u)_{ij} \equiv c_{ij}^{(u)} \frac{v}{\sqrt{2}}, \quad (\mathbf{M}'_\ell)_{ij} \equiv c_{ij}^{(\ell)} \frac{v}{\sqrt{2}}. \quad (6)$$

The diagonalization of these mass matrices determines the mass eigenstates  $d_j, u_j$  and  $\ell_j$ , which are linear combinations of the corresponding weak eigenstates  $d'_j, u'_j$  and  $\ell'_j$ , respectively.



**Fig. 2:** Tree-level FCNC couplings (green solid vertices) are absent in the SM. Therefore, very suppressed (experimentally) transitions such as  $K^0 \rightarrow \mu^+ \mu^-$  or  $K^0 - \bar{K}^0$  mixing cannot occur through tree-level exchange.

The matrix  $\mathbf{M}'_d$  can be decomposed as<sup>1</sup>  $\mathbf{M}'_d = \mathbf{H}_d \mathbf{U}_d = \mathbf{S}'_d \mathcal{M}_d \mathbf{S}_d \mathbf{U}_d$ , where  $\mathbf{H}_d \equiv \sqrt{\mathbf{M}'_d \mathbf{M}'_d{}^\dagger}$  is an Hermitian positive-definite matrix, while  $\mathbf{U}_d$  is unitary.  $\mathbf{H}_d$  can be diagonalized by a unitary matrix  $\mathbf{S}_d$ ; the resulting matrix  $\mathcal{M}_d$  is diagonal, Hermitian and positive definite. Similarly, one has  $\mathbf{M}'_u = \mathbf{H}_u \mathbf{U}_u = \mathbf{S}'_u \mathcal{M}_u \mathbf{S}_u \mathbf{U}_u$  and  $\mathbf{M}'_\ell = \mathbf{H}_\ell \mathbf{U}_\ell = \mathbf{S}'_\ell \mathcal{M}_\ell \mathbf{S}_\ell \mathbf{U}_\ell$ . In terms of the diagonal mass matrices

$$\mathcal{M}_d = \text{diag}(m_d, m_s, m_b, \dots), \quad \mathcal{M}_u = \text{diag}(m_u, m_c, m_t, \dots), \quad \mathcal{M}_\ell = \text{diag}(m_e, m_\mu, m_\tau, \dots), \quad (7)$$

the Yukawa Lagrangian takes the simpler form

$$\mathcal{L}_Y = - \left( 1 + \frac{H}{v} \right) \{ \bar{\mathbf{d}} \mathcal{M}_d \mathbf{d} + \bar{\mathbf{u}} \mathcal{M}_u \mathbf{u} + \bar{\mathbf{\ell}} \mathcal{M}_\ell \mathbf{\ell} \}, \quad (8)$$

where the mass eigenstates are defined by

$$\begin{aligned} \mathbf{d}_L &\equiv \mathbf{S}_d \mathbf{d}'_L, & \mathbf{u}_L &\equiv \mathbf{S}_u \mathbf{u}'_L, & \mathbf{\ell}_L &\equiv \mathbf{S}_\ell \mathbf{\ell}'_L, \\ \mathbf{d}_R &\equiv \mathbf{S}_d \mathbf{U}_d \mathbf{d}'_R, & \mathbf{u}_R &\equiv \mathbf{S}_u \mathbf{U}_u \mathbf{u}'_R, & \mathbf{\ell}_R &\equiv \mathbf{S}_\ell \mathbf{U}_\ell \mathbf{\ell}'_R. \end{aligned} \quad (9)$$

Note, that the Higgs couplings are proportional to the corresponding fermions masses.

Since,  $\bar{\mathbf{f}}'_L \mathbf{f}'_L = \bar{\mathbf{f}}_L \mathbf{f}_L$  and  $\bar{\mathbf{f}}'_R \mathbf{f}'_R = \bar{\mathbf{f}}_R \mathbf{f}_R$  ( $f = d, u, \ell$ ), the form of the neutral-current part of the  $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$  Lagrangian does not change when expressed in terms of mass eigenstates. Therefore, there are no flavour-changing neutral currents (FCNCs) in the SM (Fig. 2). This is a consequence of treating all equal-charge fermions on the same footing (GIM mechanism [8]), and guarantees that weak transitions such as  $B_{s,d}^0 \rightarrow \ell^+ \ell^-$ ,  $K^0 \rightarrow \mu^+ \mu^-$  or  $K^0 - \bar{K}^0$  mixing, which are known experimentally to be very suppressed, cannot happen at tree level. However,  $\bar{\mathbf{u}}'_L \mathbf{d}'_L = \bar{\mathbf{u}}_L \mathbf{S}_u \mathbf{S}'_d \mathbf{d}'_L \equiv \bar{\mathbf{u}}_L \mathbf{V} \mathbf{d}'_L$ . In general,  $\mathbf{S}_u \neq \mathbf{S}_d$ ; thus, if one writes the weak eigenstates in terms of mass eigenstates, a  $N_G \times N_G$  unitary mixing matrix  $\mathbf{V}$  appears in the quark charged-current sector as indicated in Eq. (3).

If neutrinos are assumed to be massless, we can always redefine the neutrino flavours, in such a way as to eliminate the mixing in the lepton sector:  $\bar{\nu}'_L \mathbf{\ell}'_L = \bar{\nu}'_L \mathbf{S}'_\ell \mathbf{\ell}_L \equiv \bar{\nu}_L \mathbf{\ell}_L$ . Thus, we have lepton-flavour conservation in the minimal SM without right-handed neutrinos. If sterile  $\nu_R$  fields are included in the model, one has an additional Yukawa term in Eq. (4), giving rise to a neutrino mass matrix  $(\mathbf{M}'_\nu)_{ij} \equiv c_{ij}^{(\nu)} v / \sqrt{2}$ . Thus, the model can accommodate non-zero neutrino masses and lepton-flavour violation through a lepton mixing matrix  $\mathbf{V}_L$  analogous to the one present in the quark sector. Note, however, that the total lepton number  $L \equiv L_e + L_\mu + L_\tau$  is still conserved. We know experimentally that neutrino masses are tiny and, as shown in Table 1, there are strong bounds on lepton-flavour violating decays. However, we do have a clear evidence of neutrino oscillation phenomena [9]. Moreover, since right-handed neutrinos are singlets under  $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ , the SM gauge symmetry group allows for a right-handed Majorana neutrino mass term, violating lepton number by two units. Non-zero neutrino masses clearly imply interesting new phenomena [4].

<sup>1</sup> The condition  $\det \mathbf{M}'_f \neq 0$  ( $f = d, u, \ell$ ) guarantees that the decomposition  $\mathbf{M}'_f = \mathbf{H}_f \mathbf{U}_f$  is unique:  $\mathbf{U}_f \equiv \mathbf{H}_f^{-1} \mathbf{M}'_f$ . The matrices  $\mathbf{S}_f$  are completely determined (up to phases) only if all diagonal elements of  $\mathbf{M}'_f$  are different. If there is some degeneracy, the arbitrariness of  $\mathbf{S}_f$  reflects the freedom to define the physical fields. When  $\det \mathbf{M}'_f = 0$ , the matrices  $\mathbf{U}_f$  and  $\mathbf{S}_f$  are not uniquely determined, unless their unitarity is explicitly imposed.

**Table 1:** Best published limits on lepton-flavour-violating transitions [9].

$\text{Br}(\mu^- \rightarrow X^-) \cdot 10^{12}$		(90% CL)					
$e^- \gamma$	0.42	$e^- 2\gamma$	72	$e^- e^- e^+$	1.0		
$\text{Br}(\tau^- \rightarrow X^-) \cdot 10^8$		(90% CL)					
$e^- \gamma$	3.3	$e^- e^+ e^-$	2.7	$e^- \mu^+ \mu^-$	2.7	$e^- e^- \mu^+$	1.5
$\mu^- \gamma$	4.4	$\mu^- e^+ e^-$	1.8	$\mu^- \mu^+ \mu^-$	2.1	$\mu^- \mu^- e^+$	1.7
$e^- \pi^0$	8.0	$\mu^- \pi^0$	11	$e^- \phi$	3.1	$\mu^- \phi$	8.4
$e^- \eta$	9.2	$e^- \eta'$	16	$e^- \rho^0$	1.8	$e^- \omega$	4.8
$\mu^- \eta$	6.5	$\mu^- \eta'$	13	$\mu^- \rho^0$	1.2	$\mu^- \omega$	4.7
$e^- K_S$	2.6	$e^- K^{*0}$	3.2	$e^- \bar{K}^{*0}$	3.4	$e^- K^+ \pi^-$	3.1
$\mu^- K_S$	2.3	$\mu^- K^{*0}$	5.9	$\mu^- \bar{K}^{*0}$	7.0	$\mu^- K^+ \pi^-$	4.5
$e^- K_S K_S$	7.1	$e^- K^+ K^-$	3.4	$e^- \pi^+ \pi^-$	2.3	$e^- \pi^+ K^-$	3.7
$\mu^- K_S K_S$	8.0	$\mu^- K^+ K^-$	4.4	$\mu^- \pi^+ \pi^-$	2.1	$\mu^- \pi^+ K^-$	8.6
$e^- f_0(980) \rightarrow e^- \pi^+ \pi^-$			3.2	$\mu^- f_0(980) \rightarrow \mu^- \pi^+ \pi^-$			3.4
$\text{Br}(Z \rightarrow X^0) \cdot 10^6$		(95% CL)					
$e^\pm \mu^\mp$	0.75	$e^\pm \tau^\mp$	9.8	$\mu^\pm \tau^\mp$	12		
$\text{Br}(B_{(s)}^0 \rightarrow X^0) \cdot 10^8$		(95% CL)					
$B^0 \rightarrow e^\pm \mu^\mp$	0.28	$B_s^0 \rightarrow e^\pm \mu^\mp$	1.1				
$\text{Br}(\mu^- + N \rightarrow e^- + N) \cdot 10^{11}$		(90% CL)					
Ti	0.43	Pb	4.6	S	7		

The fermion masses and the quark mixing matrix  $\mathbf{V}$  are all determined by the Yukawa couplings in Eq. (4). However, the coefficients  $c_{ij}^{(f)}$  are unknown; therefore, we have a bunch of arbitrary parameters. A general  $N_G \times N_G$  unitary matrix is characterized by  $N_G^2$  real parameters:  $N_G(N_G - 1)/2$  moduli and  $N_G(N_G + 1)/2$  phases. In the case of  $\mathbf{V}$ , many of these parameters are irrelevant because we can always choose arbitrary quark phases. Under the phase redefinitions  $u_i \rightarrow e^{i\phi_i} u_i$  and  $d_j \rightarrow e^{i\theta_j} d_j$ , the mixing matrix changes as  $\mathbf{V}_{ij} \rightarrow \mathbf{V}_{ij} e^{i(\theta_j - \phi_i)}$ ; thus,  $2N_G - 1$  phases are unobservable. The number of physical free parameters in the quark-mixing matrix then gets reduced to  $(N_G - 1)^2$ :  $N_G(N_G - 1)/2$  moduli and  $(N_G - 1)(N_G - 2)/2$  phases.

In the simpler case of two generations,  $\mathbf{V}$  is determined by a single parameter. One then recovers the Cabibbo rotation matrix [6]

$$\mathbf{V} = \begin{pmatrix} \cos \theta_C & \sin \theta_C \\ -\sin \theta_C & \cos \theta_C \end{pmatrix}. \quad (10)$$

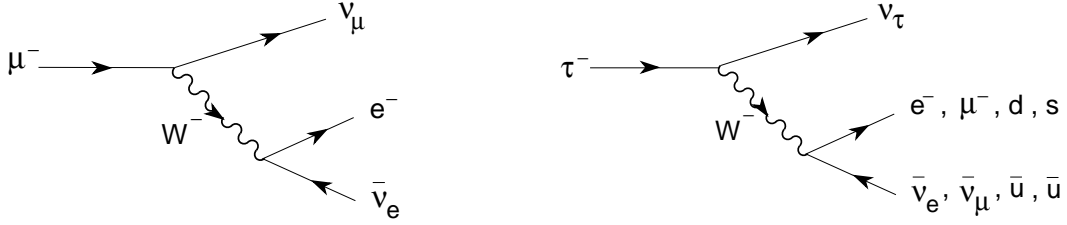
With  $N_G = 3$ , the CKM matrix is described by three angles and one phase. Different (but equivalent) representations can be found in the literature. The Particle data Group [9] advocates the use of the

following one as the ‘standard’ CKM parametrization:

$$\mathbf{V} = \begin{bmatrix} c_{12} c_{13} & s_{12} c_{13} & s_{13} e^{-i\delta_{13}} \\ -s_{12} c_{23} - c_{12} s_{23} s_{13} e^{i\delta_{13}} & c_{12} c_{23} - s_{12} s_{23} s_{13} e^{i\delta_{13}} & s_{23} c_{13} \\ s_{12} s_{23} - c_{12} c_{23} s_{13} e^{i\delta_{13}} & -c_{12} s_{23} - s_{12} c_{23} s_{13} e^{i\delta_{13}} & c_{23} c_{13} \end{bmatrix}. \quad (11)$$

Here  $c_{ij} \equiv \cos \theta_{ij}$  and  $s_{ij} \equiv \sin \theta_{ij}$ , with  $i$  and  $j$  being generation labels ( $i, j = 1, 2, 3$ ). The real angles  $\theta_{12}, \theta_{23}$  and  $\theta_{13}$  can all be made to lie in the first quadrant, by an appropriate redefinition of quark field phases; then,  $c_{ij} \geq 0, s_{ij} \geq 0$  and  $0 \leq \delta_{13} \leq 2\pi$ . Notice that  $\delta_{13}$  is the only complex phase in the SM Lagrangian. Therefore, it is the only possible source of  $\mathcal{CP}$ -violation phenomena. In fact, it was for this reason that the third generation was assumed to exist [7], before the discovery of the  $b$  and the  $\tau$ . With two generations, the SM could not explain the observed  $\mathcal{CP}$  violation in the  $K$  system.

### 3 Lepton decays



**Fig. 3:** Tree-level Feynman diagrams for  $\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu$  and  $\tau^- \rightarrow \nu_\tau X^-$  ( $X^- = e^- \bar{\nu}_e, \mu^- \bar{\nu}_\mu, d\bar{u}, s\bar{u}$ ).

The simplest flavour-changing process is the leptonic decay of the muon, which proceeds through the  $W$ -exchange diagram shown in Fig. 3. The momentum transfer carried by the intermediate  $W$  is very small compared to  $M_W$ . Therefore, the vector-boson propagator reduces to a contact interaction,

$$\frac{-g_{\mu\nu} + q_\mu q_\nu / M_W^2}{q^2 - M_W^2} \xrightarrow{q^2 \ll M_W^2} \frac{g_{\mu\nu}}{M_W^2}. \quad (12)$$

The decay can then be described through an effective local four-fermion Hamiltonian,

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}} [\bar{e} \gamma^\alpha (1 - \gamma_5) \nu_e] [\bar{\nu}_\mu \gamma_\alpha (1 - \gamma_5) \mu], \quad (13)$$

where

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2} = \frac{1}{2v^2} \quad (14)$$

is called the Fermi coupling constant.  $G_F$  is fixed by the total decay width,

$$\frac{1}{\tau_\mu} = \Gamma[\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu (\gamma)] = \frac{G_F^2 m_\mu^5}{192\pi^3} (1 + \delta_{\text{RC}}) f\left(\frac{m_e^2}{m_\mu^2}\right), \quad (15)$$

where  $f(x) = 1 - 8x + 8x^3 - x^4 - 12x^2 \ln x$ , and  $\delta_{\text{RC}} \approx \frac{\alpha}{2\pi} \left(\frac{25}{4} - \pi^2\right)$  takes into account higher-order QED corrections, which are known to  $\mathcal{O}(\alpha^2)$  [10–12]. The tiny neutrino masses can be safely neglected. The measured lifetime [13],  $\tau_\mu = (2.196\,981\,1 \pm 0.000\,002\,2) \cdot 10^{-6}$  s, implies the value

$$G_F = (1.166\,378\,7 \pm 0.000\,000\,6) \cdot 10^{-5} \text{ GeV}^{-2} \approx \frac{1}{(293 \text{ GeV})^2}. \quad (16)$$

**Table 2:** Experimental determinations of the ratios  $g_\ell/g_{\ell'}$  [17].

	$\Gamma_{\tau \rightarrow \mu}/\Gamma_{\tau \rightarrow e}$	$\Gamma_{\pi \rightarrow \mu}/\Gamma_{\pi \rightarrow e}$	$\Gamma_{K \rightarrow \mu}/\Gamma_{K \rightarrow e}$	$\Gamma_{K \rightarrow \pi \mu}/\Gamma_{K \rightarrow \pi e}$	$\Gamma_{W \rightarrow \mu}/\Gamma_{W \rightarrow e}$
$ g_\mu/g_e $	1.0018 (14)	1.0021 (16)	0.9978 (20)	1.0010 (25)	0.996 (10)
	$\Gamma_{\tau \rightarrow e}/\Gamma_{\mu \rightarrow e}$	$\Gamma_{\tau \rightarrow \pi}/\Gamma_{\pi \rightarrow \mu}$	$\Gamma_{\tau \rightarrow K}/\Gamma_{K \rightarrow \mu}$	$\Gamma_{W \rightarrow \tau}/\Gamma_{W \rightarrow \mu}$	
$ g_\tau/g_\mu $	1.0011 (15)	0.9962 (27)	0.9858 (70)	1.034 (13)	
	$\Gamma_{\tau \rightarrow \mu}/\Gamma_{\mu \rightarrow e}$	$\Gamma_{W \rightarrow \tau}/\Gamma_{W \rightarrow e}$			
$ g_\tau/g_e $	1.0030 (15)	1.031 (13)			

The decays of the  $\tau$  lepton proceed through the same  $W$ -exchange mechanism. The only difference is that several final states are kinematically allowed:  $\tau^- \rightarrow \nu_\tau e^- \bar{\nu}_e$ ,  $\tau^- \rightarrow \nu_\tau \mu^- \bar{\nu}_\mu$ ,  $\tau^- \rightarrow \nu_\tau d \bar{u}$  and  $\tau^- \rightarrow \nu_\tau s \bar{u}$ . Owing to the universality of the  $W$  couplings in  $\mathcal{L}_{CC}$ , all these decay modes have equal amplitudes (if final fermion masses and QCD interactions are neglected), except for an additional  $N_C |\mathbf{V}_{ui}|^2$  factor ( $i = d, s$ ) in the semileptonic channels, where  $N_C = 3$  is the number of quark colours. Making trivial kinematical changes in Eq. (15), one easily gets the lowest-order prediction for the total  $\tau$  decay width:

$$\frac{1}{\tau_\tau} \equiv \Gamma(\tau) \approx \Gamma(\mu) \left( \frac{m_\tau}{m_\mu} \right)^5 \left\{ 2 + N_C \left( |\mathbf{V}_{ud}|^2 + |\mathbf{V}_{us}|^2 \right) \right\} \approx \frac{5}{\tau_\mu} \left( \frac{m_\tau}{m_\mu} \right)^5, \quad (17)$$

where we have used the CKM unitarity relation  $|\mathbf{V}_{ud}|^2 + |\mathbf{V}_{us}|^2 = 1 - |\mathbf{V}_{ub}|^2 \approx 1$  (we will see later that this is an excellent approximation). From the measured muon lifetime, one has then  $\tau_\tau \approx 3.3 \cdot 10^{-13}$  s, to be compared with the experimental value  $\tau_\tau^{\text{exp}} = (2.903 \pm 0.005) \cdot 10^{-13}$  s [9]. The numerical difference is due to the effect of QCD corrections which enhance the hadronic  $\tau$  decay width by about 20%. The size of these corrections has been accurately predicted in terms of the strong coupling [14], allowing us to extract from  $\tau$  decays one of the most precise determinations of  $\alpha_s$  [15, 16].

In the SM all lepton doublets have identical couplings to the  $W$  boson. Comparing the measured decay widths of leptonic or semileptonic decays which only differ in the lepton flavour, one can test experimentally that the  $W$  interaction is indeed the same, i.e., that  $g_e = g_\mu = g_\tau \equiv g$ . As shown in Table 2, the present data verify the universality of the leptonic charged-current couplings to the 0.2% level.

#### 4 Quark mixing

In order to measure the CKM matrix elements  $\mathbf{V}_{ij}$ , one needs to study hadronic weak decays of the type  $H \rightarrow H' \ell^- \bar{\nu}_\ell$  or  $H' \rightarrow H \ell^+ \nu_\ell$  that are associated with the corresponding quark transitions  $d_j \rightarrow u_i \ell^- \bar{\nu}_\ell$  and  $u_i \rightarrow d_j \ell^+ \nu_\ell$  (Fig. 4). Since quarks are confined within hadrons, the decay amplitude

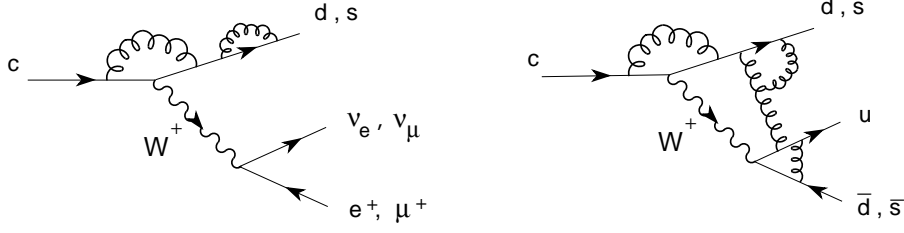
$$T[H \rightarrow H' \ell^- \bar{\nu}_\ell] \approx \frac{G_F}{\sqrt{2}} \mathbf{V}_{ij} \langle H' | \bar{u}_i \gamma^\mu (1 - \gamma_5) d_j | H \rangle [\bar{\ell} \gamma_\mu (1 - \gamma_5) \nu_\ell] \quad (18)$$

always involves an hadronic matrix element of the weak left current. The evaluation of this matrix element is a non-perturbative QCD problem, which introduces unavoidable theoretical uncertainties.

One usually looks for a semileptonic transition where the matrix element can be fixed at some kinematical point by a symmetry principle. This has the virtue of reducing the theoretical uncertainties to the level of symmetry-breaking corrections and kinematical extrapolations. The standard example is a  $0^- \rightarrow 0^-$  decay such as  $K \rightarrow \pi \ell \nu_\ell$ ,  $D \rightarrow K \ell \nu_\ell$  or  $B \rightarrow D \ell \nu_\ell$ , where, owing to parity (the vector and axial-vector currents have  $J^P = 1^-$  and  $1^+$ , respectively), only the vector current contributes:

$$\langle P'(k') | \bar{u}_i \gamma^\mu d_j | P(k) \rangle = C_{PP'} \left\{ (k + k')^\mu f_+(t) + (k - k')^\mu f_-(t) \right\}. \quad (19)$$





**Fig. 4:**  $V_{ij}$  are measured in semileptonic decays (left), where a single quark current is present. Hadronic decays (right) involve two different quark currents and are more affected by QCD effects (gluons can couple everywhere).

Here,  $C_{PP'}$  is a Clebsch–Gordan factor relating  $P \rightarrow P'$  transitions that only differ by the meson electromagnetic charges, and  $t = (k - k')^2 \equiv q^2$  is the momentum transfer. The unknown strong dynamics is fully contained in the form factors  $f_{\pm}(t)$ .

In the limit of equal quark masses,  $m_{u_i} = m_{d_j}$ , the divergence of the vector current is zero. Thus  $q_{\mu} [\bar{u}_i \gamma^{\mu} d_j] = 0$ , which implies  $f_{-}(t) = 0$ . Moreover, as shown in the appendix,  $f_{+}(0) = 1$  to all orders in the strong coupling because the associated flavour charge is a conserved quantity.<sup>2</sup> Therefore, one only needs to estimate the corrections induced by the quark mass differences.

Since  $q_{\mu} [\bar{\ell} \gamma^{\mu} (1 - \gamma_5) \nu_{\ell}] \sim m_{\ell}$ , the contribution of  $f_{-}(t)$  is kinematically suppressed in the electron and muon decay modes. The decay width can then be written as ( $\ell = e, \mu$ )

$$\Gamma(P \rightarrow P' \ell \nu) = \frac{G_F^2 M_P^5}{192\pi^3} |\mathbf{V}_{ij}|^2 C_{PP'}^2 |f_{+}(0)|^2 \mathcal{I} (1 + \delta_{\text{RC}}), \quad (20)$$

where  $\delta_{\text{RC}}$  is an electroweak radiative correction factor and  $\mathcal{I}$  denotes a phase-space integral, which in the  $m_{\ell} = 0$  limit takes the form

$$\mathcal{I} \approx \int_0^{(M_P - M_{P'})^2} \frac{dt}{M_P^8} \lambda^{3/2}(t, M_P^2, M_{P'}^2) \left| \frac{f_{+}(t)}{f_{+}(0)} \right|^2. \quad (21)$$

The usual procedure to determine  $|\mathbf{V}_{ij}|$  involves three steps:

1. Measure the shape of the  $t$  distribution. This fixes  $|f_{+}(t)/f_{+}(0)|$  and therefore determines  $\mathcal{I}$ .
2. Measure the total decay width  $\Gamma$ . Since  $G_F$  is already known from  $\mu$  decay, one gets then an experimental value for the product  $|f_{+}(0) \mathbf{V}_{ij}|$ .
3. Get a theoretical prediction for  $f_{+}(0)$ .

It is important to realize that theoretical input is always needed. Thus, the accuracy of the  $|\mathbf{V}_{ij}|$  determination is limited by our ability to calculate the relevant hadronic parameters.

#### 4.1 Determination of $|\mathbf{V}_{ud}|$ and $|\mathbf{V}_{us}|$

The conservation of the vector QCD currents in the massless quark limit allows for precise determinations of the light-quark mixings. The most accurate measurement of  $\mathbf{V}_{ud}$  is done with superallowed nuclear  $\beta$  decays of the Fermi type ( $0^+ \rightarrow 0^+$ ), where the nuclear matrix element  $\langle N' | \bar{u} \gamma^{\mu} d | N \rangle$  can be fixed by vector-current conservation. The CKM factor is obtained through the relation [19, 20],

$$|\mathbf{V}_{ud}|^2 = \frac{\pi^3 \ln 2}{ft G_F^2 m_e^5 (1 + \delta_{\text{RC}})} = \frac{(2984.48 \pm 0.05) \text{ s}}{ft (1 + \delta_{\text{RC}})}, \quad (22)$$

<sup>2</sup> This is completely analogous to the electromagnetic charge conservation in QED. The conservation of the electromagnetic current implies that the proton electromagnetic form factor does not get any QED or QCD correction at  $q^2 = 0$  and, therefore,  $Q_p = 2Q_u + Q_d = |Q_e|$ . An explicit proof can be found in Ref. [18].

where  $ft$  denotes the product of a phase-space statistical decay-rate factor and the measured half-life. In order to obtain  $|\mathbf{V}_{ud}|$ , one needs to perform a careful analysis of radiative corrections, including electroweak contributions, nuclear-structure corrections and isospin-violating nuclear effects. These nuclear-dependent corrections are quite large,  $\delta_{\text{RC}} \sim 3\text{--}4\%$ , and have a crucial role in bringing the results from different nuclei into good agreement. The weighted average of the fourteen most precise determinations yields [21]

$$|\mathbf{V}_{ud}| = 0.97417 \pm 0.00021. \quad (23)$$

The error is dominated by theoretical uncertainties stemming from nuclear Coulomb distortions and radiative corrections.

An independent determination of  $|\mathbf{V}_{ud}|$  can be obtained from neutron decay,  $n \rightarrow p e^- \bar{\nu}_e$ . The axial current also contributes in this case; therefore, one needs to use the experimental value of the axial-current matrix element at  $q^2 = 0$ ,  $\langle p | \bar{u} \gamma^\mu \gamma_5 d | n \rangle = G_A \bar{p} \gamma^\mu n$ . The present world averages,  $g_A \equiv G_A/G_V = -1.2723 \pm 0.0023$  and  $\tau_n = (880.2 \pm 1.0)$  s, imply [9, 19, 20]:

$$|\mathbf{V}_{ud}| = \left\{ \frac{(4908.7 \pm 1.9) \text{ s}}{\tau_n (1 + 3g_A^2)} \right\}^{1/2} = 0.9759 \pm 0.0016, \quad (24)$$

which is  $1.1 \sigma$  larger than (23) but less precise.

The pion  $\beta$  decay  $\pi^+ \rightarrow \pi^0 e^+ \nu_e$  offers a cleaner way to measure  $|\mathbf{V}_{ud}|$ . It is a pure vector transition, with very small theoretical uncertainties. At  $q^2 = 0$ , the hadronic matrix element does not receive isospin-breaking contributions of first order in  $m_d - m_u$ , i.e.,  $f_+(0) = 1 + \mathcal{O}[(m_d - m_u)^2]$  [22]. The small available phase space makes it possible to theoretically control the form factor with high accuracy over the entire kinematical domain [23]; unfortunately, it also implies a very suppressed branching fraction of  $\mathcal{O}(10^{-8})$ . From the currently measured value [24], one gets  $|\mathbf{V}_{ud}| = 0.9749 \pm 0.0026$  [9]. A tenfold improvement of the experimental accuracy would be needed to get a determination competitive with (23).

The standard determination of  $|\mathbf{V}_{us}|$  takes advantage of the theoretically well-understood decay amplitudes in  $K \rightarrow \pi \ell \nu_\ell$ . The high accuracy achieved in high-statistics experiments [9], supplemented with theoretical calculations of electromagnetic and isospin corrections [25, 26], allows us to extract the product  $|\mathbf{V}_{us} f_+(0)| = 0.2165 \pm 0.0004$  [27, 28], with  $f_+(0) = 1 + \mathcal{O}[(m_s - m_u)^2]$  the vector form factor of the  $K^0 \rightarrow \pi^- \ell^+ \nu_\ell$  decay [22, 29]. The exact value of  $f_+(0)$  has been thoroughly investigated since the first precise estimate by Leutwyler and Roos,  $f_+(0) = 0.961 \pm 0.008$  [30]. The most recent and precise lattice determinations exhibit a clear shift to higher values [31, 32], in agreement with the analytical chiral perturbation theory predictions at two loops [33–35]. Taking the current lattice average (with  $2 + 1 + 1$  active fermions),  $f_+(0) = 0.9706 \pm 0.0027$  [36], one obtains

$$|\mathbf{V}_{us}| = 0.2231 \pm 0.0007. \quad (25)$$

The ratio of radiative inclusive decay rates  $\Gamma[K \rightarrow \mu\nu(\gamma)]/\Gamma[\pi \rightarrow \mu\nu(\gamma)]$  provides also information on  $\mathbf{V}_{us}$  [27, 37]. With a careful treatment of electromagnetic and isospin-violating corrections, one extracts  $|\mathbf{V}_{us}/\mathbf{V}_{ud}| |f_K/f_\pi| = 0.2760 \pm 0.0004$  [28, 38, 39]. Taking for the ratio of meson decay constants the lattice average  $f_K/f_\pi = 1.1933 \pm 0.0029$  [36], one finally gets

$$\frac{|\mathbf{V}_{us}|}{|\mathbf{V}_{ud}|} = 0.2313 \pm 0.0007. \quad (26)$$

With the value of  $|\mathbf{V}_{ud}|$  in Eq. (23), this implies  $|\mathbf{V}_{us}| = 0.2253 \pm 0.0007$  that is  $2.2 \sigma$  larger than (25).

Hyperon decays are also sensitive to  $\mathbf{V}_{us}$  [40]. Unfortunately, in weak baryon decays the theoretical control on  $SU(3)$ -breaking corrections is not as good as for the meson case. A conservative estimate of these effects leads to the result  $|\mathbf{V}_{us}| = 0.226 \pm 0.005$  [41].

The accuracy of all previous determinations is limited by theoretical uncertainties. The separate measurement of the inclusive  $|\Delta S| = 0$  and  $|\Delta S| = 1$  tau decay widths provides a very clean observable to directly measure  $|\mathbf{V}_{us}|$  [17, 42] because  $SU(3)$ -breaking corrections are suppressed by two powers of the  $\tau$  mass. The present  $\tau$  decay data imply  $|\mathbf{V}_{us}| = 0.2186 \pm 0.0021$  [43], the error being dominated by the experimental uncertainties. The central value has been shifted down by the BaBar and Belle measurements, which find branching ratios smaller than previous world averages for many  $\tau$  decay modes [9]. More precise data are needed to clarify this worrisome effect. From the measured  $K^- \rightarrow \mu^- \bar{\nu}_\mu$  and  $K \rightarrow \pi \ell \nu_\ell$  decay amplitudes, one actually predicts slightly higher  $\tau^- \rightarrow \nu_\tau K^-$  and  $\tau^- \rightarrow \nu_\tau (K\pi)^-$  branching ratios [17, 44]. Since these channels are the largest contributions to the inclusive  $|\Delta S| = 1$  tau decay width, their slight underestimate has a very significant effect on  $|\mathbf{V}_{us}|$ . Using the kaon determinations for these  $\tau$  branching ratios one gets instead  $|\mathbf{V}_{us}| = 0.2213 \pm 0.0023$ .

## 4.2 Determination of $|\mathbf{V}_{cb}|$ and $|\mathbf{V}_{ub}|$

In the limit of very heavy quark masses, QCD has additional flavour and spin symmetries [45–48] that can be used to make precise determinations of  $|\mathbf{V}_{cb}|$ , either from exclusive semileptonic decays such as  $B \rightarrow D \ell \bar{\nu}_\ell$  and  $B \rightarrow D^* \ell \bar{\nu}_\ell$  [49, 50] or from the inclusive analysis of  $b \rightarrow c \ell \bar{\nu}_\ell$  transitions. In the rest frame of a heavy-light meson  $\bar{Q}q$ , with  $M_Q \gg (m_q, \Lambda_{\text{QCD}})$ , the heavy quark  $Q$  is practically at rest and acts as a static source of gluons ( $\lambda_Q \sim 1/M_Q \ll R_{\text{had}} \sim 1/\Lambda_{\text{QCD}}$ ). At  $M_Q \rightarrow \infty$ , the interaction becomes then independent of the heavy-quark mass and spin. Moreover, assuming that the charm quark is heavy enough, the  $b \rightarrow c \ell \bar{\nu}_\ell$  transition within the meson does not modify the interaction with the light quark at zero recoil, i.e., when the meson velocity remains unchanged ( $v_D = v_B$ ).

Taking the limit  $m_b > m_c \rightarrow \infty$ , all form factors characterizing the decays  $B \rightarrow D \ell \bar{\nu}_\ell$  and  $B \rightarrow D^* \ell \bar{\nu}_\ell$  reduce to a single function [45], which depends on the product of the four-velocities of the two mesons  $w \equiv v_B \cdot v_{D^{(*)}} = (M_B^2 + M_{D^{(*)}}^2 - q^2)/(2M_B M_{D^{(*)}})$ . Heavy quark symmetry determines the normalization of the rate at  $w = 1$ , the maximum momentum transfer to the leptons, because the corresponding vector current is conserved in the limit of equal  $B$  and  $D^{(*)}$  velocities. The  $B \rightarrow D^*$  mode has the additional advantage that corrections to the infinite-mass limit are of second order in  $1/m_b - 1/m_c$  at zero recoil ( $w = 1$ ) [50].

The exclusive determination of  $|\mathbf{V}_{cb}|$  is obtained from an extrapolation of the measured spectrum to  $w = 1$ . Using the CLN parametrization of the relevant form factors [51], which is based on heavy-quark symmetry and includes  $1/M_Q$  corrections, the *Heavy Flavor Averaging group* (HFLAV) [43] quotes the experimental value  $\eta_{\text{EW}} \mathcal{F}(1) |\mathbf{V}_{cb}| = (35.61 \pm 0.43) \cdot 10^{-3}$  from  $B \rightarrow D^* \ell \bar{\nu}_\ell$  data, while the measured  $B \rightarrow D \ell \bar{\nu}_\ell$  distribution results in  $\eta_{\text{EW}} \mathcal{G}(1) |\mathbf{V}_{cb}| = (41.57 \pm 1.00) \cdot 10^{-3}$ , where  $\mathcal{F}(1)$  and  $\mathcal{G}(1)$  are the corresponding form factors at  $w = 1$  and  $\eta_{\text{EW}}$  accounts for small electroweak corrections. Lattice simulations are used to estimate the deviations from unity of the two form factors at zero recoil. Using  $\eta_{\text{EW}} \mathcal{F}(1) = 0.912 \pm 0.014$  [52] and  $\eta_{\text{EW}} \mathcal{G}(1) = 1.061 \pm 0.010$  [53], one gets [43]

$$|\mathbf{V}_{cb}| = \begin{cases} (39.05 \pm 0.47_{\text{exp}} \pm 0.58_{\text{th}}) \cdot 10^{-3} & (B \rightarrow D^* \ell \bar{\nu}_\ell) \\ (39.18 \pm 0.94_{\text{exp}} \pm 0.36_{\text{th}}) \cdot 10^{-3} & (B \rightarrow D \ell \bar{\nu}_\ell) \end{cases} = (39.10 \pm 0.61) \cdot 10^{-3}. \quad (27)$$

It has been pointed out recently that the CLN parametrization is only valid within 2% and this uncertainty has not been properly taken into account in the experimental extrapolations [54–57]. Using instead the more general BGL parametrization [58], combined with lattice and light-cone sum rules information, the analysis of the most recent  $B \rightarrow D^* \ell \bar{\nu}_\ell$  Belle data [59] gives [55]

$$|\mathbf{V}_{cb}| = (40.6 \pm 1.3) \cdot 10^{-3}, \quad (28)$$

while a similar analysis of BaBar [60] and Belle [61]  $B \rightarrow D \ell \bar{\nu}_\ell$  data obtains [54]

$$|\mathbf{V}_{cb}| = (40.49 \pm 0.97) \cdot 10^{-3}. \quad (29)$$

These numbers are significantly higher than the corresponding HFLAV results in Eq. (27) and indicate the presence of underestimated uncertainties.

The inclusive determination of  $|\mathbf{V}_{cb}|$  uses the Operator Product Expansion [62, 63] to express the total  $b \rightarrow c \ell \bar{\nu}_\ell$  rate and moments of the differential energy and invariant-mass spectra in a double expansion in powers of  $\alpha_s$  and  $1/m_b$ , which includes terms of  $\mathcal{O}(\alpha_s^2)$  and up to  $\mathcal{O}(1/m_b^5)$  [12, 64–74]. The non-perturbative matrix elements of the corresponding local operators are obtained from a global fit to experimental moments of inclusive lepton energy and hadronic invariant mass distributions. The most recent analyses find [75, 76]

$$|\mathbf{V}_{cb}| = (42.00 \pm 0.63) \cdot 10^{-3}. \quad (30)$$

This value, which we will adopt in the following, agrees within errors with the exclusive determinations (28) and (29).

The presence of a light quark makes more difficult to control the theoretical uncertainties in the analogous determinations of  $|\mathbf{V}_{ub}|$ . Exclusive  $B \rightarrow \pi \ell \nu_\ell$  decays involve a non-perturbative form factor  $f_+(t)$  which is estimated through light-cone sum rules [77–80] and lattice simulations [81, 82]. The inclusive measurement requires the use of stringent experimental cuts to suppress the  $b \rightarrow X_c \ell \nu_\ell$  background that has fifty times larger rates. This induces sizeable errors in the theoretical predictions [83–91], which become sensitive to non-perturbative shape functions and depend much more strongly on  $m_b$ . The HFLAV group quotes the values [43]

$$|\mathbf{V}_{ub}| = \begin{cases} (3.67 \pm 0.09_{\text{exp}} \pm 0.12_{\text{th}}) \cdot 10^{-3} & (B \rightarrow \pi \ell \bar{\nu}_\ell) \\ (4.52 \pm 0.15_{\text{exp}} \pm 0.11_{\text{th}}) \cdot 10^{-3} & (B \rightarrow X_u \ell \bar{\nu}_\ell) \end{cases} = (3.98 \pm 0.40) \cdot 10^{-3}. \quad (31)$$

Since the exclusive and inclusive determinations of  $|\mathbf{V}_{ub}|$  disagree, we have averaged both values scaling the error by  $\sqrt{\chi^2/\text{dof}} = 3.4$ .

LHCb has extracted  $|\mathbf{V}_{ub}|/|\mathbf{V}_{cb}|$  from the measured ratio of high- $q^2$  events between the  $\Lambda_b$  decay modes into  $p\mu\nu$  ( $q^2 > 15 \text{ GeV}^2$ ) and  $\Lambda_c\mu\nu$  ( $q^2 > 7 \text{ GeV}^2$ ) [92]:

$$\frac{|\mathbf{V}_{ub}|}{|\mathbf{V}_{cb}|} = 0.080 \pm 0.004_{\text{exp}} \pm 0.004_{\text{FF}}, \quad (32)$$

where the second error is due to the limited knowledge of the relevant form factors. This ratio is compatible with the values of  $|\mathbf{V}_{cb}|$  and  $|\mathbf{V}_{ub}|$  in Eqs. (30) and (31), at the  $1.3\sigma$  level.

$|\mathbf{V}_{ub}|$  can be also extracted from the  $B^- \rightarrow \tau^- \bar{\nu}_\tau$  decay width, taking the B-meson decay constant  $f_B$  from lattice calculations [36]. Unfortunately, the current tension between the BaBar [93] and Belle [94] measurements does not allow for a very precise determination. The particle data group quotes  $|\mathbf{V}_{ub}| = (4.12 \pm 0.37 \pm 0.09) \cdot 10^{-3}$  [9], which agrees with either the exclusive or inclusive values in Eq. (31).

### 4.3 Determination of the charm and top CKM elements

The analytic control of theoretical uncertainties is more difficult in semileptonic charm decays, because the symmetry arguments associated with the light and heavy quark limits get corrected by sizeable symmetry-breaking effects. The magnitude of  $|\mathbf{V}_{cd}|$  can be extracted from  $D \rightarrow \pi \ell \nu_\ell$  and  $D \rightarrow \ell \nu_\ell$  decays, while  $|\mathbf{V}_{cs}|$  is obtained from  $D \rightarrow K \ell \nu_\ell$  and  $D_s \rightarrow \ell \nu_\ell$ , using the lattice determinations of the relevant form factor normalizations and decay constants [36]. The HFLAV group quotes the averages [43]

$$|\mathbf{V}_{cd}| = 0.216 \pm 0.005, \quad |\mathbf{V}_{cs}| = 0.997 \pm 0.017. \quad (33)$$

The difference of the ratio of double-muon to single-muon production by neutrino and antineutrino beams is proportional to the charm cross section off valence  $d$  quarks and, therefore, to  $|\mathbf{V}_{cd}|$  times the

average semileptonic branching ratio of charm mesons. This allows for an independent determination of  $|\mathbf{V}_{cd}|$ . Averaging data from several experiments, the PDG quotes [9]

$$|\mathbf{V}_{cd}| = 0.230 \pm 0.011, \quad (34)$$

which agrees with (33) but has a larger uncertainty. The analogous determination of  $|\mathbf{V}_{cs}|$  from  $\nu s \rightarrow cX$  suffers from the uncertainty of the  $s$ -quark sea content.

The top quark has only been seen decaying into bottom. From the ratio of branching fractions  $\text{Br}(t \rightarrow Wb)/\text{Br}(t \rightarrow Wq)$ , CMS has extracted [95]

$$\frac{|\mathbf{V}_{tb}|}{\sqrt{\sum_q |\mathbf{V}_{tq}|^2}} > 0.975 \quad (95\% \text{CL}), \quad (35)$$

where  $q = b, s, d$ . A more direct determination of  $|\mathbf{V}_{tb}|$  can be obtained from the single top-quark production cross section, measured at the LHC and the Tevatron. The PDG quotes the world average [9]

$$|\mathbf{V}_{tb}| = 1.009 \pm 0.031. \quad (36)$$

#### 4.4 Structure of the CKM matrix

Using the previous determinations of CKM elements, we can check the unitarity of the quark mixing matrix. The most precise test involves the elements of the first row:

$$|\mathbf{V}_{ud}|^2 + |\mathbf{V}_{us}|^2 + |\mathbf{V}_{ub}|^2 = 0.9988 \pm 0.0005, \quad (37)$$

where we have taken as reference values the determinations in Eqs. (23), (25) and (31). Radiative corrections play a crucial role at the quoted level of uncertainty, while the  $|\mathbf{V}_{ub}|^2$  contribution is negligible. This relation exhibits a  $2.4\sigma$  violation of unitarity, at the per-mill level, which calls for an independent re-evaluation of the very precise  $|\mathbf{V}_{ud}|$  value in Eq. (23) and improvements on the  $|\mathbf{V}_{us}|$  determination.

With the  $|\mathbf{V}_{cq}|^2$  values in Eqs. (30) and (33) we can also test the unitarity relation in the second row,

$$|\mathbf{V}_{cd}|^2 + |\mathbf{V}_{cs}|^2 + |\mathbf{V}_{cb}|^2 = 1.042 \pm 0.034, \quad (38)$$

and, adding the information on  $|\mathbf{V}_{tb}|$  in Eq. (36), the relation involving the third column,

$$|\mathbf{V}_{ub}|^2 + |\mathbf{V}_{cb}|^2 + |\mathbf{V}_{tb}|^2 = 1.020 \pm 0.063. \quad (39)$$

The ratio of the total hadronic decay width of the  $W$  to the leptonic one provides the sum [96,97]

$$\sum_{j=d,s,b} \left( |\mathbf{V}_{uj}|^2 + |\mathbf{V}_{cj}|^2 \right) = 2.002 \pm 0.027, \quad (40)$$

which involves the first and second rows of the CKM matrix. Although much less precise than Eq. (37), these three results test unitarity at the 3%, 6% and 1.4% level, respectively.

From Eq. (40) one can also obtain an independent estimate of  $|\mathbf{V}_{cs}|$ , using the experimental knowledge on the other CKM matrix elements, i.e.,  $|\mathbf{V}_{ud}|^2 + |\mathbf{V}_{us}|^2 + |\mathbf{V}_{ub}|^2 + |\mathbf{V}_{cd}|^2 + |\mathbf{V}_{cb}|^2 = 1.0472 \pm 0.0022$ . This gives

$$|\mathbf{V}_{cs}| = 0.977 \pm 0.014, \quad (41)$$

which is slightly more accurate than the direct determination in Eq. (33), and compatible with it.

The measured entries of the CKM matrix show a hierarchical pattern, with the diagonal elements being very close to one, the ones connecting the first two generations having a size

$$\lambda \approx |\mathbf{V}_{us}| = 0.2231 \pm 0.0007, \quad (42)$$

the mixing between the second and third families being of order  $\lambda^2$ , and the mixing between the first and third quark generations having a much smaller size of about  $\lambda^3$ . It is then quite practical to use the approximate parametrization [98]:

$$\mathbf{V} = \begin{bmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{bmatrix} + \mathcal{O}(\lambda^4), \quad (43)$$

where

$$A \approx \frac{|\mathbf{V}_{cb}|}{\lambda^2} = 0.844 \pm 0.013, \quad \sqrt{\rho^2 + \eta^2} \approx \left| \frac{\mathbf{V}_{ub}}{\lambda \mathbf{V}_{cb}} \right| = 0.425 \pm 0.043. \quad (44)$$

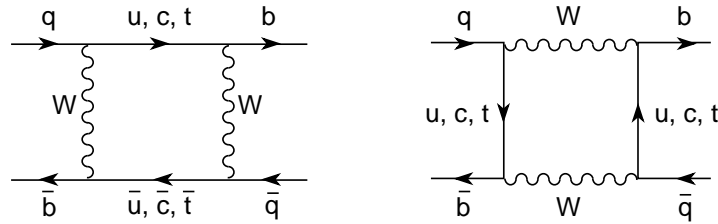
Defining to all orders in  $\lambda$  [99]  $s_{12} \equiv \lambda$ ,  $s_{23} \equiv A\lambda^2$  and  $s_{13} e^{-i\delta_{13}} \equiv A\lambda^3(\rho - i\eta)$ , Eq. (43) just corresponds to a Taylor expansion of Eq. (11) in powers of  $\lambda$ .

## 5 Meson-antimeson mixing

Additional information on the CKM parameters can be obtained from FCNC transitions, occurring at the one-loop level. An important example is provided by the mixing between the  $B_d^0$  meson and its antiparticle. This process occurs through the box diagrams shown in Fig. 5, where two  $W$  bosons are exchanged between a pair of quark lines. The mixing amplitude is proportional to

$$\langle \bar{B}_d^0 | \mathcal{H}_{\Delta B=2} | B_d^0 \rangle \sim \sum_{ij} \mathbf{V}_{id} \mathbf{V}_{ib}^* \mathbf{V}_{jd} \mathbf{V}_{jb}^* S(r_i, r_j) \sim \mathbf{V}_{td}^2 S(r_t, r_t), \quad (45)$$

where  $S(r_i, r_j)$  is a loop function [100] which depends on  $r_i \equiv m_i^2/M_W^2$ , with  $m_i$  the masses of the up-type quarks running along the internal fermionic lines. Owing to the unitarity of the CKM matrix, the mixing vanishes for equal (up-type) quark masses (GIM mechanism [8]); thus the flavour-changing transition is governed by the mass splittings between the  $u$ ,  $c$  and  $t$  quarks. Since the different CKM factors have all a similar size,  $\mathbf{V}_{ud} \mathbf{V}_{ub}^* \sim \mathbf{V}_{cd} \mathbf{V}_{cb}^* \sim \mathbf{V}_{td} \mathbf{V}_{tb}^* \sim A\lambda^3$ , the final amplitude is completely dominated by the top contribution. This transition can then be used to perform an indirect determination of  $\mathbf{V}_{td}$ .



**Fig. 5:** Box diagrams contributing to  $B^0-\bar{B}^0$  mixing ( $q = d, s$ ).

Notice that this determination has a qualitatively different character than the ones obtained before from tree-level weak decays. Now, we are going to test the structure of the electroweak theory at the quantum level. This flavour-changing transition could then be sensitive to contributions from new physics at higher energy scales. Moreover, the mixing amplitude crucially depends on the unitarity of the CKM matrix. Without the GIM mechanism embodied in the CKM mixing structure, the calculation of the analogous  $K^0 \rightarrow \bar{K}^0$  transition (replace the  $b$  by a strange quark  $s$  in the box diagrams) would have failed to explain the observed  $K^0-\bar{K}^0$  mixing by several orders of magnitude [101].



### 5.1 Mixing formalism

Since weak interactions can transform a  $P^0$  state ( $P = K, D, B$ ) into its antiparticle  $\bar{P}^0$ , these flavour eigenstates are not mass eigenstates and do not follow an exponential decay law. Let us consider an arbitrary mixture of the two flavour states,

$$|\psi(t)\rangle = a(t)|P^0\rangle + b(t)|\bar{P}^0\rangle \equiv \begin{pmatrix} a(t) \\ b(t) \end{pmatrix}, \quad (46)$$

with the time evolution (in the meson rest frame)

$$i \frac{d}{dt} |\psi(t)\rangle = \mathcal{M} |\psi(t)\rangle. \quad (47)$$

Assuming  $\mathcal{CPT}$  symmetry to hold, the  $2 \times 2$  mixing matrix can be written as

$$\mathcal{M} = \begin{pmatrix} M & M_{12} \\ M_{12}^* & M \end{pmatrix} - \frac{i}{2} \begin{pmatrix} \Gamma & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma \end{pmatrix}. \quad (48)$$

The diagonal elements  $M$  and  $\Gamma$  are real parameters, which would correspond to the mass and width of the neutral mesons in the absence of mixing. The off-diagonal entries contain the  $\Delta F = 2$  transition amplitude ( $F = S, C, B$ ):

$$M_{12} - \frac{i}{2} \Gamma_{12} = \frac{1}{2M} \left\{ \langle P^0 | \mathcal{H}_{\Delta F=2} | \bar{P}^0 \rangle + \sum_f \frac{\langle P^0 | \mathcal{H}_{\Delta F=1} | f \rangle \langle f | \mathcal{H}_{\Delta F=1} | \bar{P}^0 \rangle}{M - E_f + i\varepsilon} \right\}. \quad (49)$$

In addition to the short-distance  $\Delta F = 2$  Hamiltonian generated by the box diagrams, the mixing amplitude  $M_{12}$  receives non-local contributions involving two  $\Delta F = 1$  transitions:  $\bar{P}^0 \rightarrow f \rightarrow P^0$ . The sum extends over all possible intermediate states  $|f\rangle$ . Using the relation ( $\mathcal{P}$  denotes principal part)

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{M - E_f + i\varepsilon} = \mathcal{P} \left( \frac{1}{M - E_f} \right) - i\pi \delta(M - E_f), \quad (50)$$

one can separate the *dispersive* and *absorptive* parts of (49). The *absorptive* contribution  $\Gamma_{12}$  arises from on-shell intermediate states, i.e., all states  $|f\rangle$  into which the  $|\bar{P}^0\rangle$  and  $|P^0\rangle$  can both decay. In the SM, the  $\Delta F = 1$  Hamiltonian is generated through a single  $W^\pm$  emission, as shown in Fig. 4 for charm decay. If  $\mathcal{CP}$  were an exact symmetry,  $M_{12}$  and  $\Gamma_{12}$  would also be real parameters.

The physical eigenstates of  $\mathcal{M}$  are

$$|P_\mp\rangle = \frac{1}{\sqrt{|p|^2 + |q|^2}} \left[ p|P^0\rangle \mp q|\bar{P}^0\rangle \right], \quad (51)$$

with

$$\frac{q}{p} \equiv \frac{1 - \bar{\varepsilon}}{1 + \bar{\varepsilon}} = \left( \frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}} \right)^{1/2}. \quad (52)$$

If  $M_{12}$  and  $\Gamma_{12}$  were real then  $q/p = 1$  and the mass eigenstates  $|P_\mp\rangle$  would correspond to the  $\mathcal{CP}$ -even and  $\mathcal{CP}$ -odd states (we use the phase convention<sup>3</sup>  $\mathcal{CP}|P^0\rangle = -|\bar{P}^0\rangle$ )

$$|P_{1,2}\rangle \equiv \frac{1}{\sqrt{2}} \left( |P^0\rangle \mp |\bar{P}^0\rangle \right), \quad \mathcal{CP} |P_{1,2}\rangle = \pm |P_{1,2}\rangle. \quad (53)$$

<sup>3</sup> Since flavour is conserved by strong interactions, there is some freedom in defining the phases of flavour eigenstates. One could use  $|P_\zeta^0\rangle \equiv e^{-i\zeta}|P^0\rangle$  and  $|\bar{P}_\zeta^0\rangle \equiv e^{i\zeta}|\bar{P}^0\rangle$ , which satisfy  $\mathcal{CP}|P_\zeta^0\rangle = -e^{-2i\zeta}|\bar{P}_\zeta^0\rangle$ . Both basis are trivially related:  $M_{12}^\zeta = e^{2i\zeta}M_{12}$ ,  $\Gamma_{12}^\zeta = e^{2i\zeta}\Gamma_{12}$  and  $(q/p)_\zeta = e^{-2i\zeta}(q/p)$ . Thus,  $q/p \neq 1$  does not necessarily imply  $\mathcal{CP}$  violation.  $\mathcal{CP}$  is violated if  $|q/p| \neq 1$ ; i.e.,  $\text{Re}(\bar{\varepsilon}) \neq 0$  and  $\langle P_- | P_+ \rangle \neq 0$ . Note that  $\langle P_- | P_+ \rangle_\zeta = \langle P_- | P_+ \rangle$ . Another phase-convention-independent quantity is  $(q/p)(\bar{A}_f/A_f)$ , where  $A_f \equiv A(P^0 \rightarrow f)$  and  $\bar{A}_f \equiv -A(\bar{P}^0 \rightarrow f)$ , for any final state  $f$ .

The two mass eigenstates are no longer orthogonal when  $\mathcal{CP}$  is violated:

$$\langle P_- | P_+ \rangle = \frac{|p|^2 - |q|^2}{|p|^2 + |q|^2} = \frac{2 \operatorname{Re}(\bar{\varepsilon})}{(1 + |\bar{\varepsilon}|^2)}. \quad (54)$$

The time evolution of a state which was originally produced as a  $P^0$  or a  $\bar{P}^0$  is given by

$$\begin{pmatrix} |P^0(t)\rangle \\ |\bar{P}^0(t)\rangle \end{pmatrix} = \begin{pmatrix} g_1(t) & \frac{q}{p} g_2(t) \\ \frac{p}{q} g_2(t) & g_1(t) \end{pmatrix} \begin{pmatrix} |P^0\rangle \\ |\bar{P}^0\rangle \end{pmatrix}, \quad (55)$$

where

$$\begin{pmatrix} g_1(t) \\ g_2(t) \end{pmatrix} = e^{-iMt} e^{-\Gamma t/2} \begin{pmatrix} \cos[(x - iy)\Gamma t/2] \\ -i \sin[(x - iy)\Gamma t/2] \end{pmatrix}, \quad (56)$$

with  $M = \frac{1}{2}(M_{P_+} + M_{P_-})$ ,  $\Gamma = \frac{1}{2}(\Gamma_{P_+} + \Gamma_{P_-})$ ,

$$x \equiv \frac{\Delta M}{\Gamma}, \quad y \equiv \frac{\Delta \Gamma}{2\Gamma}, \quad (57)$$

and<sup>4</sup>

$$\Delta M \equiv M_{P_+} - M_{P_-}, \quad \Delta \Gamma \equiv \Gamma_{P_+} - \Gamma_{P_-}. \quad (58)$$

## 5.2 Experimental measurements

The main difference between the  $K^0-\bar{K}^0$  and  $B^0-\bar{B}^0$  systems stems from the different kinematics involved. The light kaon mass only allows the hadronic decay modes  $K^0 \rightarrow 2\pi$  and  $K^0 \rightarrow 3\pi$ . Since  $\mathcal{CP}|\pi\pi\rangle = +|\pi\pi\rangle$ , for both  $\pi^0\pi^0$  and  $\pi^+\pi^-$  final states, the  $\mathcal{CP}$ -even kaon state decays into  $2\pi$  whereas the  $\mathcal{CP}$ -odd one decays into the phase-space-suppressed  $3\pi$  mode. Therefore, there is a large lifetime difference and we have a short-lived  $|K_S\rangle \equiv |K_- \rangle \approx |K_1\rangle + \bar{\varepsilon}_K |K_2\rangle$  and a long-lived  $|K_L\rangle \equiv |K_+ \rangle \approx |K_2\rangle + \bar{\varepsilon}_K |K_1\rangle$  kaon, with  $\Gamma_{K_L} \ll \Gamma_{K_S} \approx 2\Gamma_{K^0}$ . One finds experimentally that  $\Delta\Gamma_{K^0} \approx -\Gamma_{K_S} \approx -2\Delta M_{K^0}$  [9]:

$$\Delta M_{K^0} = (0.5293 \pm 0.0009) \cdot 10^{10} \text{ s}^{-1}, \quad \Delta\Gamma_{K^0} = -(1.1149 \pm 0.0005) \cdot 10^{10} \text{ s}^{-1}. \quad (59)$$

Thus, the two  $K^0-\bar{K}^0$  oscillations parameters are sizeable and of similar magnitudes:  $x_{K^0} \approx -y_{K^0} \approx 1$ .

In the  $B$  system, there are many open decay channels and a large part of them are common to both mass eigenstates. Therefore, the  $|B_{\mp}\rangle$  states have a similar lifetime; i.e.,  $|\Delta\Gamma_{B^0}| \ll \Gamma_{B^0}$ . Moreover, whereas the  $B^0-\bar{B}^0$  transition is dominated by the top box diagram, the decay amplitudes get obviously their main contribution from the  $b \rightarrow c$  process. Thus,  $|\Delta\Gamma_{B^0}/\Delta M_{B^0}| \sim m_b^2/m_t^2 \ll 1$ . To experimentally measure the mixing transition requires the identification of the  $B$ -meson flavour at both its production and decay time. This can be done through flavour-specific decays such as  $B^0 \rightarrow X\ell^+\nu_\ell$  and  $\bar{B}^0 \rightarrow X\ell^-\bar{\nu}_\ell$ , where the lepton charge labels the initial  $B$  meson. In general, mixing is measured by studying pairs of  $B$  mesons so that one  $B$  can be used to *tag* the initial flavour of the other meson. For instance, in  $e^+e^-$  machines one can look into the pair production process  $e^+e^- \rightarrow B^0\bar{B}^0 \rightarrow (X\ell\nu_\ell)(Y\ell\nu_\ell)$ . In the absence of mixing, the final leptons should have opposite charges; the amount of like-sign leptons is then a clear signature of meson mixing.

Evidence for a large  $B_d^0-\bar{B}_d^0$  mixing was first reported in 1987 by ARGUS [102]. This provided the first indication that the top quark was very heavy. Since then, many experiments have analysed the mixing probability. The present world-average values are [9, 43]:

$$\Delta M_{B_d^0} = (0.5064 \pm 0.0019) \cdot 10^{12} \text{ s}^{-1}, \quad x_{B_d^0} = 0.770 \pm 0.004, \quad (60)$$

<sup>4</sup>Be aware of the different sign conventions in the literature. Quite often,  $\Delta M$  and  $\Delta\Gamma$  are defined to be positive.



while  $y_{B_d^0} = 0.001 \pm 0.005$  confirms the expected suppression of  $\Delta\Gamma_{B^0}$ .

The first direct evidence of  $B_s^0-\bar{B}_s^0$  oscillations was obtained by CDF [103]. The large measured mass difference reflects the CKM hierarchy  $|\mathbf{V}_{ts}|^2 \gg |\mathbf{V}_{td}|^2$ , implying very fast oscillations [9, 43]:

$$\begin{aligned} \Delta M_{B_s^0} &= (17.757 \pm 0.021) \cdot 10^{12} \text{ s}^{-1}, & x_{B_s^0} &= 26.72 \pm 0.09, \\ \Delta\Gamma_{B_s^0} &= -(0.086 \pm 0.006) \cdot 10^{12} \text{ s}^{-1}, & y_{B_s^0} &= -0.065 \pm 0.005. \end{aligned} \quad (61)$$

Evidence of mixing has been also obtained in the  $D^0-\bar{D}^0$  system. The present world averages [43],

$$x_{D^0} = -(0.32 \pm 0.14) \cdot 10^{-2}, \quad y_{D^0} = -\left(0.69_{-0.07}^{+0.06}\right) \cdot 10^{-2}, \quad (62)$$

confirm the SM expectation of a very slow oscillation, compared with the decay rate. Since the short-distance mixing amplitude originates in box diagrams with down-type quarks in the internal lines, it is very suppressed by the relevant combination of CKM factors and quark masses.

### 5.3 Mixing constraints on the CKM matrix

Long-distance contributions arising from intermediate hadronic states completely dominate the  $D^0-\bar{D}^0$  mixing amplitude and are very sizeable for  $\Delta M_{K^0}$ , making difficult to extract useful information on the CKM matrix. The situation is much better for  $B^0$  mesons, owing to the dominance of the short-distance top contribution which is known to next-to-leading order (NLO) in the strong coupling [104, 105]. The main uncertainty stems from the hadronic matrix element of the  $\Delta B = 2$  four-quark operator

$$\langle \bar{B}^0 | (\bar{b}\gamma^\mu(1-\gamma_5)d) (\bar{b}\gamma_\mu(1-\gamma_5)d) | B^0 \rangle \equiv \frac{8}{3} M_{B^0}^2 \xi_B^2, \quad (63)$$

which is characterized through the non-perturbative parameter  $\xi_B(\mu) \equiv f_B \sqrt{B_B(\mu)}$  [106]. The current  $(2+1)$  lattice averages [36] are  $\hat{\xi}_{B_d} = (225 \pm 9) \text{ MeV}$ ,  $\hat{\xi}_{B_s} = (274 \pm 8) \text{ MeV}$  and  $\hat{\xi}_{B_s}/\hat{\xi}_{B_d} = 1.206 \pm 0.017$ , where  $\hat{\xi}_B \approx \alpha_s(\mu)^{-3/23} \xi_B(\mu)$  is the corresponding renormalization-group-invariant quantity. Using these values, the measured mass differences in (60) and (61) imply

$$|\mathbf{V}_{tb}^* \mathbf{V}_{td}| = 0.0080 \pm 0.0003, \quad |\mathbf{V}_{tb}^* \mathbf{V}_{ts}| = 0.0388 \pm 0.0012, \quad \frac{|\mathbf{V}_{td}|}{|\mathbf{V}_{ts}|} = 0.205 \pm 0.003. \quad (64)$$

The last number takes advantage of the smaller uncertainty in the ratio  $\hat{\xi}_{B_s}/\hat{\xi}_{B_d}$ . Since  $|\mathbf{V}_{tb}| \approx 1$ , the mixing of  $B_{d,s}^0$  mesons provides indirect determinations of  $|\mathbf{V}_{td}|$  and  $|\mathbf{V}_{ts}|$ . The resulting value of  $|\mathbf{V}_{ts}|$  is in agreement with Eq. (30), satisfying the unitarity constraint  $|\mathbf{V}_{ts}| \approx |\mathbf{V}_{cb}|$ . In terms of the  $(\rho, \eta)$  parametrization of Eq. (43), one obtains

$$\sqrt{(1-\rho)^2 + \eta^2} = \begin{cases} \left| \frac{\mathbf{V}_{td}}{\lambda \mathbf{V}_{cb}} \right| = 0.86 \pm 0.04 \\ \left| \frac{\mathbf{V}_{td}}{\lambda \mathbf{V}_{ts}} \right| = 0.920 \pm 0.013 \end{cases}. \quad (65)$$

## 6 CP violation

While parity and charge conjugation are violated by the weak interactions in a maximal way, the product of the two discrete transformations is still a good symmetry of the gauge interactions (left-handed fermions  $\leftrightarrow$  right-handed antifermions). In fact,  $\mathcal{CP}$  appears to be a symmetry of nearly all observed phenomena. However, a slight violation of the  $\mathcal{CP}$  symmetry at the level of 0.2% is observed in the neutral kaon system and more sizeable signals of  $\mathcal{CP}$  violation have been established at the B factories.

Moreover, the huge matter–antimatter asymmetry present in our Universe is a clear manifestation of  $\mathcal{CP}$  violation and its important role in the primordial baryogenesis.

The  $\mathcal{CPT}$  theorem guarantees that the product of the three discrete transformations is an exact symmetry of any local and Lorentz-invariant quantum field theory, preserving micro-causality. A violation of  $\mathcal{CP}$  implies then a corresponding violation of time reversal. Since  $\mathcal{T}$  is an antiunitary transformation, this requires the presence of relative complex phases between different interfering amplitudes.

The electroweak SM Lagrangian only contains a single complex phase  $\delta_{13}$  ( $\eta$ ). This is the sole possible source of  $\mathcal{CP}$  violation and, therefore, the SM predictions for  $\mathcal{CP}$ -violating phenomena are quite constrained. The CKM mechanism requires several necessary conditions in order to generate an observable  $\mathcal{CP}$ -violation effect. With only two fermion generations, the quark mixing matrix cannot give rise to  $\mathcal{CP}$  violation; therefore, for  $\mathcal{CP}$  violation to occur in a particular process, all three generations are required to play an active role. In the kaon system, for instance,  $\mathcal{CP}$  violation can only appear at the one-loop level, where the top quark is present. In addition, all CKM matrix elements must be non-zero and the quarks of a given charge must be non-degenerate in mass. If any of these conditions were not satisfied, the CKM phase could be rotated away by a redefinition of the quark fields.  $\mathcal{CP}$ -violation effects are then necessarily proportional to the product of all CKM angles, and should vanish in the limit where any two (equal-charge) quark masses are taken to be equal. All these necessary conditions can be summarized as a single requirement on the original quark mass matrices  $\mathbf{M}'_u$  and  $\mathbf{M}'_d$  [107]:

$$\mathcal{CP} \text{ violation} \quad \iff \quad \text{Im} \left\{ \det \left[ \mathbf{M}'_u \mathbf{M}'_u{}^\dagger, \mathbf{M}'_d \mathbf{M}'_d{}^\dagger \right] \right\} \neq 0. \quad (66)$$

Without performing any detailed calculation, one can make the following general statements on the implications of the CKM mechanism of  $\mathcal{CP}$  violation:

- Owing to unitarity, for any choice of  $i, j, k, l$  (between 1 and 3),

$$\text{Im} \left[ \mathbf{V}_{ij} \mathbf{V}_{ik}^* \mathbf{V}_{lk} \mathbf{V}_{lj}^* \right] = \mathcal{J} \sum_{m,n=1}^3 \epsilon_{ilm} \epsilon_{jkn}, \quad (67)$$

$$\mathcal{J} = c_{12} c_{23} c_{13}^2 s_{12} s_{23} s_{13} \sin \delta_{13} \approx A^2 \lambda^6 \eta < 10^{-4}. \quad (68)$$

Any  $\mathcal{CP}$ -violation observable involves the product  $\mathcal{J}$  [107]. Thus, violations of the  $\mathcal{CP}$  symmetry are necessarily small.

- In order to have sizeable  $\mathcal{CP}$ -violating asymmetries  $\mathcal{A} \equiv (\Gamma - \bar{\Gamma})/(\Gamma + \bar{\Gamma})$ , one should look for very suppressed decays, where the decay widths already involve small CKM matrix elements.
- In the SM,  $\mathcal{CP}$  violation is a low-energy phenomenon, in the sense that any effect should disappear when the quark mass difference  $m_c - m_u$  becomes negligible.
- $B$  decays are the optimal place for  $\mathcal{CP}$ -violation signals to show up. They involve small CKM matrix elements and are the lowest-mass processes where the three quark generations play a direct (tree-level) role.

The SM mechanism of  $\mathcal{CP}$  violation is based on the unitarity of the CKM matrix. Testing the constraints implied by unitarity is then a way to test the source of  $\mathcal{CP}$  violation. The unitarity tests in Eqs. (37), (38), (39) and (40) involve only the moduli of the CKM parameters, while  $\mathcal{CP}$  violation has to do with their phases. More interesting are the off-diagonal unitarity conditions:

$$\begin{aligned} \mathbf{V}_{ud}^* \mathbf{V}_{us} + \mathbf{V}_{cd}^* \mathbf{V}_{cs} + \mathbf{V}_{td}^* \mathbf{V}_{ts} &= 0, \\ \mathbf{V}_{us}^* \mathbf{V}_{ub} + \mathbf{V}_{cs}^* \mathbf{V}_{cb} + \mathbf{V}_{ts}^* \mathbf{V}_{tb} &= 0, \\ \mathbf{V}_{ub}^* \mathbf{V}_{ud} + \mathbf{V}_{cb}^* \mathbf{V}_{cd} + \mathbf{V}_{tb}^* \mathbf{V}_{td} &= 0. \end{aligned} \quad (69)$$



which establishes the presence of *indirect*  $\mathcal{CP}$  violation generated by the mixing amplitude.

If the flavour of the decaying meson  $P$  is known, any observed difference between the decay rate  $\Gamma(P \rightarrow f)$  and its  $\mathcal{CP}$  conjugate  $\Gamma(\bar{P} \rightarrow \bar{f})$  would indicate that  $\mathcal{CP}$  is directly violated in the decay amplitude. One could study, for instance,  $\mathcal{CP}$  asymmetries in decays such as  $K^\pm \rightarrow \pi^\pm \pi^0$  where the pion charges identify the kaon flavour; however, no positive signals have been found in charged kaon decays. Since at least two interfering contributions are needed, let us write the decay amplitudes as

$$A[P \rightarrow f] = M_1 e^{i\phi_1} e^{i\delta_1} + M_2 e^{i\phi_2} e^{i\delta_2}, \quad A[\bar{P} \rightarrow \bar{f}] = M_1 e^{-i\phi_1} e^{i\delta_1} + M_2 e^{-i\phi_2} e^{i\delta_2}, \quad (73)$$

where  $\phi_i$  denote weak phases,  $\delta_i$  strong final-state interaction phases and  $M_i$  the moduli of the matrix elements. The rate asymmetry is given by

$$\mathcal{A}_{P \rightarrow f}^{\mathcal{CP}} \equiv \frac{\Gamma[P \rightarrow f] - \Gamma[\bar{P} \rightarrow \bar{f}]}{\Gamma[P \rightarrow f] + \Gamma[\bar{P} \rightarrow \bar{f}]} = \frac{-2M_1 M_2 \sin(\phi_1 - \phi_2) \sin(\delta_1 - \delta_2)}{|M_1|^2 + |M_2|^2 + 2M_1 M_2 \cos(\phi_1 - \phi_2) \cos(\delta_1 - \delta_2)}. \quad (74)$$

Thus, to generate a direct  $\mathcal{CP}$  asymmetry one needs: 1) at least two interfering amplitudes, which should be of comparable size in order to get a sizeable asymmetry; 2) two different weak phases [ $\sin(\phi_1 - \phi_2) \neq 0$ ], and 3) two different strong phases [ $\sin(\delta_1 - \delta_2) \neq 0$ ].

Direct  $\mathcal{CP}$  violation has been searched for in decays of neutral kaons, where  $K^0 - \bar{K}^0$  mixing is also involved. Thus, both direct and indirect  $\mathcal{CP}$  violation need to be taken into account simultaneously. A  $\mathcal{CP}$ -violation signal is provided by the ratios:

$$\eta_{+-} \equiv \frac{A(K_L \rightarrow \pi^+ \pi^-)}{A(K_S \rightarrow \pi^+ \pi^-)} = \varepsilon_K + \varepsilon'_K, \quad \eta_{00} \equiv \frac{A(K_L \rightarrow \pi^0 \pi^0)}{A(K_S \rightarrow \pi^0 \pi^0)} = \varepsilon_K - 2\varepsilon'_K. \quad (75)$$

The dominant effect from  $\mathcal{CP}$  violation in  $K^0 - \bar{K}^0$  mixing is contained in  $\varepsilon_K$ , while  $\varepsilon'_K$  accounts for direct  $\mathcal{CP}$  violation in the decay amplitudes [39]:

$$\varepsilon_K = \bar{\varepsilon}_K + i\xi_0, \quad \varepsilon'_K = \frac{i}{\sqrt{2}} \omega (\xi_2 - \xi_0), \quad \omega \equiv \frac{\text{Re}(A_2)}{\text{Re}(A_0)} e^{i(\delta_2 - \delta_0)}, \quad \xi_I \equiv \frac{\text{Im}(A_I)}{\text{Re}(A_I)}. \quad (76)$$

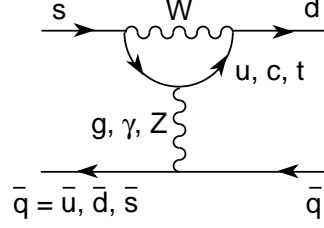
$A_I$  are the transition amplitudes into two pions with isospin  $I = 0, 2$  (these are the only two values allowed by Bose symmetry for the final  $2\pi$  state) and  $\delta_I$  their corresponding strong phase shifts. Although  $\varepsilon'_K$  is strongly suppressed by the small ratio  $|\omega| \approx 1/22$ , a non-zero value has been established through very accurate measurements, demonstrating the existence of direct  $\mathcal{CP}$  violation in K decays [109–112]:

$$\text{Re}(\varepsilon'_K/\varepsilon_K) = \frac{1}{3} \left( 1 - \left| \frac{\eta_{00}}{\eta_{+-}} \right| \right) = (16.8 \pm 1.4) \cdot 10^{-4}. \quad (77)$$

In the SM the necessary weak phases are generated through the gluonic and electroweak penguin diagrams shown in Fig. 7, involving virtual up-type quarks of the three generations in the loop. These short-distance contributions are known to NLO in the strong coupling [113, 114]. However, the theoretical prediction involves a delicate balance between the two isospin amplitudes and is sensitive to long-distance and isospin-violating effects. Using chiral perturbation theory techniques, one finds  $\text{Re}(\varepsilon'_K/\varepsilon_K) = (15 \pm 7) \cdot 10^{-4}$  [115–118], in agreement with (77) but with a large uncertainty.

Since  $\text{Re}(\varepsilon'_K/\varepsilon_K) \ll 1$ , the ratios  $\eta_{+-}$  and  $\eta_{00}$  provide a measurement of  $\varepsilon_K = |\varepsilon_K| e^{i\phi_\varepsilon}$  [9]:

$$|\varepsilon_K| = \frac{1}{3} \left( 2|\eta_{+-}| + |\eta_{00}| \right) = (2.228 \pm 0.011) \cdot 10^{-3}, \quad \phi_\varepsilon = (43.52 \pm 0.05)^\circ, \quad (78)$$


**Fig. 7:**  $\Delta S = 1$  penguin diagrams.

in perfect agreement with the semileptonic asymmetry  $\delta_L$ . In the SM  $\varepsilon_K$  receives short-distance contributions from box diagrams involving virtual top and charm quarks, which are proportional to

$$\varepsilon_K \propto \sum_{i,j=c,t} \eta_{ij} \text{Im}[\mathbf{V}_{id}\mathbf{V}_{is}^*\mathbf{V}_{jd}\mathbf{V}_{js}^*] S(r_i, r_j) \propto A^2 \lambda^6 \bar{\eta} \left\{ \eta_{tt} A^2 \lambda^4 (1 - \bar{\rho}) + P_c \right\}. \quad (79)$$

The first term shows the CKM dependence of the dominant top contribution,  $P_c$  accounts for the charm corrections [119] and the short-distance QCD corrections  $\eta_{ij}$  are known to NLO [104, 105, 120]. The measured value of  $|\varepsilon_K|$  determines an hyperbolic constraint in the  $(\bar{\rho}, \bar{\eta})$  plane, shown in Fig. 6, taking into account the theoretical uncertainty in the hadronic matrix element of the  $\Delta S = 2$  operator [36].

## 6.2 $\mathcal{CP}$ asymmetries in B decays

The semileptonic decays  $B^0 \rightarrow X^- \ell^+ \nu_\ell$  and  $\bar{B}^0 \rightarrow X^+ \ell^- \bar{\nu}_\ell$  provide the most direct way to measure the amount of  $\mathcal{CP}$  violation in the  $B^0$ - $\bar{B}^0$  mixing matrix, through

$$\begin{aligned} a_{\text{sl}}^q &\equiv \frac{\Gamma(\bar{B}_q^0 \rightarrow X^- \ell^+ \nu_\ell) - \Gamma(B_q^0 \rightarrow X^+ \ell^- \bar{\nu}_\ell)}{\Gamma(\bar{B}_q^0 \rightarrow X^- \ell^+ \nu_\ell) + \Gamma(B_q^0 \rightarrow X^+ \ell^- \bar{\nu}_\ell)} = \frac{|p|^4 - |q|^4}{|p|^4 + |q|^4} \approx 4 \text{Re}(\bar{\varepsilon}_{B_q^0}) \\ &\approx \frac{|\Gamma_{12}|}{|M_{12}|} \sin \phi_q \approx \frac{|\Delta \Gamma_{B_q^0}|}{|\Delta M_{B_q^0}|} \tan \phi_q. \end{aligned} \quad (80)$$

This asymmetry is expected to be tiny because  $|\Gamma_{12}/M_{12}| \sim m_b^2/m_t^2 \ll 1$ . Moreover, there is an additional GIM suppression in the relative mixing phase  $\phi_q \equiv \arg(-M_{12}/\Gamma_{12}) \sim (m_c^2 - m_u^2)/m_b^2$ , implying a value of  $|q/p|$  very close to 1. Therefore,  $a_{\text{sl}}^q$  could be very sensitive to new sources of  $\mathcal{CP}$  violation beyond the SM, contributing to  $\phi_q$ . The present measurements give [9, 43]

$$\text{Re}(\bar{\varepsilon}_{B_d^0}) = (-0.5 \pm 0.4) \cdot 10^{-3}, \quad \text{Re}(\bar{\varepsilon}_{B_s^0}) = (-0.15 \pm 0.70) \cdot 10^{-3}. \quad (81)$$

The large  $B^0$ - $\bar{B}^0$  mixing provides a different way to generate the required  $\mathcal{CP}$ -violating interference. There are quite a few nonleptonic final states which are reachable both from a  $B^0$  and a  $\bar{B}^0$ . For these flavour non-specific decays the  $B^0$  (or  $\bar{B}^0$ ) can decay directly to the given final state  $f$ , or do it after the meson has been changed to its antiparticle via the mixing process; i.e., there are two different amplitudes,  $A(B^0 \rightarrow f)$  and  $A(B^0 \rightarrow \bar{B}^0 \rightarrow f)$ , corresponding to two possible decay paths.  $\mathcal{CP}$ -violating effects can then result from the interference of these two contributions.

The time-dependent decay probabilities for the decay of a neutral  $B$  meson created at the time  $t_0 = 0$  as a pure  $B^0$  ( $\bar{B}^0$ ) into the final state  $f$  ( $\bar{f} \equiv \mathcal{CP} f$ ) are:

$$\begin{aligned} \Gamma[B^0(t) \rightarrow f] &\propto \frac{1}{2} e^{-\Gamma_{B^0} t} \left( |A_f|^2 + |\bar{A}_f|^2 \right) \left\{ 1 + C_f \cos(\Delta M_{B^0} t) - S_f \sin(\Delta M_{B^0} t) \right\}, \\ \Gamma[\bar{B}^0(t) \rightarrow \bar{f}] &\propto \frac{1}{2} e^{-\Gamma_{B^0} t} \left( |\bar{A}_{\bar{f}}|^2 + |A_{\bar{f}}|^2 \right) \left\{ 1 - C_{\bar{f}} \cos(\Delta M_{B^0} t) + S_{\bar{f}} \sin(\Delta M_{B^0} t) \right\}, \end{aligned} \quad (82)$$

where the tiny  $\Delta\Gamma_{B^0}$  corrections have been neglected and we have introduced the notation

$$\begin{aligned} A_f &\equiv A[B^0 \rightarrow f], & \bar{A}_f &\equiv -A[\bar{B}^0 \rightarrow f], & \bar{\rho}_f &\equiv \bar{A}_f/A_f, \\ A_{\bar{f}} &\equiv A[B^0 \rightarrow \bar{f}], & \bar{A}_{\bar{f}} &\equiv -A[\bar{B}^0 \rightarrow \bar{f}], & \rho_{\bar{f}} &\equiv A_{\bar{f}}/\bar{A}_{\bar{f}}, \\ C_f &\equiv \frac{1 - |\bar{\rho}_f|^2}{1 + |\bar{\rho}_f|^2}, & S_f &\equiv \frac{2 \operatorname{Im}\left(\frac{q}{p} \bar{\rho}_f\right)}{1 + |\bar{\rho}_f|^2}, & C_{\bar{f}} &\equiv -\frac{1 - |\rho_{\bar{f}}|^2}{1 + |\rho_{\bar{f}}|^2}, & S_{\bar{f}} &\equiv \frac{-2 \operatorname{Im}\left(\frac{p}{q} \rho_{\bar{f}}\right)}{1 + |\rho_{\bar{f}}|^2}. \end{aligned} \quad (83)$$

$\mathcal{CP}$  invariance demands the probabilities of  $\mathcal{CP}$ -conjugate processes to be identical. Thus,  $\mathcal{CP}$  conservation requires  $A_f = \bar{A}_{\bar{f}}$ ,  $A_{\bar{f}} = \bar{A}_f$ ,  $\bar{\rho}_f = \rho_{\bar{f}}$  and  $\operatorname{Im}\left(\frac{q}{p} \bar{\rho}_f\right) = \operatorname{Im}\left(\frac{p}{q} \rho_{\bar{f}}\right)$ , i.e.,  $C_f = -C_{\bar{f}}$  and  $S_f = -S_{\bar{f}}$ . Violation of any of the first three equalities would be a signal of direct  $\mathcal{CP}$  violation. The fourth equality tests  $\mathcal{CP}$  violation generated by the interference of the direct decay  $B^0 \rightarrow f$  and the mixing-induced decay  $B^0 \rightarrow \bar{B}^0 \rightarrow f$ .

For  $B^0$  mesons

$$\left. \frac{q}{p} \right|_{B_q^0} \approx \sqrt{\frac{M_{12}^*}{M_{12}}} \approx \frac{\mathbf{V}_{tb}^* \mathbf{V}_{tq}}{\mathbf{V}_{tb} \mathbf{V}_{tq}^*} \equiv e^{-2i\phi_q^M}, \quad (84)$$

where  $\phi_d^M = \beta + \mathcal{O}(\lambda^4)$  and  $\phi_s^M = -\beta_s + \mathcal{O}(\lambda^6)$ . The angle  $\beta$  is defined in Eq. (70), while  $\beta_s \equiv \arg[-(\mathbf{V}_{ts} \mathbf{V}_{tb}^*) / (\mathbf{V}_{cs} \mathbf{V}_{cb}^*)] = \lambda^2 \eta + \mathcal{O}(\lambda^4)$  is the equivalent angle in the  $B_s^0$  unitarity triangle, which is predicted to be tiny. Therefore, the mixing ratio  $q/p$  is given by a known weak phase.

An obvious example of final states  $f$  which can be reached both from the  $B^0$  and the  $\bar{B}^0$  are  $\mathcal{CP}$  eigenstates; i.e., states such that  $\bar{f} = \zeta_f f$  ( $\zeta_f = \pm 1$ ). In this case,  $A_{\bar{f}} = \zeta_f A_f$ ,  $\bar{A}_{\bar{f}} = \zeta_f \bar{A}_f$ ,  $\rho_{\bar{f}} = 1/\bar{\rho}_f$ ,  $C_{\bar{f}} = C_f$  and  $S_{\bar{f}} = S_f$ . A non-zero value of  $C_f$  or  $S_f$  signals then  $\mathcal{CP}$  violation. The ratios  $\bar{\rho}_f$  and  $\rho_{\bar{f}}$  depend in general on the underlying strong dynamics. However, for  $\mathcal{CP}$  self-conjugate final states, all dependence on the strong interaction disappears if only one weak amplitude contributes to the  $B^0 \rightarrow f$  and  $\bar{B}^0 \rightarrow f$  transitions [121, 122]. In this case, we can write the decay amplitude as  $A_f = M e^{i\phi^D} e^{i\delta_s}$ , with  $M = M^*$  and  $\phi^D$  and  $\delta_s$  weak and strong phases. The ratios  $\bar{\rho}_f$  and  $\rho_{\bar{f}}$  are then given by

$$\rho_{\bar{f}} = \bar{\rho}_f^* = \zeta_f e^{2i\phi^D}. \quad (85)$$

The modulus  $M$  and the unwanted strong phase cancel out completely from these two ratios;  $\rho_{\bar{f}}$  and  $\bar{\rho}_f$  simplify to a single weak phase, associated with the underlying weak quark transition. Since  $|\rho_{\bar{f}}| = |\bar{\rho}_f| = 1$ , the time-dependent decay probabilities become much simpler. In particular,  $C_f = 0$  and there is no longer any dependence on  $\cos(\Delta M_{B^0} t)$ . Moreover, the coefficients of the sinusoidal terms are then fully known in terms of CKM mixing angles only:  $S_f = S_{\bar{f}} = -\zeta_f \sin[2(\phi_q^M + \phi^D)] \equiv -\zeta_f \sin(2\Phi)$ . In this ideal case, the time-dependent  $\mathcal{CP}$ -violating decay asymmetry

$$\mathcal{A}_{\bar{B}^0 \rightarrow \bar{f}}^{\mathcal{CP}} \equiv \frac{\Gamma[\bar{B}^0(t) \rightarrow \bar{f}] - \Gamma[B^0(t) \rightarrow f]}{\Gamma[\bar{B}^0(t) \rightarrow \bar{f}] + \Gamma[B^0(t) \rightarrow f]} = -\zeta_f \sin(2\Phi) \sin(\Delta M_{B^0} t) \quad (86)$$

provides a direct and clean measurement of the CKM parameters [123].

When several decay amplitudes with different phases contribute,  $|\bar{\rho}_f| \neq 1$  and the interference term will depend both on CKM parameters and on the strong dynamics embodied in  $\bar{\rho}_f$ . The leading contributions to  $\bar{b} \rightarrow \bar{q}' q' \bar{q}$  are either the tree-level  $W$  exchange or penguin topologies generated by gluon ( $\gamma$ ,  $Z$ ) exchange. Although of higher order in the strong (electroweak) coupling, penguin amplitudes are logarithmically enhanced by the virtual  $W$  loop and are potentially competitive. Table 3 contains the CKM factors associated with the two topologies for different  $B$  decay modes into  $\mathcal{CP}$  eigenstates.

The gold-plated decay mode is  $B_d^0 \rightarrow J/\psi K_S$ . In addition of having a clean experimental signature, the two topologies have the same (zero) weak phase. The  $\mathcal{CP}$  asymmetry provides then a clean

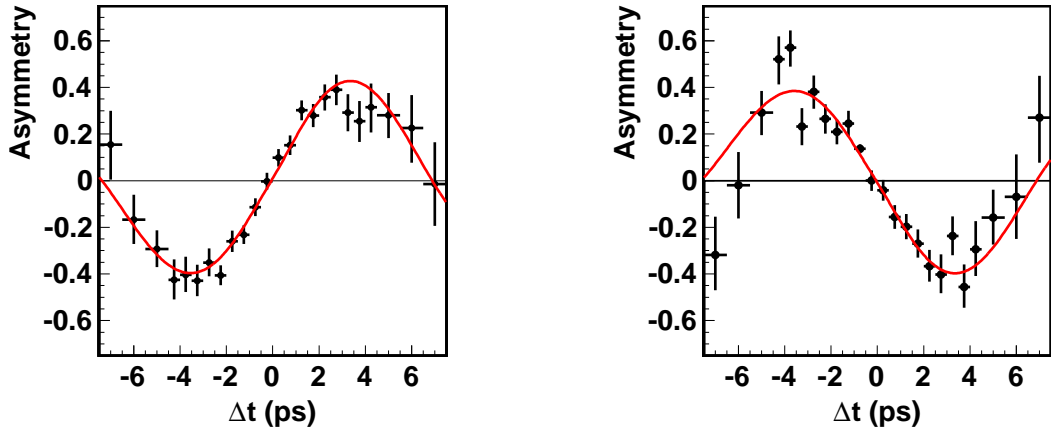
**Table 3:** CKM factors and relevant angle  $\Phi$  for some  $B$  decays into  $\mathcal{CP}$  eigenstates.

Decay	Tree-level CKM	Penguin CKM	Exclusive channels	$\Phi$
$\bar{b} \rightarrow \bar{c}c\bar{s}$	$A\lambda^2$	$-A\lambda^2$	$B_d^0 \rightarrow J/\psi K_S, J/\psi K_L$ $B_s^0 \rightarrow D_s^+ D_s^-, J/\psi \phi$	$\beta$ $-\beta_s$
$\bar{b} \rightarrow \bar{s}s\bar{s}$		$-A\lambda^2$	$B_d^0 \rightarrow K_S \phi, K_L \phi$ $B_s^0 \rightarrow \phi \phi$	$\beta$ $-\beta_s$
$\bar{b} \rightarrow \bar{d}d\bar{s}$		$-A\lambda^2$	$B_s^0 \rightarrow K_S K_S, K_L K_L$	$-\beta_s$
$\bar{b} \rightarrow \bar{c}c\bar{d}$	$-A\lambda^3$	$A\lambda^3(1 - \rho - i\eta)$	$B_d^0 \rightarrow D^+ D^-, J/\psi \pi^0$ $B_s^0 \rightarrow J/\psi K_S, J/\psi K_L$	$\approx \beta$ $\approx -\beta_s$
$\bar{b} \rightarrow \bar{u}u\bar{d}$	$A\lambda^3(\rho + i\eta)$	$A\lambda^3(1 - \rho - i\eta)$	$B_d^0 \rightarrow \pi^+ \pi^-, \rho^0 \pi^0, \omega \pi^0$ $B_s^0 \rightarrow \rho^0 K_{S,L}, \omega K_{S,L}, \pi^0 K_{S,L}$	$\approx \beta + \gamma$ $\neq \gamma - \beta_s$
$\bar{b} \rightarrow \bar{s}s\bar{d}$		$A\lambda^3(1 - \rho - i\eta)$	$B_d^0 \rightarrow K_S K_S, K_L K_L, \phi \pi^0$ $B_s^0 \rightarrow K_S \phi, K_L \phi$	$0$ $-\beta - \beta_s$

measurement of the mixing angle  $\beta$ , without strong-interaction uncertainties. Fig. 8 shows the Belle measurement [124] of time-dependent  $\bar{b} \rightarrow \bar{c}c\bar{s}$  asymmetries for  $\mathcal{CP}$ -odd ( $B_d^0 \rightarrow J/\psi K_S, B_d^0 \rightarrow \psi' K_S, B_d^0 \rightarrow \chi_{c1} K_S$ ) and  $\mathcal{CP}$ -even ( $B_d^0 \rightarrow J/\psi K_L$ ) final states. A very nice oscillation is manifest, with opposite signs for the two different choices of  $\zeta_f = \pm 1$ . Including the information obtained from other  $\bar{b} \rightarrow \bar{c}c\bar{s}$  decays, one gets the world average [43]:

$$\sin(2\beta) = 0.691 \pm 0.017. \quad (87)$$

Fitting an additional  $\cos(\Delta M_{B^0} t)$  term in the measured asymmetries results in  $C_f = -0.004 \pm 0.015$  [43], confirming the expected null result. An independent measurement of  $\sin 2\beta$  can be obtained from  $\bar{b} \rightarrow \bar{s}s\bar{s}$  and  $\bar{b} \rightarrow \bar{d}d\bar{s}$  decays, which only receive penguin contributions and, therefore, could be more sensitive to new-physics corrections in the loop diagram. These modes give  $\sin(2\beta) = 0.655 \pm 0.032$  [43], in perfect agreement with (87).



**Fig. 8:** Time-dependent asymmetries for  $\mathcal{CP}$ -odd ( $B_d^0 \rightarrow J/\psi K_S, B_d^0 \rightarrow \psi' K_S, B_d^0 \rightarrow \chi_{c1} K_S; \zeta_f = -1$ ; left) and  $\mathcal{CP}$ -even ( $B_d^0 \rightarrow J/\psi K_L; \zeta_f = +1$ ; right) final states, measured by Belle [124].

Eq. (87) determines the angle  $\beta$  up to a four-fold ambiguity:  $\beta, \frac{\pi}{2} - \beta, \pi + \beta$  and  $\frac{3\pi}{2} - \beta$ . The solution  $\beta = (21.9 \pm 0.7)^\circ$  is in remarkable agreement with the other phenomenological constraints on the unitarity triangle in Fig. 6. The ambiguity has been resolved through a time-dependent analysis of the



Dalitz plot distribution in  $B_d^0 \rightarrow D^{(*)}h^0$  decays ( $h^0 = \pi^0, \eta, \omega$ ), showing that  $\cos(2\beta) = 0.91 \pm 0.25$  and  $\beta = (22.5 \pm 4.6)^\circ$  [125]. This proves that  $\cos(2\beta)$  is positive with a  $3.7\sigma$  significance, while the multifold solution  $\frac{\pi}{2} - \beta = (68.1 \pm 0.7)^\circ$  is excluded at the  $7.3\sigma$  level.

A determination of  $\beta + \gamma = \pi - \alpha$  can be obtained from  $\bar{b} \rightarrow \bar{u}u\bar{d}$  decays, such as  $B_d^0 \rightarrow \pi\pi$  or  $B_d^0 \rightarrow \rho\rho$ . However, the penguin contamination that carries a different weak phase can be sizeable. The time-dependent asymmetry in  $B_d^0 \rightarrow \pi^+\pi^-$  shows indeed a non-zero value for the  $\cos(\Delta M_{B^0} t)$  term,  $C_f = -0.27 \pm 0.04$  [43], providing evidence of direct  $\mathcal{CP}$  violation and indicating the presence of an additional decay amplitude; therefore,  $S_f = -0.68 \pm 0.04 \neq \sin 2\alpha$ . One could still extract useful information on  $\alpha$  (up to 16 mirror solutions), using the isospin relations among the  $B_d^0 \rightarrow \pi^+\pi^-$ ,  $B_d^0 \rightarrow \pi^0\pi^0$  and  $B^+ \rightarrow \pi^+\pi^0$  amplitudes and their  $\mathcal{CP}$  conjugates [126]; however, only a loose constraint is obtained, given the limited experimental precision on  $B_d^0 \rightarrow \pi^0\pi^0$ . Much stronger constraints are obtained from  $B_d^0 \rightarrow \rho^+\rho^-, \rho^0\rho^0$  because one can use the additional polarization information of two vectors in the final state to resolve the different contributions and, moreover, the small branching fraction  $\text{Br}(B_d^0 \rightarrow \rho^0\rho^0) = (9.6 \pm 1.5) \cdot 10^{-7}$  [9] implies a very small penguin contribution. Additional information can be obtained from  $B_d^0, \bar{B}_d^0 \rightarrow \rho^\pm\pi^\mp, a_1^\pm\pi^\mp$ , although the final states are not  $\mathcal{CP}$  eigenstates. Combining all pieces of information results in [9]

$$\alpha = (93 \pm 5)^\circ. \quad (88)$$

The angle  $\gamma$  cannot be determined in  $\bar{b} \rightarrow u\bar{u}\bar{d}$  decays such as  $B_s^0 \rightarrow \rho^0 K_S$  because the colour factors in the hadronic matrix element enhance the penguin amplitude with respect to the tree-level contribution. Instead,  $\gamma$  can be measured through the tree-level decays  $B \rightarrow D^{(*)}K^{(*)}$  ( $\bar{b} \rightarrow \bar{u}c\bar{s}$ ) and  $B \rightarrow \bar{D}^{(*)}K^{(*)}$  ( $\bar{b} \rightarrow \bar{c}u\bar{s}$ ), using final states accessible in both  $D^{(*)0}$  and  $\bar{D}^{(*)0}$  decays and playing with the interference of both amplitudes [127–129]. The sensitivity can be optimized with Dalitz-plot analyses of the  $D^0, \bar{D}^0$  decay products. The extensive experimental studies performed so far result in [43]

$$\gamma = (74.0_{-6.4}^{+5.8})^\circ. \quad (89)$$

Another ambiguous solution with  $\gamma \leftrightarrow \gamma + \pi$  also exists.

Mixing-induced  $\mathcal{CP}$  violation has been also searched for in the decays of  $B_s^0$  and  $\bar{B}_s^0$  mesons into  $J/\psi\phi, \psi(2S)\phi, J/\psi K^+K^-, J/\psi\pi^+\pi^-$  and  $D_s^+D_s^-$ . From the corresponding time-dependent  $\mathcal{CP}$  asymmetries,<sup>5</sup> one extracts correlated constraints on  $\Delta\Gamma_s$  and the weak phase  $\phi_s^{ccs} \equiv 2\Phi_s^{ccs} \approx 2\phi_s^M \approx -2\beta_s$  in Eq. (86), which are shown in Fig. 9. They lead to [43]

$$\beta_s = (0.86 \pm 0.95)^\circ, \quad (90)$$

in good agreement with the SM prediction  $\beta_s \approx \eta\lambda^2 \approx 1^\circ$ .

### 6.3 Global fit of the unitarity triangle

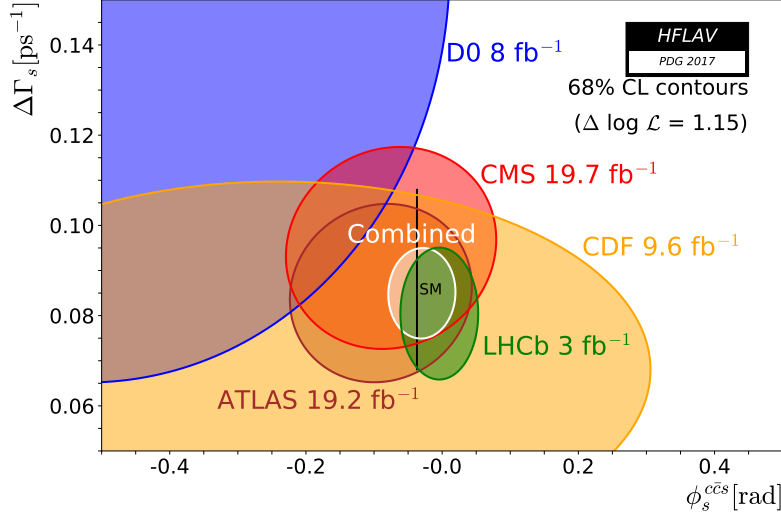
The CKM parameters can be more accurately determined through a global fit to all available measurements, imposing the unitarity constraints and taking properly into account the theoretical uncertainties. The global fit shown in Fig. 6 uses frequentist statistics and gives [108]

$$\lambda = 0.2251 \pm 0.0003, \quad A = 0.825_{-0.011}^{+0.007}, \quad \bar{\rho} = 0.160_{-0.007}^{+0.008}, \quad \bar{\eta} = 0.345 \pm 0.006. \quad (91)$$

This implies  $\mathcal{J} = (3.10_{-0.06}^{+0.05}) \cdot 10^{-5}$ ,  $\alpha = (92.0_{-1.1}^{+1.3})^\circ$ ,  $\beta = (22.6 \pm 0.4)^\circ$  and  $\gamma = (65.4_{-1.2}^{+1.0})^\circ$ . Similar results are obtained by the UTfit group [130], using instead a Bayesian approach and a slightly different treatment of theoretical uncertainties.

<sup>5</sup>The  $\Delta\Gamma_{B_s^0}$  corrections to Eq. (82) must be taken into account at the current level of precision.





**Fig. 9:** 68% CL regions in  $\Delta\Gamma_{B_s^0}$  and  $\phi_s^{c\bar{c}s}$ , extracted from  $\bar{b} \rightarrow \bar{c}c\bar{s}$   $\mathcal{CP}$  asymmetries of  $B_s^0$  mesons [43]. The vertical black line shows the SM prediction [131–133].

#### 6.4 Direct $\mathcal{CP}$ violation in $B$ decays

The big data samples accumulated at the B factories and the collider experiments have established the presence of direct  $\mathcal{CP}$  violation in several decays of  $B$  mesons. The most significant signals are [9, 43]

$$\begin{aligned}
 \mathcal{A}_{\bar{B}_d^0 \rightarrow K^- \pi^+}^{\mathcal{CP}} &= -0.082 \pm 0.006, & \mathcal{A}_{\bar{B}_d^0 \rightarrow \bar{K}^{*0} \eta}^{\mathcal{CP}} &= 0.19 \pm 0.05, & \mathcal{A}_{\bar{B}_d^0 \rightarrow K^{*-} \pi^+}^{\mathcal{CP}} &= -0.22 \pm 0.06, \\
 C_{B_d^0 \rightarrow \pi^+ \pi^-} &= -0.27 \pm 0.04, & \gamma_{B \rightarrow D^{(*)} K^{(*)}} &= (74.0^{+5.8}_{-6.4})^\circ, & \mathcal{A}_{B_s^0 \rightarrow K^+ \pi^-}^{\mathcal{CP}} &= 0.26 \pm 0.04, \\
 \mathcal{A}_{B^- \rightarrow K^- \rho^0}^{\mathcal{CP}} &= 0.37 \pm 0.10, & \mathcal{A}_{B^- \rightarrow K^- \eta}^{\mathcal{CP}} &= -0.37 \pm 0.08, & \mathcal{A}_{B^- \rightarrow K^- \pi^+ \pi^-}^{\mathcal{CP}} &= 0.027 \pm 0.008, \\
 \mathcal{A}_{B^- \rightarrow \pi^- \pi^+ \pi^-}^{\mathcal{CP}} &= 0.057 \pm 0.013, & \mathcal{A}_{B^- \rightarrow K^- K^+ K^-}^{\mathcal{CP}} &= -0.033 \pm 0.008, \\
 \mathcal{A}_{B^- \rightarrow K^- f_2(1270)}^{\mathcal{CP}} &= -0.68^{+0.19}_{-0.17}, & \mathcal{A}_{B^- \rightarrow K^- K^+ \pi^-}^{\mathcal{CP}} &= -0.118 \pm 0.022. \quad (92)
 \end{aligned}$$

Unfortunately, owing to the unavoidable presence of strong phases, a real theoretical understanding of the corresponding SM predictions is still lacking. Progress in this direction is needed to perform meaningful tests of the CKM mechanism of  $\mathcal{CP}$  violation and pin down any possible effects from new physics beyond the SM framework.

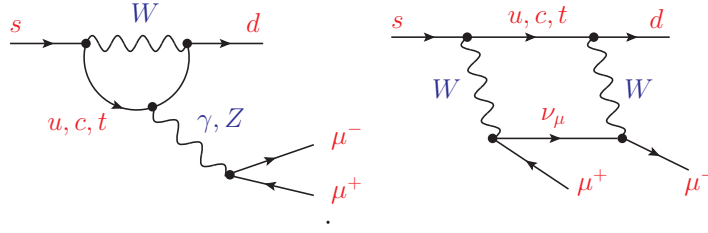
### 7 Rare decays

Complementary and very valuable information can be obtained from rare decays that in the SM are strongly suppressed by the GIM mechanism. These processes are then sensitive to new-physics contributions with a different flavour structure. Well-known examples are the  $\bar{K}^0 \rightarrow \mu^+ \mu^-$  decay modes [9, 134],

$$\text{Br}(K_L \rightarrow \mu^+ \mu^-) = (6.84 \pm 0.11) \cdot 10^{-9}, \quad \text{Br}(K_S \rightarrow \mu^+ \mu^-) < 1.0 \cdot 10^{-9} \quad (95\% \text{ CL}), \quad (93)$$

which tightly constrain any hypothetical flavour-changing ( $s \rightarrow d$ ) tree-level coupling of the  $Z$  boson. In the SM, these decays receive short-distance contributions from the penguin and box diagrams displayed in Fig. 10. Owing to the unitarity of the CKM matrix, the resulting amplitude vanishes for equal up-type quark masses, which entails a heavy suppression:

$$\mathcal{M} \propto \sum_{i=u,c,t} \mathbf{V}_{is} \mathbf{V}_{id}^* F(m_i^2/M_W^2) = \mathbf{V}_{cs} \mathbf{V}_{cd}^* \tilde{F}(m_c^2/M_W^2) + \mathbf{V}_{ts} \mathbf{V}_{td}^* \tilde{F}(m_t^2/M_W^2), \quad (94)$$



**Fig. 10:** Short-distance penguin (left) and box (right) SM contributions to  $\bar{K}^0 \rightarrow \mu^+ \mu^-$ .

where  $F(x)$  is a loop function and  $\tilde{F}(x) \equiv F(x) - F(0)$ . In addition to the loop factor  $g^4/(16\pi^2)$ , the charm contribution carries a mass suppression  $\lambda m_c^2/M_W^2$ , while the top term is proportional to  $\lambda^5 A^2(1 - \rho + i\eta) m_t^2/M_W^2$ . The large top mass compensates the strong Cabibbo suppression so that the top contribution is finally larger than the charm one. However, the total short-distance contribution to the  $K_L$  decay,  $\text{Br}(K_L \rightarrow \mu^+ \mu^-)_{\text{sd}} = (0.79 \pm 0.12) \cdot 10^{-9}$  [119, 135], is nearly one order of magnitude below the experimental measurement (93), while  $\text{Br}(K_S \rightarrow \mu^+ \mu^-)_{\text{sd}} \approx 1.7 \cdot 10^{-13}$  [39].

The decays  $K_L \rightarrow \mu^+ \mu^-$  and  $K_S \rightarrow \mu^+ \mu^-$  are actually dominated by long-distance contributions, through  $K_L \rightarrow \pi^0, \eta, \eta' \rightarrow \gamma\gamma \rightarrow \mu^+ \mu^-$  and  $K_S \rightarrow \pi^+ \pi^- \rightarrow \gamma\gamma \rightarrow \mu^+ \mu^-$ . The absorptive component from two on-shell intermediate photons nearly saturates the measured  $\text{Br}(K_L \rightarrow \mu^+ \mu^-)$  [136], while the  $K_S$  decay rate is predicted to be  $\text{Br}(K_S \rightarrow \mu^+ \mu^-) = 5.1 \cdot 10^{-12}$  [39, 137]. These decays can be rigorously analysed with chiral perturbation theory techniques [39], but the strong suppression of their short-distance contributions does not make possible to extract useful information on the CKM parameters. Nevertheless, it is possible to predict the longitudinal polarization  $P_L$  of either muon in the  $K_L$  decay, a  $\mathcal{CP}$ -violating observable which in the SM is dominated by indirect  $\mathcal{CP}$  violation from  $K^0 - \bar{K}^0$  mixing:  $|P_L| = (2.6 \pm 0.4) \cdot 10^{-3}$  [137].

Other interesting kaon decay modes such as  $K^0 \rightarrow \gamma\gamma$ ,  $K \rightarrow \pi\gamma\gamma$  and  $K \rightarrow \pi\ell^+\ell^-$  are also governed by long-distance physics [39, 138]. Of particular interest is the decay  $K_L \rightarrow \pi^0 e^+ e^-$ , since it is sensitive to new sources of CP violation. At lowest order in  $\alpha$  the decay proceeds through  $K_2^0 \rightarrow \pi^0 \gamma^*$  that violates  $\mathcal{CP}$ , while the  $\mathcal{CP}$ -conserving contribution through  $K_L^0 \rightarrow \pi^0 \gamma^* \gamma^*$  is suppressed by an additional power of  $\alpha$  and it is found to be below  $10^{-12}$  [39]. The  $K_L \rightarrow \pi^0 e^+ e^-$  transition is then dominated by the  $\mathcal{O}(\alpha)$  CP-violating contributions [138], both from  $K^0 - \bar{K}^0$  mixing and direct  $\mathcal{CP}$  violation. The estimated rate  $\text{Br}(K_L \rightarrow \pi^0 e^+ e^-) = (3.1 \pm 0.9) \cdot 10^{-11}$  [39, 139, 140] is only a factor ten smaller than the present (90% C.L.) upper bound of  $2.8 \cdot 10^{-10}$  [141] and should be reachable in the near future.

The decays  $K^\pm \rightarrow \pi^\pm \nu \bar{\nu}$  and  $K_L \rightarrow \pi^0 \nu \bar{\nu}$  provide a more direct access to CKM information because long-distance effects play a negligible role. The decay amplitudes are dominated by  $Z$ -penguin and  $W$ -box loop diagrams of the type shown in Fig. 10 (replacing the final muons by neutrinos), and are proportional to the hadronic  $K\pi$  matrix element of the  $\Delta S = 1$  vector current, which (assuming isospin symmetry) is extracted from  $K \rightarrow \pi \ell \nu_\ell$  decays. The small long-distance and isospin-violating corrections can be estimated with chiral perturbation theory. The neutral decay is  $\mathcal{CP}$  violating and proceeds almost entirely through direct  $\mathcal{CP}$  violation (via interference with mixing). Taking the CKM inputs from other observables, the predicted SM rates are [142–144]:

$$\text{Br}(K^+ \rightarrow \pi^+ \nu \bar{\nu})_{\text{th}} = (9.1 \pm 0.7) \cdot 10^{-11}, \quad \text{Br}(K_L \rightarrow \pi^0 \nu \bar{\nu})_{\text{th}} = (3.0 \pm 0.3) \cdot 10^{-11}. \quad (95)$$

The uncertainties are largely parametrical, due to CKM input, the charm and top masses and  $\alpha_s(M_Z)$ . On the experimental side, the charged kaon mode was already observed [145], while only an upper bound on the neutral mode has been achieved [146]:

$$\text{Br}(K^+ \rightarrow \pi^+ \nu \bar{\nu}) = (1.73_{-1.05}^{+1.15}) \cdot 10^{-10}, \quad \text{Br}(K_L \rightarrow \pi^0 \nu \bar{\nu}) < 2.6 \cdot 10^{-8} \quad (90\% \text{ C.L.}). \quad (96)$$

The CERN NA62 experiment is currently collecting  $K^+ \rightarrow \pi^+ \nu \bar{\nu}$  data with the goal of measuring the branching ratio with an accuracy of 10% [147]. The neutral mode is being searched for by the KOTO experiment at J-PARC [148], which aims to reach the SM predicted branching ratio. These experiments will start to seriously probe the new-physics potential of these decays.

The inclusive decay  $\bar{B} \rightarrow X_s \gamma$  provides another powerful test of the SM flavour structure at the quantum loop level. It proceeds through a  $b \rightarrow s \gamma$  penguin diagram with an on-shell photon. The present experimental world average,  $\text{Br}(\bar{B} \rightarrow X_s \gamma)_{E_\gamma \geq 1.6 \text{ GeV}} = (3.32 \pm 0.15) \cdot 10^{-4}$  [43], agrees very well with the SM theoretical prediction,  $\text{Br}(\bar{B} \rightarrow X_s \gamma)_{E_\gamma \geq 1.6 \text{ GeV}}^{\text{th}} = (3.36 \pm 0.23) \cdot 10^{-4}$  [149, 150], obtained at the next-to-next-to-leading order.

Evidence for the  $B_s^0 \rightarrow \mu^+ \mu^-$  process has been reported by the CMS and LHCb collaborations [151], and LHCb has subsequently observed this decay mode with a  $7.8 \sigma$  significance [152]:<sup>6</sup>

$$\overline{\text{Br}}(B_s^0 \rightarrow \mu^+ \mu^-) = (3.0 \pm 0.6^{+0.3}_{-0.2}) \cdot 10^{-9}, \quad \overline{\text{Br}}(B_d^0 \rightarrow \mu^+ \mu^-) < 3.4 \cdot 10^{-10} \quad (95\% \text{ CL}). \quad (97)$$

These experimental results agree with the SM predictions  $\overline{\text{Br}}(B_s^0 \rightarrow \mu^+ \mu^-) = (3.65 \pm 0.23) \cdot 10^{-9}$  and  $\overline{\text{Br}}(B_d^0 \rightarrow \mu^+ \mu^-) = (1.06 \pm 0.09) \cdot 10^{-10}$  [154]. Other interesting decays with  $B$  mesons are  $\bar{B} \rightarrow K^{(*)} l^+ l^-$  and  $\bar{B} \rightarrow K^{(*)} \nu \bar{\nu}$  [83].

## 8 Flavour constraints on new physics

The CKM matrix provides a very successful description of flavour phenomena, as it is clearly exhibited in the unitarity test of Fig. 6, showing how very different observables converge into a single choice of CKM parameters. This is a quite robust and impressive result. One can perform separate tests with different subsets of measurements, according to their CP-conserving or CP-violating nature, or splitting them into tree-level and loop-induced processes. In all cases, one finally gets a closed triangle and similar values for the fitted CKM entries [108, 130]. However, the SM mechanism of flavour mixing and  $\mathcal{CP}$  violation is conceptually quite unsatisfactory because it does not provide any dynamical understanding of the numerical values of fermion masses, and mixings. We completely ignore the reasons why the fermion spectrum contains such a hierarchy of different masses, spanning many orders of magnitude, and which fundamental dynamics is behind the existence of three flavour generations and their observed mixing structure.

The phenomenological success of the SM puts severe constraints on possible scenarios of new physics. The absence of any clear signals of new phenomena in the LHC searches is pushing the hypothetical new-physics scale at higher energies, above the TeV. The low-energy implications of new dynamics beyond the SM can then be analysed in terms of an effective Lagrangian, containing only the known SM fields:

$$\mathcal{L}_{\text{eff}} = L_{\text{SM}} + \sum_{d>4} \sum_k \frac{c_k^{(d)}}{\Lambda_{\text{NP}}^{d-4}} O_k^{(d)}. \quad (98)$$

The effective Lagrangian is organised as an expansion in terms of dimension- $d$  operators  $O_k^{(d)}$ , invariant under the SM gauge group, suppressed by the corresponding powers of the new-physics scale  $\Lambda_{\text{NP}}$ . The dimensionless couplings  $c_k^{(d)}$  encode information on the underlying dynamics. The lowest-order term in this dimensional expansion is the SM Lagrangian that contains all allowed operators of dimension 4.

At low energies, the terms with lower dimensions dominate the physical transition amplitudes. There is only one operator with  $d = 5$  (up to Hermitian conjugation and flavour assignments), which violates lepton number by two units and is then related with the possible existence of neutrino Majorana masses [155]. Taking  $m_\nu \gtrsim 0.05 \text{ eV}$ , one estimates a very large lepton-number-violating scale

<sup>6</sup> $\overline{\text{Br}}$  is the time-integrated branching ratio, which for  $B_s^0$  is slightly affected by the sizeable value of  $\Delta\Gamma_s$  [153].

$\Lambda_{\text{NP}}/c^{(5)} \lesssim 10^{15}$  GeV [17]. Assuming lepton-number conservation, the first signals of new phenomena should be associated with  $d = 6$  operators.

One can easily analyse the possible impact of  $\Delta F = 2$  ( $F = S, C, B$ ) four-quark operators ( $d = 6$ ), such as the SM left-left operator in Eq. (63) that induces a  $\Delta B = 2$  transition. Since the SM box diagram provides an excellent description of the data, hypothetical new-physics contributions can only be tolerated within the current uncertainties, which puts stringent upper bounds on the corresponding Wilson coefficients  $\tilde{c}_k \equiv c_k^{(6)}/\Lambda_{\text{NP}}^2$ . For instance,  $\Delta M_{B_d^0}$  and  $S_{J/\psi K_S}$  imply that the real (CP-conserving) and imaginary (CP-violating) parts of  $\tilde{c}_k$  must be below  $10^{-6}$   $\text{TeV}^{-2}$  for a  $(\bar{b}_L \gamma^\mu d_L)^2$  operator, and nearly one order of magnitude smaller ( $10^{-7}$ ) for  $(\bar{b}_R d_L)(\bar{b}_L d_R)$  [156]. Stronger bounds are obtained in the kaon system from  $\Delta M_{K^0}$  and  $\varepsilon_K$ . For the  $(\bar{s}_L \gamma^\mu d_L)^2$  operator the real (imaginary) coefficient must be below  $10^{-6}$  ( $3 \cdot 10^{-9}$ ), while the corresponding bounds for  $(\bar{s}_R d_L)(\bar{s}_L d_R)$  are  $7 \cdot 10^{-9}$  ( $3 \cdot 10^{-11}$ ), in the same  $\text{TeV}^{-2}$  units [156]. Taking the coefficients  $c_k^{(6)} \sim \mathcal{O}(1)$ , this would imply  $\Lambda_{\text{NP}} > 3 \cdot 10^3$  TeV for  $B_d^0$  and  $\Lambda_{\text{NP}} > 3 \cdot 10^5$  TeV for  $K^0$ . Therefore, two relevant messages emerge from the data:

1. A generic flavour structure with coefficients  $c_k^{(6)} \sim \mathcal{O}(1)$  is completely ruled out at the TeV scale.
2. New flavour-changing physics at  $\Lambda_{\text{NP}} \sim 1$  TeV could only be possible if the corresponding Wilson coefficients  $c_k^{(6)}$  inherit the strong SM suppressions generated by the GIM mechanism.

The last requirement can be satisfied by assuming that the up and down Yukawa matrices are the only sources of quark-flavour symmetry breaking (minimal flavour violation) [157–159]. In the absence of Yukawa interactions, the SM Lagrangian has a  $\mathcal{G} \equiv U(3)_{L_L} \otimes U(3)_{Q_L} \otimes U(3)_{\ell_R} \otimes U(3)_{u_R} \otimes U(3)_{d_R}$  global flavour symmetry, because one can rotate arbitrarily in the 3-generation space each one of the five SM fermion components in Eq. (2). The Yukawa matrices are the only explicit breakings of this large symmetry group. Assuming that the new physics does not introduce any additional breakings of the flavour symmetry  $\mathcal{G}$  (beyond insertions of Yukawa matrices), one can easily comply with the flavour bounds. Otherwise, flavour data provide very strong constraints on models with additional sources of flavour symmetry breaking and probe physics at energy scales not directly accessible at accelerators.

The subtle SM cancellations suppressing FCNC transitions would be easily destroyed in the presence of new physics contributions. To better appreciate the non-generic nature of the flavour structure, let us analyse the simplest extension of the SM scalar sector with a second Higgs doublet, which increases the number of quark Yukawas:

$$\mathcal{L}_Y = - \sum_{a=1}^2 \left\{ \bar{Q}'_L \mathcal{Y}_d^{(a)'} \phi_a d'_R + \bar{Q}'_L \mathcal{Y}_u^{(a)'} \phi_a^c u'_R + \bar{L}'_L \mathcal{Y}_\ell^{(a)'} \phi_a \ell'_R \right\} + \text{h.c.}, \quad (99)$$

where  $\phi_a^T = (\phi_a^{(+)}, \phi_a^{(0)})$  are the two scalar doublets,  $\phi_a^c$  their  $\mathcal{C}$ -conjugate fields, and  $Q'_L$  and  $L'_L$  the left-handed quark and lepton doublets, respectively. All fermion fields are written as three-dimensional flavour vectors and  $\mathcal{Y}_f^{(a)'}$  are  $3 \times 3$  complex matrices in flavour space. With an appropriate scalar potential, the neutral components of the scalar doublets acquire vacuum expectation values  $\langle 0 | \phi_a^{(0)} | 0 \rangle = v_a e^{i\theta_a}$ . It is convenient to make a  $U(2)$  transformation in the space of scalar fields,  $(\phi_1, \phi_2) \rightarrow (\Phi_1, \Phi_2)$ , so that only the first doublet has a non-zero vacuum expectation value  $v = (v_1^2 + v_2^2)^{1/2}$ .  $\Phi_1$  plays then the same role as the SM Higgs doublet, while  $\Phi_2$  does not participate in the electroweak symmetry breaking. In this scalar basis the Yukawa interactions become more transparent:

$$\sum_{a=1}^2 \mathcal{Y}_{d,\ell}^{(a)'} \phi_a = \sum_{a=1}^2 Y_{d,\ell}^{(a)'} \Phi_a, \quad \sum_{a=1}^2 \mathcal{Y}_u^{(a)'} \phi_a^c = \sum_{a=1}^2 Y_u^{(a)'} \Phi_a^c. \quad (100)$$

The fermion masses originate from the  $\Phi_1$  couplings, because  $\Phi_1$  is the only field acquiring a vacuum expectation value:

$$M'_f = Y_f^{(1)'} \frac{v}{\sqrt{2}}. \quad (101)$$

The diagonalization of these fermion mass matrices proceeds in exactly the same way as in the SM, and defines the fermion mass eigenstates  $d_i$ ,  $u_i$ ,  $\ell_i$ , with diagonal mass matrices  $M_f$ , as described in Section 2. However, in general, one cannot diagonalize simultaneously all Yukawa matrices, i.e., in the fermion mass-eigenstate basis the matrices  $Y_f^{(2)}$  remain non-diagonal, giving rise to dangerous flavour-changing transitions mediated by neutral scalars. The appearance of FCNC interactions represents a major phenomenological shortcoming, given the very strong experimental bounds on this type of phenomena.

To avoid this disaster, one needs to implement ad-hoc dynamical restrictions to guarantee the suppression of FCNC couplings at the required level. Unless the Yukawa couplings are very small or the scalar bosons very heavy, a specific flavour structure is required by the data. The unwanted non-diagonal neutral couplings can be eliminated requiring the alignment in flavour space of the Yukawa matrices [160]:

$$Y_{d,\ell}^{(2)} = \varsigma_{d,\ell} Y_{d,\ell}^{(1)} = \frac{\sqrt{2}}{v} \varsigma_{d,\ell} M_{d,\ell}, \quad Y_u^{(2)} = \varsigma_u^* Y_u^{(1)} = \frac{\sqrt{2}}{v} \varsigma_u^* M_u, \quad (102)$$

with  $\varsigma_f$  arbitrary complex proportionality parameters.<sup>7</sup>

Flavour alignment constitutes a very simple implementation of minimal flavour violation. It results in a very specific dynamical structure, with all fermion-scalar interactions being proportional to the corresponding fermion masses. The Yukawas are fully characterized by the three complex alignment parameters  $\varsigma_f$ , which introduce new sources of CP violation. The aligned two-Higgs doublet model Lagrangian satisfies the flavour constraints [162–172], and leads to a rich collider phenomenology with five physical scalar bosons [173–180]:  $h$ ,  $H$ ,  $A$  and  $H^\pm$ .

## 9 Flavour anomalies

The experimental data exhibit a few deviations from the SM predictions. For instance, Table 2 shows a  $2.6\sigma$  ( $2.4\sigma$ ) violation of lepton universality in  $|g_\tau/g_\mu|$  ( $|g_\tau/g_e|$ ) at the 1% level, from  $W \rightarrow \ell\nu$  decays, that is difficult to reconcile with the precise 0.15% limits extracted from virtual  $W^\pm$  transitions, shown also in the same table. In fact, it has been demonstrated that it is not possible to accommodate this deviation from universality with an effective Lagrangian and, therefore, such a signal could only be explained by the introduction of new light degrees of freedom that so far remain undetected [181]. Thus, the most plausible explanation is a small problem (statistical fluctuation or underestimated systematics) in the LEP-2 measurements that will remain unresolved until more precise high-statistics  $W \rightarrow \ell\nu$  data samples become available.

Some years ago BaBar reported a non-zero  $\mathcal{CP}$  asymmetry in  $\tau^\pm \rightarrow \pi^\pm K_S \nu$  decays at the level of  $3.6 \cdot 10^{-3}$  [182], the same size than the SM expectation from  $K^0-\bar{K}^0$  mixing [183, 184] but with the opposite sign, which represents a  $2.8\sigma$  anomaly. So far, Belle has only reached a null result with a smaller  $10^{-2}$  sensitivity and, therefore, has not been able to either confirm or refute the asymmetry. Nevertheless, on very general grounds, it has been shown that the BaBar signal is incompatible with other sets of data ( $K^0$  and  $D^0$  mixing, neutron electric dipole moment) [185].

Another small flavour anomaly was triggered by the unexpected large value of  $\text{Br}(B^- \rightarrow \tau^- \bar{\nu})$ , found in 2006 by Belle [186] and later confirmed by BaBar [93], which implied values of  $|\mathbf{V}_{ub}|$  higher than the ones measured in semileptonic decays or extracted from global CKM fits. While the BaBar results remain unchanged, the reanalysis of the full Belle data sample resulted in a sizeable  $\sim 40\%$

<sup>7</sup>Actually, since one only needs that  $Y_f^{(1)'}$  and  $Y_f^{(2)'}$  can be simultaneously diagonalized, in full generality the factors  $\varsigma_f$  could be 3-dimensional diagonal matrices in the fermion mass-eigenstate basis (generalized alignment) [161]. The fashionable models of types I, II, X and Y, usually considered in the literature, are particular cases of the flavour-aligned Lagrangian with all  $\varsigma_f$  parameters real and fixed in terms of  $\tan\beta = v_2/v_1$  [160].

reduction of the measured central value [94], eliminating the discrepancy with the SM but leaving a pending disagreement with the BaBar results.

In the last few years a series of anomalies have emerged in  $b \rightarrow c\tau\nu$  and  $b \rightarrow s\mu^+\mu^-$  transitions. Given the difficulty of the experimental analyses, the results should be taken with some caution and further studies with larger data sets are still necessary. Nevertheless, these anomalies exhibit a quite consistent pattern that makes them intriguing.

### 9.1 $b \rightarrow c\tau\nu$ decays

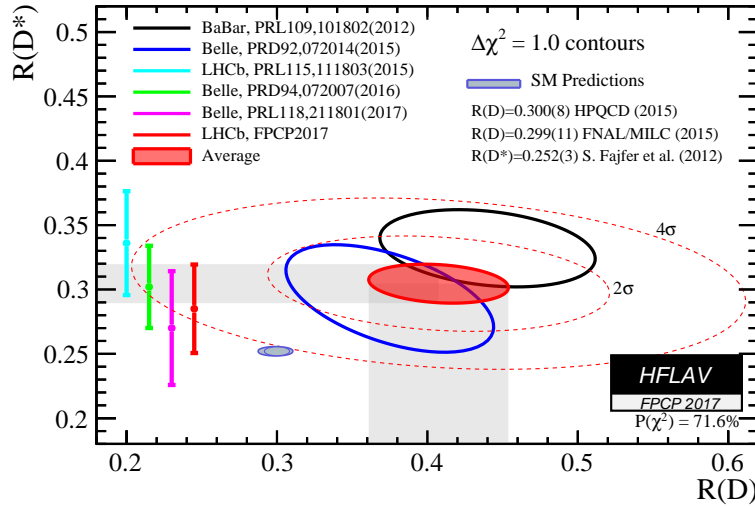
In 2012 the BaBar collaboration [187] observed an excess in  $B \rightarrow D^{(*)}\tau\nu_\tau$  decays with respect to the SM predictions [188], indicating a violation of lepton-flavour universality at the 30% level. The measured observables are the ratios

$$R(D) \equiv \frac{\text{Br}(B \rightarrow D\tau\nu)}{\text{Br}(B \rightarrow D\ell\nu)}, \quad R(D^*) \equiv \frac{\text{Br}(B \rightarrow D^*\tau\nu)}{\text{Br}(B \rightarrow D^*\ell\nu)}, \quad (103)$$

with  $\ell = e, \mu$ , where many sources of experimental and theoretical errors cancel. The effect has been later confirmed by LHCb [189] ( $D^*$  mode only) and Belle [190] (Fig. 11). Although the results of the last two experiments are slightly closer to the SM expected values,  $R(D) = 0.300 \pm 0.008$  [36, 53, 191] and  $R(D^*) = 0.252 \pm 0.003$  [188], the resulting world averages [43]

$$R(D) = 0.407 \pm 0.039 \pm 0.024, \quad R(D^*) = 0.304 \pm 0.013 \pm 0.007, \quad (104)$$

exceed the SM predictions by  $2.3\sigma$  and  $3.4\sigma$ , respectively. Combining the two measurements, with a correlation of  $-20\%$ , this implies a  $4.1\sigma$  deviation from the SM [43], which is a very large effect for a tree-level SM transition.



**Fig. 11:** Measurements of  $R(D)$  and  $R(D^*)$  and their average compared with the SM predictions [36, 53, 188, 191]. The continuous (dashed) ellipses display the  $1\sigma$  ( $2\sigma$  and  $4\sigma$ ) regions [43].

A worrisome aspect to keep in mind is that the latest  $R(D^*)$  measurements of both LHCb and Belle, with the  $\tau$  reconstructed with hadronic instead of leptonic decay modes, exhibit a shift to smaller values, approaching the SM prediction. On the theoretical side, the recent re-evaluation of form-factor extrapolations related to the exclusive  $|\mathbf{V}_{cb}|$  determination favours a slightly larger prediction  $R(D^*) = 0.260 \pm 0.008$  [55], with an enlarged uncertainty, while a smaller error is assigned to the other ratio,  $R(D) = 0.299 \pm 0.003$  [54].



Different new-physics explanations of the anomaly have been put forward.<sup>8</sup> A natural candidate for an additional tree-level contribution to these decays would be a charged scalar boson, coupling to fermions proportionally to their masses, as predicted in flavour-aligned two-Higgs-doublet models [160, 162]. This could easily increase the value of  $R(D)$ , since the scalar form factor contribution to the  $\tau$  mode is already sizeable in the SM. However, to accommodate a large increase of  $R(D^*)$ , which is sensitive to several vector and axial-vector form factors, requires too big scalar contributions that are in tension with the measured  $q^2$  distribution. The usual two-Higgs doublet models (types I, II, X and Y) cannot be made compatible with the data, but a (generalized) flavour-aligned scenario with the simultaneous presence of left- and right-handed couplings to the quarks improves considerably the global fit [192]. Taking into account the known constraints from the  $B_c$  decay width and the inclusive  $b \rightarrow X_c \tau \nu$  branching ratio, an explanation of the anomaly in terms of an intermediate charged scalar would require values of  $R(D^*)$   $1\sigma$  smaller than the present experimental central value. The scalar contribution could be disentangled through measurements of the differential distributions and the polarization of the final  $\tau$  and/or  $D^*$  [192].

A different possibility is the existence of some new-physics contribution that only manifests in the Wilson coefficient of the SM operator  $(\bar{c}_L \gamma^\mu b_L)(\bar{\tau}_L \gamma_\mu \nu_L)$  [193–195]. This would imply a universal enhancement of all  $b \rightarrow c \tau \nu$  transitions,

$$\frac{R(D)}{R(D)_{\text{SM}}} = \frac{R(D^*)}{R(D^*)_{\text{SM}}} = \frac{R(J/\psi)}{R(J/\psi)_{\text{SM}}}, \quad (105)$$

in agreement with the experimental ratios  $R(D)$  and  $R(D^*)$ , and the recent LHCb observation of the  $B_c \rightarrow J/\psi \tau \nu$  decay [196],

$$R(J/\psi) \equiv \frac{\text{Br}(B_c \rightarrow J/\psi \tau \nu)}{\text{Br}(B_c \rightarrow J/\psi \ell \nu)} = 0.71 \pm 0.17 \pm 0.18, \quad (106)$$

which is  $2\sigma$  above the SM expected value  $R(J/\psi) \sim 0.25\text{--}0.28$  [197–200]. Writing the four-fermion left-left operator in terms of  $SU(2)_L \otimes U(1)_Y$  invariant operators at the electroweak scale, and imposing that the experimental bounds on  $\text{Br}(b \rightarrow s \nu \bar{\nu})$  are satisfied, this possibility would imply rather large rates in  $b \rightarrow s \tau^+ \tau^-$  transitions [193–195], but still safely below the current upper limits [201].

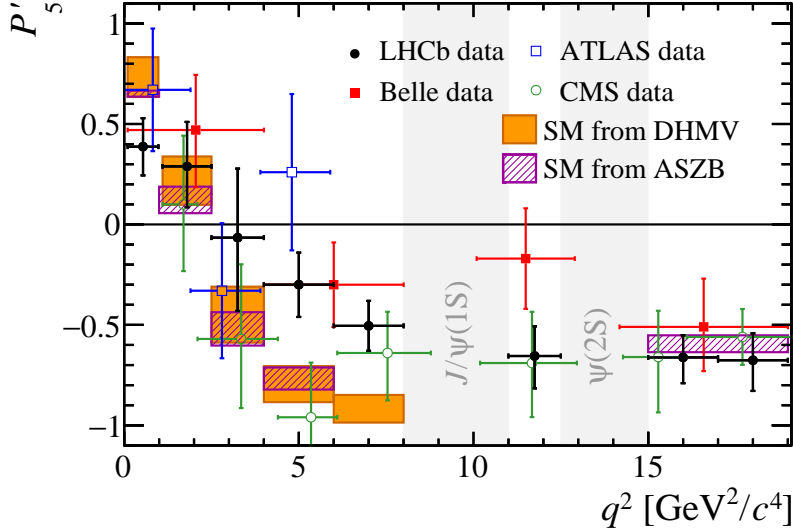
## 9.2 $b \rightarrow s \ell^+ \ell^-$ decays

The rates of several  $b \rightarrow s \mu^+ \mu^-$  transitions have been found at LHCb to be consistently lower than their SM predictions:  $B_s^0 \rightarrow \phi \mu^+ \mu^-$  [202],  $B_d^0 \rightarrow K^0 \mu^+ \mu^-$  [203],  $B^+ \rightarrow K^+ \mu^+ \mu^-$  [203, 204],  $B^+ \rightarrow K^{*+} \mu^+ \mu^-$  [203],  $B_d^0 \rightarrow K^{*0} \mu^+ \mu^-$  [205] and  $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$  [206]. The angular and invariant-mass distributions of the final products in  $B \rightarrow K^* \mu^+ \mu^-$  have been also studied by ATLAS [207], BaBar [208], Belle [209], CDF [210], CMS [211] and LHCb [205]. The rich variety of angular dependences in the four-body  $K \pi \mu^+ \mu^-$  final state allows one to disentangle different sources of dynamical contributions. Particular attention has been put in the so-called *optimised observables*  $P'_i(q^2)$ , where  $q^2$  is the dilepton invariant-mass squared, which are specific combinations of angular observables that are free from form-factor uncertainties in the heavy quark-mass limit [212]. A sizeable discrepancy with the SM prediction [213–215], shown in Fig. 12, has been identified in two adjacent bins of the  $P'_5$  distribution, at large  $K^*$  recoil [205, 207, 209, 211]. Belle has also included  $K^* e^+ e^-$  final states in the analysis, but the results for this electron mode are compatible with the SM expectations [209]. The hadronic uncertainties of the SM predictions are not easy to quantify, but have been scrutinised by several groups [216–221].

LHCb has also reported violations of lepton flavour universality, through the ratios [222]

$$R_K \equiv \left. \frac{\text{Br}(B^+ \rightarrow K^+ \mu^+ \mu^-)}{\text{Br}(B^+ \rightarrow K^+ e^+ e^-)} \right|_{q^2 \in [1, 6] \text{ GeV}^2} = 0.745^{+0.090}_{-0.074} \pm 0.036, \quad (107)$$

<sup>8</sup>A long, but not exhaustive, list of relevant references is given in Refs. [192].



**Fig. 12:** Comparison between the SM predictions for  $P'_5$  [213–215] and the experimental measurements [223].

and

$$R_{K^{*0}} \equiv \frac{\text{Br}(B_d^0 \rightarrow K^{*0} \mu^+ \mu^-)}{\text{Br}(B_d^0 \rightarrow K^{*0} J/\psi(\rightarrow \mu^+ \mu^-))} \bigg/ \frac{\text{Br}(B_d^0 \rightarrow K^{*0} e^+ e^-)}{\text{Br}(B_d^0 \rightarrow K^{*0} J/\psi(\rightarrow e^+ e^-))} \quad (108)$$

in two different  $q^2$  bins [224],

$$R_{K^{*0}} \big|_{q^2 \in [0.045, 1.1] \text{ GeV}^2} = 0.66_{-0.07}^{+0.11} \pm 0.03, \quad R_{K^{*0}} \big|_{q^2 \in [1.1, 6] \text{ GeV}^2} = 0.69_{-0.07}^{+0.11} \pm 0.05. \quad (109)$$

These observables constitute a much cleaner probe of new physics because most theoretical uncertainties cancel out [225–228]. In the SM, the only difference between the muon and electron channels is the lepton mass. The SM theoretical predictions,  $R_K = 1.00 \pm 0.01_{\text{QED}}$ ,  $R_{K^{*0}}[0.045, 1.1] = 0.906 \pm 0.028_{\text{th}}$  and  $R_{K^{*0}}[1.1, 6] = 1.00 \pm 0.01_{\text{QED}}$  [229], deviate from the experimental measurements by  $2.6 \sigma$ ,  $2.1 \sigma$  and  $2.6 \sigma$ , respectively.

Global fits to the  $b \rightarrow s \ell^+ \ell^-$  data with an effective low-energy Lagrangian

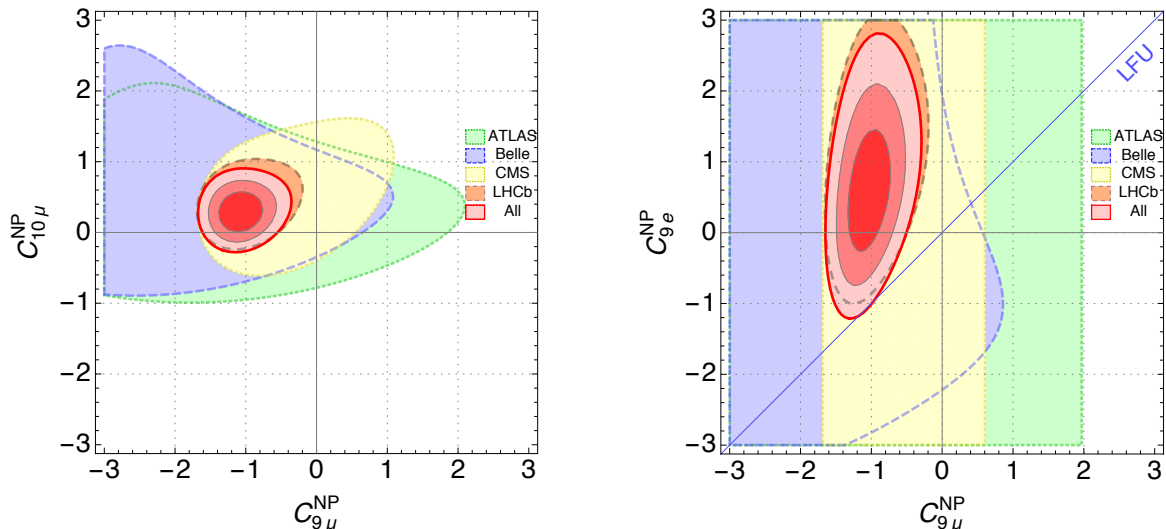
$$\mathcal{L}_{\text{eff}} = \frac{G_F}{\sqrt{2}} \mathbf{V}_{tb} \mathbf{V}_{ts}^* \frac{\alpha}{\pi} \sum_{i,\ell} C_{i,\ell} O_i^\ell \quad (110)$$

show a clear preference for new-physics contributions to the operators  $O_9^\ell = (\bar{s}_L \gamma_\mu b_L)(\bar{\ell} \gamma^\mu \ell)$  and  $O_{10}^\ell = (\bar{s}_L \gamma_\mu b_L)(\bar{\ell} \gamma^\mu \gamma_5 \ell)$ , with  $\ell = \mu$  [230–234]. Fig. 13 exhibits the need for  $C_{9,\mu}^{\text{NP}} \neq 0$  and rather large, of the order of  $\sim -25\%$  of the SM contribution to  $C_{9,\mu}$  [230]. The statistical significance for non-zero values of  $C_{10,\mu}^{\text{NP}}$ ,  $C_{9,e}^{\text{NP}}$  and  $C_{10,e}^{\text{NP}}$  is low, although a reasonable fit to the data could be obtained with  $C_{9,\mu}^{\text{NP}} - C_{9,e}^{\text{NP}} - C_{10,\mu}^{\text{NP}} + C_{10,e}^{\text{NP}} \sim -1.4$  [232]. Many attempts to explain these results within specific scenarios of new physics have been put forward (new gauge bosons, leptoquarks, new fermions and/or scalars, etc.). A representative list of references can be found in Refs. [235–237].

## 10 Discussion

The flavour structure of the SM is one of the main pending questions in our understanding of weak interactions. Although we do not know the reasons of the observed family replication, we have learnt experimentally that the number of SM generations is just three (and no more). Therefore, we must study as precisely as possible the few existing flavours, to get some hints on the dynamics responsible for their observed structure.





**Fig. 13:** Allowed 1, 2 and 3  $\sigma$  regions in the  $(C_{9,\mu}^{\text{NP}}, C_{10,\mu}^{\text{NP}})$  and  $(C_{9,\mu}^{\text{NP}}, C_{9,e}^{\text{NP}})$  planes, from a global fit to  $b \rightarrow s\ell^+\ell^-$  data [230]. The contours from specific experiments are show at 3  $\sigma$ .

In the SM all flavour dynamics originate in the fermion mass matrices, which determine the measurable masses and mixings. Thus, flavour is related through the Yukawa interactions with the scalar sector, the part of the SM Lagrangian that is more open to theoretical speculations. At present, we totally ignore the underlying dynamics responsible for the vastly different scales exhibited by the fermion spectrum and the particular values of the measured mixings. The SM Yukawa matrices are just a bunch of arbitrary parameters to be fitted to data, which is conceptually unsatisfactory.

The SM incorporates a mechanism to generate  $\mathcal{CP}$  violation, through the single phase naturally occurring in the CKM matrix. This mechanism, deeply rooted into the unitarity structure of  $\mathbf{V}$ , implies very specific requirements for  $\mathcal{CP}$  violation to show up. The CKM matrix has been thoroughly investigated in dedicated experiments and a large number of  $\mathcal{CP}$ -violating processes have been studied in detail. The flavour data seem to fit into the SM framework, confirming that the fermion mass matrices are the dominant source of flavour-mixing phenomena. However, a fundamental explanation of the flavour dynamics is still lacking.

At present, a few flavour anomalies have been identified in  $b \rightarrow c\tau\nu$  and  $b \rightarrow s\mu^+\mu^-$  transitions. Whether they truly represent the first signals of new phenomena, or just result from statistical fluctuations and/or underestimated systematics remains to be understood. New experimental input from LHC and Belle-II should soon clarify the situation. Very valuable information on the flavour dynamics is also expected from BESS-III and from several kaon (NA62, KOTO) and muon (MEG-II, Mu2e, COMET, PRISM/PRIME) experiments, complementing the high-energy searches for new phenomena at LHC. Unexpected surprises may well be discovered, probably giving hints of new physics at higher scales and offering clues to the problems of fermion mass generation, quark mixing and family replication.

## Acknowledgements

I want to thank the organizers for the charming atmosphere of this school and all the students for their many interesting questions and comments. This work has been supported in part by the Spanish Government and ERDF funds from the EU Commission [Grants FPA2017-84445-P and FPA2014-53631-C2-1-P], by Generalitat Valenciana [Grant Prometeo/2017/053] and by the Spanish Centro de Excelencia Severo Ochoa Programme [Grant SEV-2014-0398].

## Appendices

### A Conservation of the vector current

In the limit where all quark masses are equal, the QCD Lagrangian remains invariant under global  $SU(N_f)$  transformations of the quark fields in flavour space, where  $N_f$  is the number of (equal-mass) quark flavours. This guarantees the conservation of the corresponding Noether currents  $V_{ij}^\mu = \bar{u}_i \gamma^\mu d_j$ . In fact, using the QCD equations of motion, one immediately finds that

$$\partial_\mu V_{ij}^\mu \equiv \partial_\mu (\bar{u}_i \gamma^\mu d_j) = i (m_{u_i} - m_{d_j}) \bar{u}_i d_j, \quad (\text{A.1})$$

which vanishes when  $m_{u_i} = m_{d_j}$ . In momentum space, this reads  $q_\mu V_{ij}^\mu = \mathcal{O}(m_{u_i} - m_{d_j})$ , with  $q_\mu$  the corresponding momentum transfer.

Let us consider a  $0^-(k) \rightarrow 0^-(k')$  weak transition mediated by the vector current  $V_{ij}^\mu$ . The relevant hadronic matrix element is given in Eq. (19) and contains two form factors  $f_\pm(q^2)$ . The conservation of the vector current implies that  $f_-(q^2)$  is identically zero when  $m_{u_i} = m_{d_j}$ . Therefore,

$$\langle P'_i(k') | V_{ij}^\mu(x) | P_j(k) \rangle = e^{iq \cdot x} C_{PP'} (k + k')^\mu f_+(q^2). \quad (\text{A.2})$$

We have made use of translation invariance to write  $V_{ij}^\mu(x) = e^{iP \cdot x} V_{ij}^\mu(0) e^{-iP \cdot x}$ , with  $P^\mu$  the four-momentum operator. This determines the dependence of the matrix element on the space-time coordinate, with  $q^\mu = (k' - k)^\mu$ .

The conserved Noether charges

$$\mathcal{N}_{ij} = \int d^3x V_{ij}^0(x) = \int d^3x u_i^\dagger(x) d_j(x), \quad (\text{A.3})$$

annihilate one quark  $d_j$  and create instead one  $u_i$  (or replace  $\bar{u}_i$  by  $\bar{d}_j$ ), transforming the meson  $P$  into  $P'$  (up to a trivial Clebsh-Gordan factor  $C_{PP'}$  that, for light-quarks, takes the value  $1/\sqrt{2}$  when  $P'$  is a  $\pi^0$  and is 1 otherwise). Thus,

$$\langle P'(k') | \mathcal{N}_{ij} | P(k) \rangle = C_{PP'} \langle P'(k') | P(k) \rangle = C_{PP'} (2\pi)^3 2k^0 \delta^{(3)}(\vec{k}' - \vec{k}). \quad (\text{A.4})$$

On the other side, inserting in this matrix element the explicit expression for  $\mathcal{N}_{ij}$  in (A.3) and making use of (A.2),

$$\langle P'(k') | \mathcal{N}_{ij} | P(k) \rangle = \int d^3x \langle P'(k') | V_{ij}^0(x) | P(k) \rangle = C_{PP'} (2\pi)^3 2k^0 \delta^{(3)}(\vec{q}) f_+(0). \quad (\text{A.5})$$

Comparing Eqs. (A.4) and (A.5), one finally obtains the result

$$f_+(0) = 1. \quad (\text{A.6})$$

Therefore, the flavour symmetry  $SU(N_f)$  determines the normalization of the form factor at  $q^2 = 0$ , when  $m_{u_i} = m_{d_j}$ . It is possible to prove that the deviations from 1 are of second order in the symmetry-breaking quark mass difference, i.e.,  $f_+(0) = 1 + \mathcal{O}[(m_{u_i} - m_{d_j})^2]$  [22].

## References

- [1] S.L. Glashow, *Nucl. Phys.* **22** (1961) 579.
- [2] S. Weinberg, *Phys. Rev. Lett.* **19** (1967) 1264.
- [3] A. Salam, in *Elementary Particle Theory*, ed. N. Svartholm (Almqvist and Wiksells, Stockholm, 1969), p. 367.

- [4] A. Pich, *The Standard Model of Electroweak Interactions*, Reports CERN-2006-003 and CERN-2007-005, ed. R. Fleischer, p. 1; arXiv:hep-ph/0502010, arXiv:0705.4264 [hep-ph].
- [5] A. Pich, *Flavour Physics and CP Violation*, Proceedings 6th CERN – Latin-American School of High-Energy Physics (CLASHEP 2011, Natal, Brazil, March 33 - April 5, 2011), Report CERN-2013-003, p.119, ed. C. Grojean, M. Mulders and M. Spiropulu, arXiv:1112.4094 [hep-ph].
- [6] N. Cabibbo, *Phys. Rev. Lett.* **10** (1963) 531.
- [7] M. Kobayashi and T. Maskawa, *Prog. Theor. Phys.* **42** (1973) 652.
- [8] S.L. Glashow, J. Iliopoulos and L. Maiani, *Phys. Rev.* **D2** (1970) 1285.
- [9] C. Patrignani *et al.* [Particle Data Group], *Chin. Phys.* **C40** (2016) 100001; <http://pdg.lbl.gov>.
- [10] W.J. Marciano and A. Sirlin, *Phys. Rev. Lett.* **61** (1988) 1815.
- [11] T. van Ritbergen and R.G. Stuart, *Phys. Rev. Lett.* **82** (1999) 488; *Nucl. Phys.* **B564** (2000) 343.
- [12] A. Pak and A. Czarnecki, *Phys. Rev. Lett.* **100** (2008) 241807.
- [13] MuLan Collaboration, *Phys. Rev. Lett.* **106** (2011) 041803.
- [14] E. Braaten, S. Narison and A. Pich, *Nucl. Phys.* **B373** (1992) 581.
- [15] A. Pich and A. Rodríguez-Sánchez, *Phys. Rev.* **D94** (2016) 034027.
- [16] A. Pich, EPJ Web Conf. **137** (2017) 01016.
- [17] A. Pich, *Prog. Part. Nucl. Phys.* **75** (2014) 41.
- [18] A. Pich, *Flavourdynamics*, arXiv:hep-ph/9601202.
- [19] W.J. Marciano and A. Sirlin, *Phys. Rev. Lett.* **96** (2006) 032002; **71** (1993) 3629.
- [20] A. Czarnecki, W.J. Marciano and A. Sirlin, *Phys. Rev.* **D70** (2004) 093006.
- [21] J. C. Hardy and I. S. Towner, *Phys. Rev.* **C91** (2015) 025501.
- [22] M. Ademollo and R. Gatto, *Phys. Rev. Lett.* **13** (1964) 264.
- [23] V. Cirigliano, M. Knecht, H. Neufeld and H. Pichl, *Eur. Phys. J.* **C27** (2003) 255.
- [24] D. Poganec *et al.*, *Phys. Rev. Lett.* **93** (2004) 181803.
- [25] V. Cirigliano, M. Giannotti and H. Neufeld, *JHEP* **0811** (2008) 006.
- [26] A. Kastner and H. Neufeld, *Eur. Phys. J.* **C57** (2008) 541.
- [27] M. Antonelli *et al.* [FlaviaNet Working Group on Kaon Decays], *Eur. Phys. J.* **C69** (2010) 399.
- [28] M. Moulson, arXiv:1411.5252 [hep-ex].
- [29] R.E. Behrends and A. Sirlin, *Phys. Rev. Lett.* **4** (1960) 186.
- [30] H. Leutwyler and M. Roos, *Z. Phys.* **C25** (1984) 91.
- [31] A. Bazavov *et al.*, *Phys. Rev. Lett.* **112** (2014) 112001.
- [32] N. Carrasco *et al.*, *Phys. Rev.* **D93** (2016) 114512.
- [33] J. Bijnens and P. Talavera, *Nucl. Phys.* **B669** (2003) 341.
- [34] M. Jamin, J.A. Oller and A. Pich, *JHEP* **0402** (2004) 047.
- [35] V. Cirigliano *et al.*, *JHEP* **0504** (2005) 006.
- [36] S. Aoki *et al.* [FLAG Working Group], *Eur. Phys. J.* **C77** (2017) 112.
- [37] W.J. Marciano, *Phys. Rev. Lett.* **93** (2004) 231803.
- [38] V. Cirigliano and H. Neufeld, *Phys. Lett.* **B700** (2011) 7.
- [39] V. Cirigliano, G. Ecker, H. Neufeld, A. Pich and J. Portolés, *Rev. Mod. Phys.* **84** (2012) 399.
- [40] N. Cabibbo, E.C. Swallow and R. Winston, *Annu. Rev. Nucl. Part. Sci.* **53** (2003) 39; *Phys. Rev. Lett.* **92** (2004) 251803.
- [41] V. Mateu and A. Pich, *JHEP* **0510** (2005) 041.
- [42] E. Gámiz, M. Jamin, A. Pich, J. Prades and F. Schwab, *Phys. Rev. Lett.* **94** (2005) 011803; *JHEP* **0301** (2003) 060.

- [43] Y. Amhis *et al.* [HFLAV Collaboration], *Eur. Phys. J.* **C77** (2017) 895; <https://hflav.web.cern.ch/>.
- [44] M. Antonelli, V. Cirigliano, A. Lusiani and E. Passemar, *JHEP* **1310** (2013) 070.
- [45] N. Isgur and M. Wise, *Phys. Lett.* **B232** (1989) 113; **B237** (1990) 527.
- [46] B. Grinstein, *Nucl. Phys.* **B339** (1990) 253.
- [47] E. Eichten and B. Hill, *Phys. Lett.* **B234** (1990) 511.
- [48] H. Georgi, *Phys. Lett.* **B240** (1990) 447.
- [49] M. Neubert, *Phys. Lett.* **B264** (1991) 455.
- [50] M. Luke, *Phys. Lett.* **B252** (1990) 447.
- [51] I. Caprini, L. Lellouch and M. Neubert, *Nucl. Phys.* **B530** (1998) 153.
- [52] J. A. Bailey *et al.* [Fermilab Lattice and MILC Collaborations], *Phys. Rev.* **D89** (2014) 114504.
- [53] J. A. Bailey *et al.* [MILC Collaboration], *Phys. Rev.* **D92** (2015) 034506.
- [54] D. Bigi and P. Gambino, *Phys. Rev.* **D94** (2016) 094008.
- [55] D. Bigi, P. Gambino and S. Schacht, *Phys. Lett.* **B769** (2017) 441; *JHEP* **1711** (2017) 061.
- [56] B. Grinstein and A. Kobach, *Phys. Lett.* **B771** (2017) 359.
- [57] F. U. Bernlochner, Z. Ligeti, M. Papucci and D. J. Robinson, *Phys. Rev.* **D95** (2017) 115008 [Err: **D97** (2018) 059902]; **D96** (2017) 091503.
- [58] C. G. Boyd, B. Grinstein and R. F. Lebed, *Phys. Rev.* **D56** (1997) 6895.
- [59] Belle Collaboration, arXiv:1702.01521 [hep-ex].
- [60] BaBar Collaboration, *Phys. Rev. Lett.* **104** (2010) 011802.
- [61] Belle Collaboration, *Phys. Rev.* **D93** (2016) 032006.
- [62] I.I.Y. Bigi *et al.*, *Phys. Rev. Lett.* **71** (1993) 496; *Phys. Lett.* **B323** (1994) 408.
- [63] A.V. Manohar and M.B. Wise, *Phys. Rev.* **D49** (1994) 1310.
- [64] M. Gremm and A. Kapustin, *Phys. Rev.* **D55** (1997) 6924.
- [65] D. Benson, I.I. Bigi, T. Mannel and N. Uraltsev, *Nucl. Phys.* **B665** (2003) 367.
- [66] C.W. Bauer, Z. Ligeti, M. Luke, A.V. Manohar and M. Trott, *Phys. Rev.* **D70** (2004) 094017.
- [67] P. Gambino and N. Uraltsev, *Eur. Phys. J.* **C34** (2004) 181.
- [68] D. Benson, I.I. Bigi and N. Uraltsev, *Nucl. Phys.* **B710** (2005) 371.
- [69] O. Buchmuller and H. Flacher, *Phys. Rev.* **D73** (2006) 073008.
- [70] T. Mannel, S. Turczyk and N. Uraltsev, *JHEP* **1011** (2010) 109.
- [71] P. Gambino, *JHEP* **1109** (2011) 055.
- [72] A. Alberti, T. Ewerth, P. Gambino and S. Nandi, *Nucl. Phys.* **B870** (2013) 16.
- [73] A. Alberti, P. Gambino and S. Nandi, *JHEP* **1401** (2014) 147.
- [74] T. Mannel, A. A. Pivovarov and D. Rosenthal, *Phys. Rev.* **D92** (2015) 054025.
- [75] A. Alberti, P. Gambino, K. J. Healey and S. Nandi, *Phys. Rev. Lett.* **114** (2015) 061802.
- [76] P. Gambino, K. J. Healey and S. Turczyk, *Phys. Lett.* **B763** (2016) 60.
- [77] P. Ball and R. Zwicky, *Phys. Rev.* **D71** (2005) 014015.
- [78] G. Duplancic *et al.*, *JHEP* **0804** (2008) 014.
- [79] A. Khodjamirian *et al.*, *Phys. Rev.* **D83** (2011) 094031.
- [80] A. Bharucha, *JHEP* **1205** (2012) 092.
- [81] J. A. Bailey *et al.* [Fermilab Lattice and MILC Collaborations], *Phys. Rev.* **D92** (2015) 014024.
- [82] J. M. Flynn *et al.*, *Phys. Rev.* **D91** (2015) 074510.
- [83] M. Antonelli *et al.*, *Phys. Rept.* **494** (2010) 197.
- [84] B.O. Lange *et al.*, *Phys. Rev.* **D72** (2005) 073006; *Nucl. Phys.* **B699** (2004) 335.
- [85] J.R. Andersen and E. Gardi, *JHEP* **0601** (2006) 097.

- [86] P. Gambino, P. Giordano, G. Ossola and N. Uraltsev, *JHEP* **0710** (2007) 058.
- [87] U. Aglietti, F. Di Lodovico, G. Ferrera and G. Ricciardi, *Eur. Phys. J.* **C59** (2009) 831.
- [88] C. Greub, M. Neubert and B. D. Pecjak, *Eur. Phys. J.* **C65** (2010) 501.
- [89] G. Paz, *JHEP* **0906** (2009) 083.
- [90] Z. Ligeti *et al.*, *Phys. Rev.* **D82** (2010) 033003; **D78** (2008) 114014; arXiv:1101.3310 [hep-ph].
- [91] P. Gambino and J.F. Kamenik, *Nucl. Phys.* **B840** (2010) 424.
- [92] LHCb Collaboration, *Nature Phys.* **11** (2015) 743.
- [93] BaBar Collaboration, *Phys. Rev.* **D77** (2008) 011107; **D81** (2010) 051101; **D88** (2013) 031102.
- [94] Belle Collaboration, *Phys. Rev. Lett.* **110** (2013) 131801. *Phys. Rev.* **D92** (2015) 051102.
- [95] CMS Collaboration, *Phys. Lett.* **B736** (2014) 33.
- [96] The ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group and the SLD Electroweak and Heavy Flavour Groups, arXiv:1012.2367 [hep-ex]; <http://www.cern.ch/LEPEWWG/>.
- [97] The ALEPH, DELPHI, L3, OPAL and SLD Collaborations, the LEP Electroweak Working Group and the SLD Electroweak and Heavy Flavour Groups, *Phys. Rept.* **427** (2006) 257.
- [98] L. Wolfenstein, *Phys. Rev. Lett.* **51** (1983) 1945.
- [99] A.J. Buras, M.E. Lautenbacher and G. Ostermaier, *Phys. Rev.* **D50** (1994) 3433.
- [100] T. Inami and C.S. Lim, *Progr. Theor. Phys.* **65** (1981) 297.
- [101] M.K. Gaillard and B.W. Lee, *Phys. Rev.* **D10** (1974) 897.
- [102] ARGUS Collaboration, *Phys. Lett.* **B192** (1987) 245.
- [103] CDF Collaboration, *Phys. Rev. Lett.* **97** (2006) 242003.
- [104] A.J. Buras, M. Jamin and P.H. Weisz, *Nucl. Phys.* **B347** (1990) 491.
- [105] S. Herrlich and U. Nierste, *Nucl. Phys.* **B419** (1994) 292; **B476** (1996) 27.
- [106] A. Pich and J. Prades, *Phys. Lett.* **B346** (1995) 342.
- [107] C. Jarlskog, *Phys. Rev. Lett.* **55** (1985) 1039; *Z. Phys.* **C29** (1985) 491.
- [108] CKMfitter Group, *Eur. Phys. J.* **C41** (2005) 1; <http://ckmfitter.in2p3.fr/>.
- [109] NA48 Collaboration, *Phys. Lett.* **B544** (2002) 97; **B465** (1999) 335; *Eur. Phys. J.* **C22** (2001) 231.
- [110] KTeV Collaboration, *Phys. Rev.* **D83** (2011) 092001; **D67** (2003) 012005; *Phys. Rev. Lett.* **83** (1999) 22.
- [111] NA31 Collaboration, *Phys. Lett.* **B317** (1993) 233; **B206** (1988) 169.
- [112] E731 Collaboration, *Phys. Rev. Lett.* **70** (1993) 1203.
- [113] A.J. Buras, M. Jamin and M.E. Lautenbacher, *Nucl. Phys.* **B408** (1993) 209; *Phys. Lett.* **B389** (1996) 749.
- [114] M. Ciuchini *et al.*, *Phys. Lett.* **B301** (1993) 263; *Z. Phys.* **C68** (1995) 239.
- [115] E. Pallante and A. Pich, *Phys. Rev. Lett.* **84** (2000) 2568; *Nucl. Phys.* **B592** 294.
- [116] E. Pallante, A. Pich and I. Scimemi, *Nucl. Phys.* **B617** (2001) 441.
- [117] V. Cirigliano, A. Pich, G. Ecker and H. Neufeld, *Phys. Rev. Lett.* **91** (2003) 162001; *Eur. Phys. J.* **C33** (2004) 369
- [118] H. Gisbert and A. Pich, *Rep. Prog. Phys.* (2018) in print, arXiv:1712.06147 [hep-ph].
- [119] G. Buchalla, A. J. Buras and M. E. Lautenbacher, *Rev. Mod. Phys.* **68** (1996) 1125.
- [120] J. Brod and M. Gorbahn, *Phys. Rev.* **D82** (2010) 094026.
- [121] A.B. Carter and A.I. Sanda, *Phys. Rev. Lett.* **45** (1980) 952; *Phys. Rev.* **D23** (1981) 1567.
- [122] I.I. Bigi and A.I. Sanda, *Nucl. Phys.* **B193** (1981) 85.
- [123] P. Krawczyk *et al.*, *Nucl. Phys.* **B307** (1988) 19.

- [124] Belle collaboration, *Phys. Rev. Lett.* **108** (2012) 171802.
- [125] BaBar and Belle Collaborations, arXiv:1804.06152 [hep-ex].
- [126] M. Gronau and D. London, *Phys. Rev. Lett.* **65** (1990) 3381.
- [127] M. Gronau and D. London, *Phys. Lett.* **B253** (1991) 483.
- [128] M. Gronau and D. Wyler, *Phys. Lett.* **B265** (1991) 172.
- [129] D. Atwood *et al.*, *Phys. Rev. Lett.* **78** (1997) 3257; *Phys. Rev.* **D63** (2001) 036005.
- [130] UTfit Collaboration, *JHEP* **0507** (2005) 028; <http://www.utfit.org/UTfit/>.
- [131] A. Lenz and U. Nierste, *JHEP* **0706** (2007) 072.
- [132] M. Artuso, G. Borissov and A. Lenz, *Rev. Mod. Phys.* **88** (2016) 045002.
- [133] T. Jubb, M. Kirk, A. Lenz and G. Tetlalmatzi-Xolocotzi, *Nucl. Phys.* **B915** (2017) 431.
- [134] LHCb Collaboration, *Eur. Phys. J.* **C77** (2017) 678.
- [135] M. Gorbahn and U. Haisch, *Phys. Rev. Lett.* **97** (2006) 122002.
- [136] D. Gomez Dumm and A. Pich, *Phys. Rev. Lett.* **80** (1998) 4633.
- [137] G. Ecker and A. Pich, *Nucl. Phys.* **B366** (1991) 189.
- [138] G. Ecker, A. Pich and E. de Rafael, *Nucl. Phys.* **B303** (1988) 665; **B291** (1987) 692; *Phys. Lett.* **B189** (1987) 363.
- [139] A. J. Buras, M. E. Lautenbacher, M. Misiak and M. Munz, *Nucl. Phys.* **B423** (1994) 349.
- [140] G. Buchalla, G. D'Ambrosio and G. Isidori, *Nucl. Phys.* **B672** (2003) 387.
- [141] KTeV Collaboration, *Phys. Rev. Lett.* **93** (2004) 021805.
- [142] A. J. Buras, M. Gorbahn, U. Haisch and U. Nierste, *Phys. Rev. Lett.* **95** (2005) 261805; *JHEP* **0611** (2006) 002 [Err: **1211** (2012) 167].
- [143] J. Brod, M. Gorbahn and E. Stamou, *Phys. Rev.* **D83** (2011) 034030.
- [144] A. J. Buras, D. Buttazzo, J. Girrbach-Noe and R. Knegjens, *JHEP* **1511** (2015) 033.
- [145] E949 Collaboration, *Phys. Rev. Lett.* **101** (2008) 191802.
- [146] E391a Collaboration, *Phys. Rev. Lett.* **100** (2008) 201802; *Phys. Rev.* **D81** (2010) 072004.
- [147] G. Anelli *et al.*, CERN-SPSC-2005-013, CERN-SPSC-P-326 (2005).
- [148] B. Beckford [KOTO Collaboration], arXiv:1710.01412 [hep-ex].
- [149] M. Misiak *et al.*, *Phys. Rev. Lett.* **114** (2015) 221801.
- [150] M. Czakon *et al.*, *JHEP* **1504** (2015) 168.
- [151] CMS and LHCb Collaborations, *Nature* **522** (2015) 68.
- [152] LHCb Collaboration, *Phys. Rev. Lett.* **118** (2017) 191801.
- [153] K. De Bruyn *et al.*, *Phys. Rev. Lett.* **109** (2012) 041801.
- [154] C. Bobeth *et al.*, *Phys. Rev. Lett.* **112** (2014) 101801.
- [155] S. Weinberg, *Phys. Rev. Lett.* **43** (1979) 1566.
- [156] G. Isidori, Y. Nir and G. Perez, *Ann. Rev. Nucl. Part. Sci.* **60** (2010) 355.
- [157] L. J. Hall and L. Randall, *Phys. Rev. Lett.* **65** (1990) 2939.
- [158] R. S. Chivukula and H. Georgi, *Phys. Lett.* **B188** (1987) 99.
- [159] G. D'Ambrosio, G. F. Giudice, G. Isidori and A. Strumia, *Nucl. Phys.* **B645** (2002) 155.
- [160] A. Pich and P. Tuzón, *Phys. Rev.* **D80** (2009) 091702.
- [161] A. Peñuelas and A. Pich, *JHEP* **84** (2017).
- [162] M. Jung, A. Pich and P. Tuzón, *JHEP* **1011** (2010) 003; *Phys. Rev.* **D83** (2011) 074011.
- [163] M. Jung, X. Q. Li and A. Pich, *JHEP* **1210** (2012) 063.
- [164] M. Jung and A. Pich, *JHEP* **1404** (2014) 076.
- [165] X. Q. Li, J. Lu and A. Pich, *JHEP* **1406** (2014) 022.

- [166] V. Ilisie, *JHEP* **1504** (2015) 077.
- [167] A. Cherchiglia, P. Kneschke, D. Stöckinger and H. Stöckinger-Kim, *JHEP* **1701** (2017) 007.
- [168] Q. Chang, P. F. Li and X. Q. Li, *Eur. Phys. J.* **C75** (2015) 594.
- [169] Q. Y. Hu, X. Q. Li and Y. D. Yang, *Eur. Phys. J.* **C77** (2017) 190; **C77** (2017) 228.
- [170] N. Cho, X. Q. Li, F. Su and X. Zhang, *Adv. High Energy Phys.* **2017** (2017) 2863647.
- [171] T. Han, S. K. Kang and J. Sayre, *JHEP* **1602** (2016) 097.
- [172] T. Enomoto and R. Watanabe, *JHEP* **1605** (2016) 002.
- [173] A. Celis, V. Ilisie and A. Pich, *JHEP* **1307** (2013) 053; **1312** (2013) 095.
- [174] G. Abbas, A. Celis, X. Q. Li, J. Lu and A. Pich, *JHEP* **1506** (2015) 005.
- [175] V. Ilisie and A. Pich, *JHEP* **1409** (2014) 089.
- [176] W. Altmannshofer, S. Gori and G. D. Kribs, *Phys. Rev.* **D86** (2012) 115009.
- [177] Y. Bai, V. Barger, L. L. Everett and G. Shaughnessy, *Phys. Rev.* **D87** (2013) 115013.
- [178] L. Duarte, G. A. González-Sprinberg and J. Vidal, *JHEP* **1311** (2013) 114.
- [179] C. Ayala, G. A. González-Sprinberg, R. Martinez and J. Vidal, *JHEP* **1703** (2017) 128.
- [180] L. Wang and X. F. Han, *JHEP* **1404** (2014) 128.
- [181] A. Filipuzzi, J. Portolés and M. González-Alonso, *Phys. Rev.* **D85** (2012) 116010.
- [182] BaBar Collaboration, *Phys. Rev.* **D85** (2012) 031102 [Err: **D85** (2012) 099904].
- [183] I. I. Bigi and A. I. Sanda, *Phys. Lett.* **B625** (2005) 47.
- [184] Y. Grossman and Y. Nir, *JHEP* **1204** (2012) 002.
- [185] V. Cirigliano, A. Crivellin and M. Hoferichter, *Phys. Rev. Lett.* **120** (2018) 141803.
- [186] Belle Collaboration, *Phys. Rev. Lett.* **97** (2006) 251802; *Phys. Rev.* **D82** (2010) 071101.
- [187] BaBar Collaboration, *Phys. Rev. Lett.* **109** (2012) 101802.
- [188] S. Fajfer, J. F. Kamenik and I. Nisandzic, *Phys. Rev.* **D85** (2012) 094025.
- [189] LHCb Collaboration, *Phys. Rev. Lett.* **115** (2015) 111803 [Err: **115** (2015) 159901]; **120** (2018) 171802; *Phys. Rev.* **D97** (2018) 072013.
- [190] Belle Collaboration, *Phys. Rev.* **D92** (2015) 072014; **D94** (2016) 072007; **D97** (2018) 012004; *Phys. Rev. Lett.* **118** (2017) 211801.
- [191] H. Na *et al.* [HPQCD Collaboration], *Phys. Rev.* **D92** (2015) 054510 [Err: **D93** (2016) 119906].
- [192] A. Celis, M. Jung, X. Q. Li and A. Pich, *JHEP* **1301** (2013) 054; *Phys. Lett.* **B771** (2017) 168.
- [193] R. Alonso, B. Grinstein and J. Martin Camalich, *JHEP* **1510** (2015) 184.
- [194] A. Crivellin, D. Müller and T. Ota, *JHEP* **1709** (2017) 040.
- [195] B. Capdevila *et al.*, *Phys. Rev. Lett.* **120** (2018) 181802.
- [196] LHCb Collaboration, *Phys. Rev. Lett.* **120** (2018) 121801.
- [197] A. Y. Anisimov, I. M. Narodetsky, C. Semay and B. Silvestre-Brac, *Phys. Lett.* **B452** (1999) 129.
- [198] V. V. Kiselev, hep-ph/0211021.
- [199] M. A. Ivanov, J. G. Korner and P. Santorelli, *Phys. Rev.* **D73** (2006) 054024.
- [200] E. Hernandez, J. Nieves and J. M. Verde-Velasco, *Phys. Rev.* **D74** (2006) 074008.
- [201] C. Bobeth and U. Haisch, *Acta Phys. Polon.* **B44** (2013) 127.
- [202] LHCb Collaboration, *JHEP* **1509** (2015) 179; **1307** (2013) 084.
- [203] LHCb Collaboration, *JHEP* **1406** (2014) 133.
- [204] LHCb Collaboration, *JHEP* **1302** (2013) 105; *Eur. Phys. J.* **C77** (2017) 161.
- [205] LHCb Collaboration, *JHEP* **1611** (2016) 047 [Err: **1704** (2017) 142]; **1602** (2016) 104; **1308** (2013) 131; *Phys. Rev. Lett.* **111** (2013) 191801.
- [206] LHCb Collaboration, *JHEP* **1506** (2015) 115; *Phys. Lett.* **B725** (2013) 25.

- [207] ATLAS Collaboration, arXiv:1805.04000 [hep-ex].
- [208] BaBar Collaboration, *Phys. Rev.* **D93** (2016) 052015.
- [209] Belle Collaboration, *Phys. Rev. Lett.* **118** (2017) 111801; **103** (2009) 171801; arXiv:1604.04042 [hep-ex].
- [210] CDF Collaboration, *Phys. Rev. Lett.* **108** (2012) 081807.
- [211] CMS Collaboration, *Phys. Lett.* **B781** (2018) 517; *Phys. Lett.* **B753** (2016) 424.
- [212] S. Descotes-Genon, J. Matias, M. Ramon and J. Virto, *JHEP* **1301** (2013) 048.
- [213] S. Descotes-Genon, L. Hofer, J. Matias and J. Virto, *JHEP* **1412** (2014) 125.
- [214] W. Altmannshofer and D. M. Straub, *Eur. Phys. J.* **C75** (2015) 382.
- [215] A. Bharucha, D. M. Straub and R. Zwicky, *JHEP* **1608** (2016) 098.
- [216] B. Capdevila, S. Descotes-Genon, L. Hofer and J. Matias, *JHEP* **1704** (2017) 016.
- [217] S. Jäger and J. Martin Camalich, *Phys. Rev.* **D93** (2016) 014028.
- [218] M. Ciuchini *et al.*, *JHEP* **1606** (2016) 116; PoS ICHEP **2016** (2016) 584.
- [219] V. G. Chobanova *et al.*, *JHEP* **1707** (2017) 025.
- [220] C. Bobeth, M. Chrzaszcz, D. van Dyk and J. Virto, arXiv:1707.07305 [hep-ph].
- [221] T. Blake, U. Egede, P. Owen, G. Pomery and K. A. Petridis, arXiv:1709.03921 [hep-ph].
- [222] LHCb Collaboration, *Phys. Rev. Lett.* **113** (2014) 151601.
- [223] F. Dettori [LHCb Collaboration], arXiv:1805.05073 [hep-ex].
- [224] LHCb Collaboration, *JHEP* **1708** (2017) 055.
- [225] C. Bobeth, G. Hiller, D. van Dyk and C. Wacker, *JHEP* **1201** (2012) 107.
- [226] G. Hiller and F. Kruger, *Phys. Rev.* **D69** (2004) 074020.
- [227] C. Bobeth, G. Hiller and G. Piranishvili, *JHEP* **0712** (2007) 040.
- [228] C. Bouchard *et al.* [HPQCD Collaboration], *Phys. Rev. Lett.* **111** (2013) 162002 [Err: **112** (2014) 149902].
- [229] M. Bordone, G. Isidori and A. Pattori, *Eur. Phys. J.* **C76** (2016) 440.
- [230] B. Capdevila, A. Crivellin, S. Descotes-Genon, J. Matias and J. Virto, *JHEP* **1801** (2018) 093.
- [231] S. Descotes-Genon, L. Hofer, J. Matias and J. Virto, *JHEP* **1606** (2016) 092.
- [232] W. Altmannshofer, P. Stangl and D. M. Straub, *Phys. Rev.* **D96** (2017) 055008.
- [233] L. S. Geng *et al.*, *Phys. Rev.* **D96** (2017) 093006.
- [234] T. Hurth, F. Mahmoudi, D. Martinez Santos and S. Neshatpour, *Phys. Rev.* **D96** (2017) 095034.
- [235] S. Fajfer, N. Košnik and L. Vale Silva, *Eur. Phys. J.* **C78** (2018) 275.
- [236] D. Marzocca, arXiv:1803.10972 [hep-ph].
- [237] J. E. Camargo-Molina, A. Celis and D. A. Farougy, arXiv:1805.04917 [hep-ph].



# Beyond the Standard Model

*M. Mondragon*

Instituto de Fisica, Universidad Nacional Autonoma de Mexico UNAM, CD MX 01000, Mexico

## Abstract

I present a brief outline of some of the different paths for extending the Standard Model. These include supersymmetry, Grand Unified Theories, extra dimensions, and multi-Higgs models. The aim is to give graduate students an overview of some of the theoretical motivations to go into a specific extension, and a few examples of how the experimental searches go hand in hand with these efforts.

Supersymmetry, Grand Unified Theories, Extra dimensions, Multi-Higgs models

## 1 Introduction

The Standard Model is an extremely successful description of the elementary particles and its interactions around the Fermi scale, but it leaves a number of unanswered questions that lead naturally to the conclusion that it is the low energy limit of a more fundamental theory. Along with these unanswered question comes a large number of free parameters whose value can only be determined experimentally. Among the unanswered questions in the Standard Model are:

- Are there more than three generations of particles?
- Why are the fundamental particles masses so different?
- How is the Higgs mass stabilized, i.e. what solves the hierarchy problem?
- Is there more than one Higgs?
- What is the nature of the neutrinos, are they Dirac or Majorana?
- What is dark matter?
- Why is there more matter than anti-matter?
- Is there more CP violation?
- Why only left-handed particles feel the electroweak interaction?
- Is it possible to unify quantum mechanics with gravity?

The collection of models and theories that make the research that extends the Standard Model to answer some of these, and many more, questions is known as physics Beyond the Standard Model (BSM). In here we will briefly explain some of the more popular efforts to go beyond the Standard Model. There are many excellent reviews devoted to each topic, in here we give only an overview of some of these extensions with an emphasis on the interplay between theory and experiment.

## 2 Symmetries

Modern physics is built on the observation that there are symmetries in Nature. The usual or traditional way of relating different parameters or sectors of a theory has been by adding symmetries. The Lagrangian of the theory has to respect the symmetries of the system, which means it remains invariant under symmetry transformations. This requirement means that the mathematical consistency of a proposed model can lead to new discoveries.

There are different types of symmetries: Continuous, discrete, global, local, hidden, broken, accidental... the Standard Model incorporates all of these kind of symmetries.

Continuous symmetries are implemented by unitary transformations, since they are continuously connected to the identity. Continuous symmetries lead to conserved quantities, as stated by Noether's theorem [1], which are useful to classify or label the symmetries. These systems might have local or global conservation laws; global if the symmetry acts equally everywhere and local if the symmetry acts differently on every point.

On the other hand, discrete symmetries, as their name states, represent non-continuous transformations. In quantum mechanics they have associated a multiplicative quantum number.

The quantum field theory (QFT) that describes the SM is based on a combination of space-time symmetries and internal symmetries. Space-time symmetries act on the coordinates of space-time, they include rotations, translations, and general Lorentz and Poincaré transformations, which are global symmetries, and general coordinate transformations, which are local symmetries. Thus, the SM has all the transformations of special and general relativity.

On the other hand, internal symmetries refer to the transformations of the different fields. Internal symmetries are related to gauge invariance, which refers to the fact that different field configurations may describe the same observables. A gauge transformation on a field implies that also the kinetic term has to be transformed, so as to leave the Lagrangian invariant. This leads to the definition of the covariant derivative, where a vector or gauge potential is introduced, ensuring the invariance of the Lagrangian under gauge transformations. There are many choices of gauge potential that will describe the same physics. From the conservation of internal symmetries follows the conservation of colour and electric charge, for instance.

Often symmetries are not exact or explicit, but may be broken or hidden. A symmetry can be broken spontaneously, like in the Higgs mechanism, where the vacuum expectation value of the Higgs field is different from zero, and thus does not respect the symmetries of the original Lagrangian. Thus, a symmetry may be hidden in the sense that we may just perceive the symmetries in the vacuum state, and not the more general symmetry group of the complete Lagrangian.

A symmetry can also be broken explicitly, whereby there are terms in the Lagrangian which do not respect it, and thus are not invariant under the symmetry transformation.

The Standard Model is based on a number of symmetries: Poincaré invariance in four dimensions and gauge invariance under the group  $SU(3) \times SU(2)_L \times U(1)_Y$ . The electroweak part  $SU(2)_L \times U(1)_Y$  is spontaneously broken by the Higgs mechanism to  $U(1)_{EM}$ . Besides Lorentz and gauge invariance the SM exhibits some discrete symmetries, which can be exact or broken. Among these are charge conjugation (C), parity (P), and the combined effect of both CP, as well as time reversal (T), which are broken in electroweak interactions. The SM Lagrangian is invariant under the action of CPT, as required by the CPT Theorem, which states that any localized Lorentz invariant gauge theory is invariant under the combined action of C, P, and T. There are also accidental symmetries in the SM, that is, they appear without having required invariance under any particular symmetry, but as a consequence of the field content and other symmetries and properties of the Lagrangian. These include baryon and lepton number conservation, which are exact at the classical level.

### 3 Why go beyond the SM?

As already mentioned, the SM is very successful in describing the fundamental particles and their interactions at the electroweak scale, but it leaves a number of puzzles or unanswered questions that lead to the conclusion that it is the low energy limit of a more fundamental theory. The first deviations of the SM appeared with the neutrino masses and the existence of dark matter, the latter was actually proposed even before the SM was conceived. The SM was formulated with neutrino masses equal to zero, and it might not look like a great departure from it just to add extra terms to the Lagrangian to give mass to the neutrinos, the same way as for the other matter particles. But neutrino masses are tiny, and thus the associated Yukawa coupling would be several orders of magnitude smaller than the other matter cou-

plings, enhancing the mass hierarchy among the fundamental particles. A popular way to explain the smallness of neutrino masses is to assume the right-handed neutrinos, which are sterile, acquire a large Majorana mass. The diagonalisation of the neutrino mixing matrix which includes the Dirac mass term (of order of the electroweak scale) and the Majorana mass term leads to one very heavy and one very light eigenvalue, thus the name seesaw mechanism (see for instance [2, 3]).

Amongst the puzzles that lead to the proposal of different models BSM are some of the questions mentioned in the introduction, and more. The list is long and it grows as we go into the details. Let us look at some of the more important ones:

- The hierarchy problem refers to the fact that we expect the fundamental scale to be the Planck scale,  $M_P = 10^{19} \text{ GeV}$ , but the particle masses are “of the order” of the electroweak scale  $\sim \mathcal{O}(100) \text{ GeV}$ , i.e. 17 orders of magnitude below  $M_P$ . This scale is dictated by the spontaneous electroweak symmetry breaking, where the Higgs fields acquire a vacuum expectation value (vev) different from zero, and through its self-interaction a mass, which we know now to be  $\sim 125 \text{ GeV}$ .

On the other hand, the Higgs field, being a scalar, has radiative corrections that grow quadratically with the energy, and not logarithmically like the other particles. The quantum corrections to the Higgs mass would drive it to the Planck scale, which is assumed to be the fundamental one

$$M_h^2 \propto M^2(\Lambda^2) + \delta m_h^2 = M^2(\Lambda^2) - Cg^2\Lambda^2, \quad (1)$$

where  $\Lambda$  is an energy scale,  $g$  the SU(2) gauge coupling and  $C$  some constant. Thus, an incredible fine-tuning must exist between the bare mass and the radiative corrections to cancel almost completely among themselves, and render the Higgs mass  $\sim 125 \text{ GeV}$ . This enormous degree of fine-tuning is clearly a naturality issue. The hierarchy and naturalness (or lack thereof) problems have led to the conclusion that around  $\sim \mathcal{O}(1) \text{ TeV}$  there must appear new physics, or rather physics unknown to us.

- The cosmological constant is interpreted as the energy of the vacuum, and recent observations suggest its value is close to zero but not identically zero. In quantum field theory, assuming the fundamental theory to be at the Planck scale, the cosmological constant is predicted to be  $\sim 120$  orders of magnitude bigger than the measured value. This is an even bigger naturalness problem than the one of the Higgs mass.

- There are only three generations of particles, as is inferred from the Z width and from Big Bang Nucleosynthesis. The masses of the fundamental fermions are much lower than the Planck scale, as above mentioned. But there are also large differences among them, between the up and the top quarks there are five orders of magnitude in mass. The mixing in the quark sector is not very large, the CKM matrix largest elements are in the diagonal. Also in the lepton sector, if we take into account neutrino masses (which strictly speaking are zero in the SM), the discrepancy between the largest and the lightest mass is at least five orders of magnitude, if not bigger. On the other hand, the mixing angles in the PMNS matrix are fairly large. This pattern of masses and mixings is not understood, but it might suggest an underlying symmetry at work (or not).

- We have enough evidence of the existence of dark matter, and we will assume that it is a particle. This particle cannot be one of the SM ones. The neutrinos could be a fraction of the total dark matter (DM), but DM cannot be composed entirely of neutrinos, since it would be in contradiction with the large structure formation of the Universe. Thus, we need particles beyond the ones in the SM to explain dark matter (see [4, 5] and references therein). It is usually assumed that DM is only composed of one type of particle, but there is no reason that DM could be actually made of more than one type of particle.

The solution to some of these problems has been to follow the traditional path, i.e. to add symmetries, but it is also possible to add particles and/or interactions, or to add more space-time dimensions, or combinations of all of them. Of course, any attempt to extend the SM has to go hand in hand with experimental data, and the SM has to emerge as the low-energy theory of any theory beyond it. For extended recent reviews and textbooks on BSM physics from different perspectives, see for instance [6–13].

#### 4 Grand Unified Theories – GUTs

In the context of adding more symmetries, one could think that there is a larger gauge symmetry group which contains the SM one. This is the idea behind Grand Unified Theories or GUTs, where the electroweak and strong interactions are unified in a single interaction at high energies. This larger symmetry group is realized at very high energies, which corresponds to the beginning of the Universe, and the symmetry is broken down to the SM gauge group, which is what we observe today. The unification idea is very attractive, what we perceive as three separate interactions at low energies is in reality only one, gravity is not included in this scheme. Grand unification can work because of the behaviour of the different gauge couplings as they move up in energy, which is calculated using the Renormalization Group Equations (RGEs). Whereas the inverse of the strong and weak coupling increase with energy, the inverse of the electromagnetic coupling decreases, thus they all tend to a similar value at high energies (see Fig. 4). If unified, we have two parameters, the unification scale  $M_{GUT}$  and the unification coupling  $\alpha_{GUT}$  rather than three. But reducing one parameter is not the only achievement of a unified theory. Because now the particles are unified in a larger symmetry group, they are related through the symmetry, which gives some nice surprises and predictions (see for instance [14, 15]).

The simplest, and one of the most studied, examples of a GUT is the one based in the gauge group  $SU(5)$ .  $SU(5)$  can accommodate all the particles of the SM and can be broken exactly into its gauge group,

$$SU(5) \rightarrow SU(3) \times SU(2) \times U(1) \quad (2)$$

The left-handed quarks and leptons are accommodated in irreducible representations (irreps) of  $SU(5)$  as follows  $\bar{\mathbf{5}} = [\mathbf{d}^c, \mathbf{L}]$  and  $\mathbf{10} = [\mathbf{Q}, \mathbf{u}^c, \mathbf{e}^c]$ , and the left-handed anti-neutrinos are in a singlet irrep  $\mathbf{1}$  (the original model did not include those, since there was not yet evidence of neutrino masses). In matrix notation

$$\bar{\mathbf{5}} = \begin{pmatrix} d^c \\ d^c \\ d^c \\ e \\ -\nu_e \end{pmatrix}, \quad \mathbf{10} = \begin{pmatrix} 0 & u^c & -u^c & -u & -d \\ & 0 & u^c & -u & d \\ & & 0 & -u & -d \\ & & & 0 & -e^+ \end{pmatrix}. \quad (3)$$

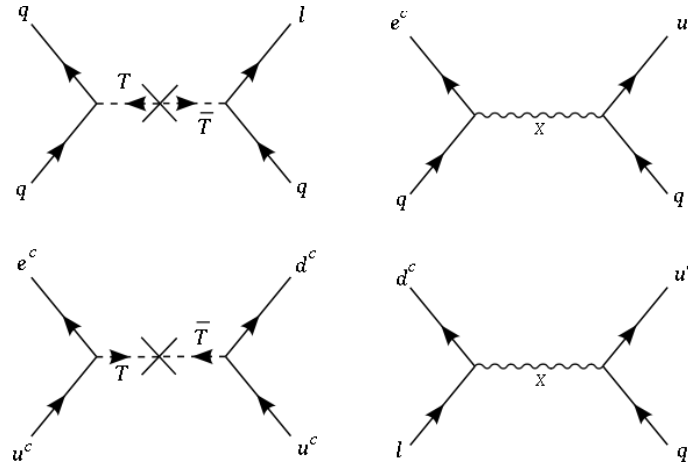
The breaking of the  $SU(5)$  group to the SM one is achieved by the vacuum expectation value of the adjoint representation, the  $\mathbf{24}$ . The gauge bosons fit in the adjoint irrep  $\mathbf{24}$ , which also contains twelve fractionally charged bosons with both lepton and baryon number. These “exotic” bosons are called leptoquarks, and are denoted by X and Y. Since the leptons and quarks are combined in the same irreps, baryon and lepton numbers are not conserved, although the combination B-L is still conserved. Thus, unless the X and Y bosons are very heavy, the violation of B and L can lead to proton decay through exchange of leptoquarks, which is a signature of many GUTs. These decay modes obey the selection rule  $\Delta(B - L) = 0$ , and are mediated by effective operators of dimension 6. In Fig. 1 examples of diagrams that lead to proton decay through these dimension six operators are shown.

The Higgs boson can be accommodated in a  $\mathbf{5}$  or a  $\bar{\mathbf{5}}$  irrep, which will also decompose in triplets of  $SU(3)$  and doublets of  $SU(2)$  once the  $SU(5)$  group is broken. Again, we have a coloured and an electroweak part mixed in the same irrep. These coloured Higgs triplets violate baryon and lepton number and can mediate fast proton decay via the exchange of the scalar triplet, unless they are very heavy. Thus, there must be a fine tuning to get the doublets at the electroweak scale, and at the same time leave the triplets at the GUT scale. This is referred to as the doublet-triplet splitting fine tuning problem.

The proton lifetime in these type of GUTs is

$$\tau \sim \frac{M_{GUT}^4}{\alpha_{GUT}^2 m_p^5} \sim 10^{30} - 10^{31} \text{ yrs}, \quad (4)$$

which excludes them by the bound on the proton lifetime set by super-Kamiokande  $\tau > 10^{34} \text{ yrs}$ , ruling out the simplest models. Moreover, letting the gauge couplings of the SM run to high energies



**Fig. 1:** Dimension six operators that mediate proton decay in  $SU(5)$ , the diagrams on the left are mediated by the coloured Higgs triplet and anti-triplet, the ones on the right by the X boson.

through their RGE shows that they do not quite unify, see upper panel of Fig. 4. If one incorporates supersymmetry in the game, then unification may be achieved, as the lower panel of Fig. 4 shows. Although in SUSY GUTs the proton decay prediction remains, the proton lifetime may be larger than in non-SUSY models, within the experimental bounds.

On the other hand, Grand Unified theories give an explanation of charge quantization. Since the charge operator is a generator of  $SU(5)$ , the sum of all charges in each irrep must add to zero, leading to  $Q(d^c) = -1/3Q(e)$ . They give an approximation to the value of the Weinberg angle, and also approximations to some quark mass ratios.

Besides the minimal  $SU(5)$  sketched here, several other GUTs have been studied. For instance, another popular GUT is  $SO(10)$ , which has all fermions in the same irrep  $\mathbf{16}$ , and a right handed neutrino may be included naturally. Other groups have different features like the Pati-Salam group  $SU(4)_c \times SU(2)_L \times SU(2)$ , which proposes a fourth colour charge and also predicts the existence of magnetic monopoles, or the flipped  $SU(5) \times U(1)$ , where the assignment of the particles to the irreps is  $\bar{\mathbf{5}} = [u^c, \mathbf{L}]$  and  $\mathbf{10} = [Q, d^c, \nu^c]$ , and  $\mathbf{1} = e^c$ .

## 5 Supersymmetry

Particles come in two fundamental versions, bosons and fermions. Matter particles are fermions, with half-integer spin and obey Fermi-Dirac statistics. Gauge particles, which are the ones that “carry” the interactions, are vector bosons with spin one and obey Bose-Einstein statistics. The Higgs boson (or perhaps Higgs bosons) is a scalar with zero spin, and thus also obeys Bose-Einstein statistics. Supersymmetry (SUSY) relates bosons and fermions, so what in the SM are distinct types of particles, in SUSY theories are related through a symmetry transformation. There are many excellent reviews on supersymmetry, both from the formal as well as from the more phenomenological point of view (see for instance [16–18, 21]). In here we will expose the main motivation for SUSY and some of models that have been more studied.

Through SUSY it is possible to transform bosons into fermions and viceversa

$$Q|Boson \rangle = |Fermion \rangle; \quad Q|Fermion \rangle = |Boson \rangle. \quad (5)$$

The extension of the Coleman-Mandula Theorem [19] due to Haag, Lopuszanski and Sohnius [20] tells us how to build an interacting quantum field theory with such restrictions. One has to generalize the Lie algebra to a graded Lie algebra with anti-commutators and commutators, in order to include the fermionic

generators. Thus, a supersymmetric theory must have the Poincaré generators  $P^M$  for translations in space and time, and  $M^{M,N}$  for the Lorentz boosts and rotations, plus the spinor generators  $Q_\alpha^A$  ( $\alpha = 1, 2$  and  $A = 1, \dots, \mathcal{N}$ ) corresponding to spins  $(A, B) = (\frac{1}{2}, 0)$  and  $Q_{\dot{\alpha}}^A$  with  $(A, B) = (0, \frac{1}{2})$ .

If  $O_a$  are operators of a Lie algebra, a graded Lie algebra satisfies the relations

$$O_a O_b - (-1)^{\eta_a \eta_b} O_b O_a = i C^e{}_{ab} O_e, \quad (6)$$

where the *gradings*  $\eta_a$  take values

$$\eta_a = \begin{cases} 0 & : O_a \text{ bosonic generator} \\ 1 & : O_a \text{ fermionic generator} \end{cases}. \quad (7)$$

We will only address the case of N=1 supersymmetry, i.e. only one set of  $Q$  and  $Q^\dagger$ . The possibility of having several copies of these generators leads to extended supersymmetries ( $N = 2, N = 4$ ), but these theories do not have chiral fermions, which makes them unsuitable as direct extensions of the SM.

The particle states are described by irreducible representations of the SUSY algebra, called supermultiplets. The supermultiplet contains equal number of both fermionic and bosonic degrees of freedom. The particle states in each supermultiplet are called superpartners and they differ by 1/2 unit of spin. In N=1 SUSY, a chiral supermultiplet contains a Weyl fermion and two real scalars. The Weyl fermion has two degrees of freedom corresponding to each helicity state, and each real scalar has one degree of freedom, which are usually described together as one complex scalar. A gauge supermultiplet contains a massless spin-1 gauge boson and its superpartner, a gluino, which is a massless spin-1/2 Weyl fermion, both of which have two helicity states. They both transform in the adjoint representation of the gauge group which is self-conjugate, therefore the left- and right-handed components of the gauginos have the same transformation properties, thus they are Majorana in nature. The Higgs boson belongs also to a chiral supermultiplet, with a spin-1/2 superpartner, the Higgsino. But in SUSY theories one Higgs supermultiplet is not enough to guarantee an anomaly free theory. The Higgs supermultiplet has a gauge anomaly, which is cancelled if another Higgs supermultiplet with opposite hypercharge is added. The members of a supermultiplet must have the same coupling strengths and also the same mass, thus

$$\sum m_{bosons}^2 = \sum m_{fermions}^2. \quad (8)$$

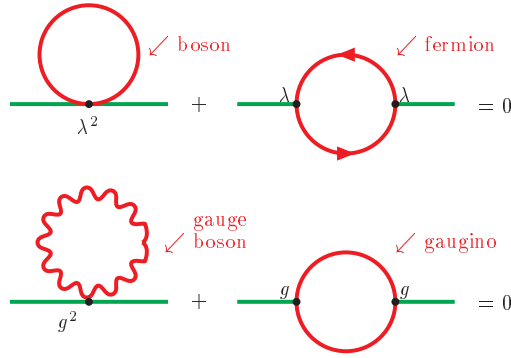
In the context of the radiative corrections to the Higgs mass this means that the quadratic corrections to the Higgs mass coming from the bosons are exactly cancelled by their fermionic superpartners, order by order in perturbation theory.

One of the main motivations to study SUSY models is that they provide a solution to the hierarchy problem, in a natural way. The contributions to the radiative corrections to the Higgs mass that come from the SM particles are cancelled exactly, order by order, by their superpartners, see Fig. 2. This happens thanks to the sum rule, Eq. (8), and an extra (-1) sign that comes from the Fermi-Dirac statistics. This is exact only if SUSY is unbroken, once it is broken the squared mass difference between bosons and fermions is proportional to the susy breaking scale squared  $M_{SUSY}^2$ . This value is expected to be  $\sim \mathcal{O}(1) TeV$ , to be consistent with the scale of the Higgs mass, and it is also the value that is required to have good unification of couplings in the minimal SUSY extension of the MSSM.

## 5.1 The Minimal Supersymmetric Extension of the Standard Model

By far the most phenomenologically viable SUSY model studied is the Minimal Supersymmetric Standard Model (MSSM). It has N=1 SUSY, to include chiral fermions, and two Higgs supermultiplets, to be free of gauge anomalies. The gauge group is the same as the SM one,  $SU(3) \times SU(2) \times U(1)$ , but now the particles are arranged in supermultiplets. The matter or chiral supermultiplets contain the quarks and leptons and their respective superpartners, squarks and sleptons, which are spin-0 scalars. There are two





**Fig. 2:** Cancellation of divergencies to the Higgs boson mass from contributions from the superpartners, figure from ref [21].

Particles	Bosons	Fermions	Charges under SM
quarks, squarks $Q_i$	$(\tilde{u}, \tilde{d})_L$	$(u, d)_L$	$(\mathbf{3}, \mathbf{2}, 1/6)$
$\tilde{u}$	$\tilde{u}_R^*$	$u_R^\dagger$	$(\bar{\mathbf{3}}, \mathbf{2}, -2/3)$
$\tilde{d}$	$\tilde{d}_R^*$	$d_R^\dagger$	$(\bar{\mathbf{3}}, \mathbf{2}, 1/3)$
leptons, sleptons $L_i$	$(\tilde{\nu}, \tilde{e})_L$	$(\nu, e)_L$	$(\mathbf{1}, \mathbf{2}, -1/2)$
$\tilde{e}$	$\tilde{e}_R^*$	$e_R^\dagger$	$(\bar{\mathbf{1}}, \mathbf{1}, 1)$
Higgs bosons $H_u$	$(H_u^+, H_u^0)$	$(\tilde{H}_u^+, \tilde{H}_u^0)$	$(\bar{\mathbf{1}}, \mathbf{2}, 1/2)$
$H_d$	$(H_d^0, H_d^-)$	$(\tilde{H}_d^0, \tilde{H}_d^-)$	$(\bar{\mathbf{1}}, \mathbf{2}, -1/2)$
Bino, B	$\tilde{B}^0$	$B^0$	$(\mathbf{1}, \mathbf{1}, 0)$
Wino, W, Zino, Z	$(\tilde{W}^\pm, \tilde{Z})$	$(W^\pm, Z)$	$(\mathbf{1}, \mathbf{3}, 0)$
Gluino, g	$\tilde{g}$	$g$	$(\mathbf{8}, \mathbf{1}, 0)$

**Table 1:** Particle content of the MSSM

chiral supermultiplets with opposite hypercharge that contain the Higgs bosons and their corresponding superpartners, the Higgsinos. The gauge supermultiplets, which contain the spin-1 vector bosons and their superpartners the gauginos, which are spin-1/2 Majorana fermions.

In supersymmetric theories the interactions are determined by the superpotential, which is a holomorphic function of the chiral superfields. It respects SUSY and gauge invariance. The superpotential contains the Yukawa interactions and mass terms, in the MSSM it is

$$W_{MSSM} = \epsilon_{ij}(y_{ab}^U Q_a^j U_b^c H_2^i + y_{ab}^D Q_a^j D_b^c H_1^i + y_{ab}^L L_a^j E_b^c H_1^i + \mu H_u^i H_d^j) \quad (9)$$

where the superfields in the potential are given in Table 5.1

In principle there could be terms in the superpotential that break baryon or lepton number, like for instance  $LQ\tilde{d}^c$ . These terms would mediate really fast proton decay through  $p \rightarrow e^+ + \pi^0$ . In order to forbid these terms a new multiplicative discrete parity is introduced, R symmetry. R symmetry is defined as

$$R = (-1)^{3(B-L)+2S}, \quad (10)$$

thus all supersymmetry particles have  $R = -1$  and the SM particles and all the Higgs bosons have  $R = 1$ . Besides forbidding proton decay, one immediate consequence of R parity is that the lightest supersymmetric particle (LSP) is stable. This feature may provide with a good candidate to dark matter, if the LSP is electrically neutral. Another consequence of R parity is that superparticles are created in pairs.

## 5.2 SUSY breaking in the MSSM

If SUSY exists in Nature it has to be a broken symmetry, since no superpartners of the SM particles have been observed. The way to break SUSY consistent with phenomenological observations, and to preserve the solution to the hierarchy problem is to add soft breaking terms to the SUSY Lagrangian. Soft breaking terms do not reintroduce quadratic divergences in the theory, they have positive mass dimension, and are super-renormalizable. The soft breaking part of the MSSM Lagrangian has a similar form to the superpotential

$$\begin{aligned}
 -\mathcal{L}_{\text{Breaking}} = & \sum_i m_{0i}^2 |\varphi_i|^2 + \left( \frac{1}{2} \sum_{\alpha} M_{\alpha} \tilde{\lambda}_{\alpha} \tilde{\lambda}_{\alpha} + B H_1 H_2 \right. \\
 & \left. + A_{ab}^U \tilde{Q}_a \tilde{U}_b^c H_2 + A_{ab}^D \tilde{Q}_a \tilde{D}_b^c H_1 + A_{ab}^L \tilde{L}_a \tilde{E}_b^c H_1 + h.c. \right), \tag{11}
 \end{aligned}$$

where  $\varphi_i$  are the scalar fields,  $\tilde{\lambda}_{\alpha}$  are the gaugino fields,  $\tilde{Q}, \tilde{U}, \tilde{D}$  and  $\tilde{L}, \tilde{E}$  are the squark and slepton fields, respectively, and  $H_{u,d}$  are the Higgs fields. In principle there can be over 100 soft breaking terms, but they are constrained by requiring the absence of Flavour Changing Neutral Currents (FCNCs) and CP violation. These constraints appear naturally if one imposes “universal” soft breaking terms, which means that all the soft scalars are proportional to the identity matrix  $\mathbb{1}_{3 \times 3}$ , the trilinear  $A$  terms are proportional to the Yukawa couplings, and there are no extra CP violating phases besides the usual CKM one

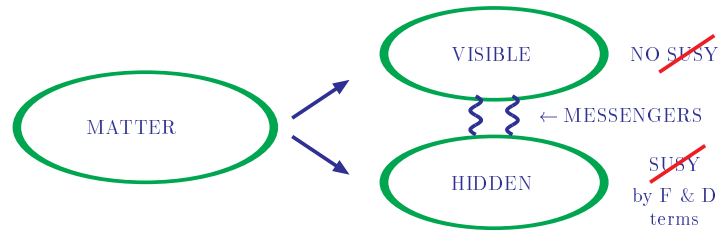
$$\tilde{m}_{Q,u,d,L,e}^2 \propto \mathbb{1}_{Q,u,d,L,e}; \quad a_{u,d,e} \propto A_{u,d,e} Y_{u,d,e}. \tag{12}$$

Usually, universal gaugino masses at the unification scale are assumed too. These universality conditions are usually taken as boundary conditions for the RGEs at a higher scale, most commonly the GUT scale. They are assumed to originate from an underlying more fundamental theory, where SUSY is dynamically broken, although such theory is not known. The general assumption is that the MSSM is connected to an unknown or “hidden” sector, that communicates to the matter or “visible” sector through so-called messengers. Thus, the effective theory that is left after the hidden and visible sectors interact through the messenger has a Lagrangian with soft SUSY breaking terms. The most studied types of SUSY breaking through hidden sectors are:

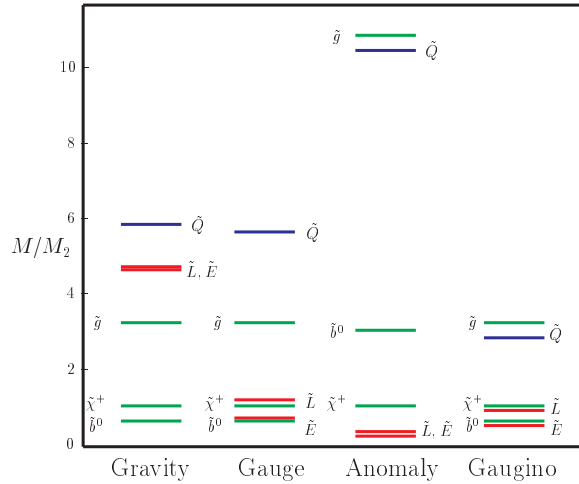
- Gravity mediated, the sectors interact through gravity. The SUSY breaking scale is of the order of the gravitino mass, the superpartner of the graviton. Although it is very attractive due to the fact that gravity exists and is felt by all particles, there does not exist yet a theory of quantum gravity.
- Gauge mediated, the SUSY breaking is communicated to the observable sector through the known gauge interactions that involve the messenger particles in loop diagrams. These messenger particles are the SM particles plus particles from a unified theory. The LSP is the gravitino.
- Anomaly mediated, SUSY breaking appears at loop level through a superconformal anomaly. An interesting feature in this scenario is that the soft terms that appear are renormalization group invariant, i.e. they are valid at any renormalization scale. One drawback of the simplest anomaly mediated scenario is that the slepton is tachyonic, although there are ways out of this problem.
- Gaugino mediated, it is based on brane theory. It is assumed that the SM lives on a brane, while gravity and other fields propagate in the bulk. SUSY breaking happens in a different brane and it is communicated to our brane through the gauginos. This is the least studied of the four hidden sector proposals.

All these scenarios generate soft breaking terms of the form given in Eq. (11). To each one of these scenarios correspond particular boundary conditions at the unification scale, that after evolving to the electroweak scale through the RGEs, give different predictions for the s-spectrum. Examples of such different s-spectra are shown in Fig. 3.





SPARTICLE SPECTRA



**Fig. 3:** The upper part shows the hidden sector scenario, whereas the lower part shows the different particle-spectra that come from the different scenarios, both figures from ref [21].

### 5.3 Higgs potential and masses in the MSSM

The tree level Higgs potential of the MSSM should be compatible also with electroweak symmetry breaking. We require that it breaks the  $SU(2) \times U(1)$  group to the electromagnetic  $U(1)_{EM}$ . At the minimum it is always possible to rotate away the vev's of one of the Higgs fields, for instance  $H_u^+ = 0$ , which at the minimum  $\partial V / \partial H_u^+$  implies that also  $H_d^- = 0$ . Thus, the potential is given by

$$\begin{aligned}
 V_{tree}(H_u, H_d) &= (|\mu^2| + m_{H_u}^2)|H_u|^2 + (|\mu^2| + m_{H_d}^2)|H_d|^2 - b(H_u^0 H_d^0 + c.c.) \\
 &+ \frac{g^2 + g'^2}{8} (|H_u^0|^2 - |H_d^0|^2)^2
 \end{aligned} \tag{13}$$

where  $g$  is the  $SU(2)$  gauge coupling and  $g'$  the  $U(1)$  one,  $m_{H_u}^2$  and  $m_{H_d}^2$  are the soft breaking mass terms for the Higgs bosons, and  $\mu$  their mixing term in the superpotential. The quartic couplings are fixed in terms of the gauge couplings. The requirement that the potential is bounded from below gives positivity conditions on the potential parameters

$$m_1^2 + m_2^2 > 2|b|. \tag{14}$$

To guarantee electroweak symmetry breaking, a linear combination of the Higgs fields has to have a negative square mass term, which happens if

$$m_3^2 > m_1^2 m_2^2. \tag{15}$$

These conditions are not satisfied at the GUT scale in general, but since they are scale dependent, i.e. “running parameters”, they might be realized at lower energies after evolving the RGEs. Thus, the symmetry breaking is driven through radiative corrections. Because of this feature the mechanism is known as radiative electroweak symmetry breaking. Furthermore, after electroweak breaking the Higgs vevs and the gauge couplings must be related to the  $Z$  boson mass through

$$m_Z^2 = (g^2 + g'^2)(v_u^2 + v_d^2)/4, \quad (16)$$

where

$$\langle H_u^0 \rangle = v_u/\sqrt{2}, \quad \langle H_d^0 \rangle = v_d/\sqrt{2}, \quad v = v_u^2 + v_d^2 \simeq 246.218 \text{ GeV},$$

and we define  $\tan \beta = v_u/v_d$ . Eq. (16) imposes further conditions on the  $\mu$  and  $b$  parameters

$$b = \frac{(m_{H_d}^2 - m_{H_u}^2) \tan 2\beta + M_Z^2 \sin 2\beta}{2}, \quad (17)$$

$$\mu^2 = \frac{m_{H_u}^2 \sin^2 \beta - m_{H_d}^2 \cos^2 \beta}{2} - \frac{M_Z^2}{2}. \quad (18)$$

To find the Higgs bosons masses, it is necessary to expand the potential around the minimum and separate it in real and imaginary parts. The two complex Higgs doublets have eight real degrees of freedom. The real parts correspond to the CP-even Higgs bosons and the imaginary parts to the CP-odd ones and to the Goldstone bosons. After electroweak symmetry breaking the three Goldstone bosons become the longitudinal modes of the  $W^\pm$  and  $Z$  bosons, as in the SM. The other five degrees of freedom will give rise to four massive scalars, two of which are neutral  $h$  and  $H$ , two charged  $H^\pm$ , plus a massive neutral pseudoscalar  $A$ . Of these, one can be easily identified with the SM Higgs boson, namely  $h$ , which is naturally lighter than the other four.

After SUSY and electroweak symmetry breaking all the SM, Higgs bosons and the s-particles acquire their physical masses. The tree level masses of the Higgs bosons are:

$$M_Z = 2b/\sin 2\beta \quad (19)$$

$$M_{h,H} = \frac{1}{2}(M_A^2 + M_Z^2 \mp \sqrt{(M_A^2 + M_Z^2)^2 - 4M_A^2 M_Z^2 \cos 2\beta}) \quad (20)$$

$$M_{H^\pm} = M_A^2 + M_W^2 \quad (21)$$

It is clear from the minimisation conditions and the expressions for the Higgs' masses that there are only two free parameters at tree level, which can be taken as  $M_A$  and  $\tan \beta$ . Moreover, whereas  $M_A, M_H$ , and  $M_{H^\pm}$  can be very large, the lightest Higgs mass  $m_h$  is bounded from above

$$M_h \leq \min(M_A, M_Z) |\cos 2\beta| \leq M_Z. \quad (22)$$

Radiative corrections to  $M_h$ , particularly from the RGE running quartic coupling, which is proportional to the top quark mass, and threshold corrections from integrating out the stop quarks in the loop, are very large and change this bound to

$$M_h \lesssim 135 \text{ GeV}. \quad (23)$$

The actual value of  $M_h$  in any specific model will depend on the s-spectrum, particularly on the stop quark masses, and the degree of mixing among themselves. This result is encouraging for the MSSM as an extension of the SM in the sense that it has a natural decoupling limit, which reduces to the SM, but the details of the predictions depend sensibly on the details of the soft breaking terms.

Another interesting feature of the MSSM is that after electroweak symmetry breaking the neutral states can mix among themselves, the same happens to the charged states. This gives rise to four mass

eigenstates known as neutralinos  $\chi_{1,2,3,4}^0$  which are a mixture of the neutral gauginos ( $\tilde{W}^0, \tilde{B}$ ) and neutral Higgsinos ( $\tilde{H}_u^0, \tilde{H}_d^0$ ); and two charged mass eigenstates, the charginos  $\chi_{1,2}^\pm$  which are a mixture of the charged Higgsinos ( $\tilde{H}_u^\pm, \tilde{H}_d^\pm$ ) and the charged Winos ( $\tilde{W}^\pm$ ).

The mass matrices of the neutralinos are

$$M^{(0)} = \begin{pmatrix} M_1 & 0 & -M_Z \cos \beta \sin_W & M_Z \sin \beta \sin_W \\ 0 & M_2 & M_Z \cos \beta \cos_W & -M_Z \sin \beta \cos_W \\ -M_Z \cos \beta \sin_W & M_Z \cos \beta \cos_W & 0 & -\mu \\ M_Z \sin \beta \sin_W & -M_Z \sin \beta \cos_W & -\mu & 0 \end{pmatrix}, \quad (24)$$

the mass matrices of the charginos are

$$M^{(c)} = \begin{pmatrix} M_2 & \sqrt{2}M_W \sin \beta \\ \sqrt{2}M_W \cos \beta & \mu \end{pmatrix}, \quad (25)$$

and the mass matrices for the scalar quarks are of the form

$$\begin{pmatrix} \tilde{m}_{fL}^2 & m_t(A_f - \mu f_{ac}\beta) \\ m_f(A_f - \mu f_{ac}\beta) & \tilde{m}_{fR}^2 \end{pmatrix},$$

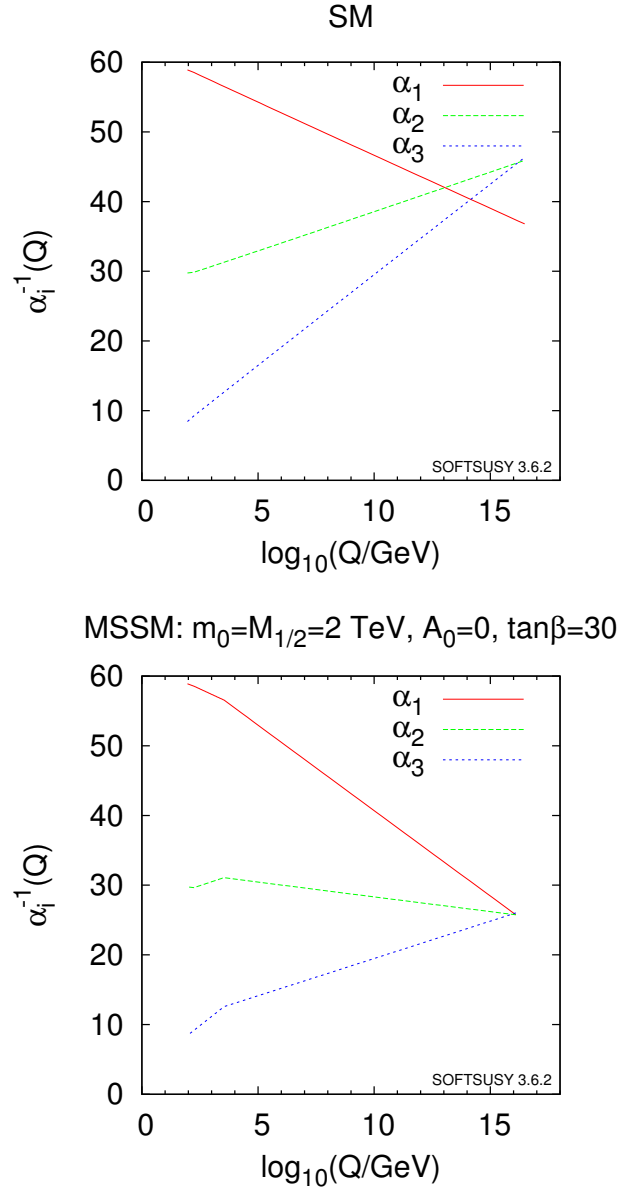
where  $f$  is any of the quarks and leptons, and  $f_{ac}\beta = \tan \beta$  for down type quarks and leptons, and  $f_{ac}\beta = \cot \beta$  for up type quarks. The mixing in the first two families is practically negligible, but can be sizeable in the third one.

As can be seen from the expressions of the mass matrices, the sparticle spectrum (s-spectrum) depends on the choice of soft breaking parameters, given in Eq. (11) and the  $\mu$  Higgsino mixing term appearing in the superpotential (9). Thus, different soft breaking terms will lead to different allowed parameter space at low energies.

To make the study of the MSSM more tractable the number of free parameters is reduced in different ways. One popular way is the constrained MSSM (cMSSM), in which the universality relations Eq. (12) are taken a step further, making all the soft scalar masses at a high scale equal,  $\tilde{m}_0$ , all the trilinear terms equal  $A$ , and all the gaugino masses equal  $m_{1/2}$ . This leaves only five free parameters:  $\tilde{m}_0, m_{1/2}, A, \tan \beta, \text{sign} \mu$ . The cMSSM is inspired in minimal supergravity models, which have the same five parameters, but with extra relations among themselves at the unification scale. Most experimental searches are based on the cMSSM or variations of it. This model might be too constrained and even unrealistic, but opening slightly the parameter space at the unification scale can widen it at experimental scales. For instance, relaxing any of the strict universality relations changes the phenomenology at low energies.

Besides giving a solution to the hierarchy problem, there are two other widely studied features that make the MSSM and other SUSY models attractive.

- The neutralino LSP has been proposed as a good candidate for dark matter, since it is neutral and has only electroweak interactions. The neutralinos are a combination of the neutral gauginos, which means that for different values of the soft breaking terms the LSP will be different. This in turn means that the admixture of the lightest neutralino (Higgsino, Wino and Bino) will lead to different types of DM (for an extensive review on the subject see ref. [22]). The combination of experimental collider searches and direct DM detection experiments can be used to restrict the allowed parameter space of a particular model. An example of this is provided in [23], where DM is assumed to be the lightest neutralino, and regions in parameter space that agree with the dark matter relic abundance  $\Omega$  and experimental constraints are studied for different SUSY models.
- The MSSM with  $\mu > 0$  might provide a solution to the discrepancy between the experimental and the theoretical value of the anomalous magnetic moment of the muon  $g - 2$  [24]. On the other hand, lepton flavour violating processes have been proposed also to alleviate or solve entirely this problem (see for instance [25, 26]).



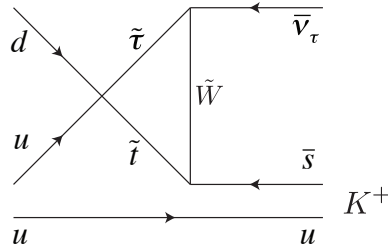
**Fig. 4:** Unification of couplings in the SM (up) and in the MSSM (bottom), where the SUSY breaking is assumed to be at  $2 \text{ TeV}$ . From ref. [27].

There are a number of low energy experimental constraints that reduce the parameter space for the cMSSM. The most stringent ones are the ones that come from FCNC's and flavour physics. Among the most used constraints are:

- The branching ratio  $b \rightarrow s\gamma$ .
- The branching ratio  $B \rightarrow \mu^+\mu^-$ .
- Constraints on direct searches on SUSY particles.
- Dark matter constraints, if the dark matter is composed 100% of the LSP, which is usually the neutralino.

## 5.4 SUSY GUTs

The unification of couplings, which was not good in the SM, turns out to work rather well in SUSY GUT models, see Fig. 4, assuming the SUSY breaking happens around a few TeV (for recent reviews see [28] and references therein). The idea is basically the same as in GUTs, but with supersymmetry added. The



**Fig. 5:** Dimension five operators that mediate proton decay in SUSY GUTs, from [29].

matter and gauge contents are assigned to the same irreps as in non-SUSY GUTs, but now they refer to the superfields, where the SM particles appear in the supermultiplets together with their superpartners.

In SUSY GUTs there are more ways for the proton to decay through dimension five operators, see Fig. 5. These decay modes are the dominant ones, with  $\tau_p \sim 10^{34}$  yrs. The coloured Higgsino triplets in SUSY GUTs can give rise to dimension five operators also, thus they have to be heavier than the GUT scale to suppress proton decay.

Dimension six operators exist, like in ordinary GUTs, but since the unification scale is large  $M_{GUT} \sim 10^{16}$  the proton lifetime coming from these processes is  $\tau_p \sim 10^{35}$  yrs. Dimension four operators do not appear in the models considered here, since R parity is conserved.

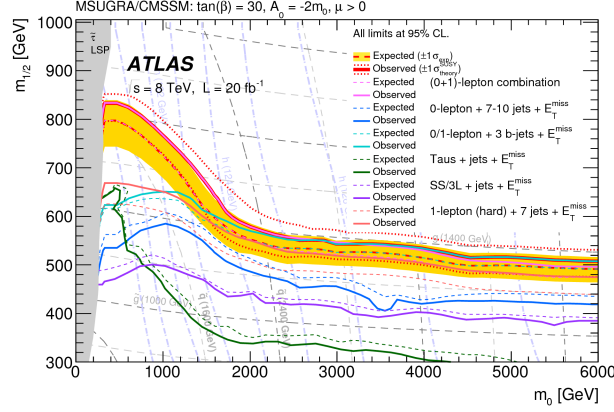
Many of the constraints on the parameter space in the MSSM come from the assumption that there is an underlying unified theory behind it. That is the origin of the hidden sector SUSY breaking assumption and of the universality conditions that were discussed in the previous subsection.

An intriguing feature of SUSY GUTs is that they can be made finite to all-loops in perturbation theory, which can lead to good predictions for the third generation of quark masses, the Higgs mass, and a prediction for a relatively heavy s-spectrum (see for instance [30] and references therein).

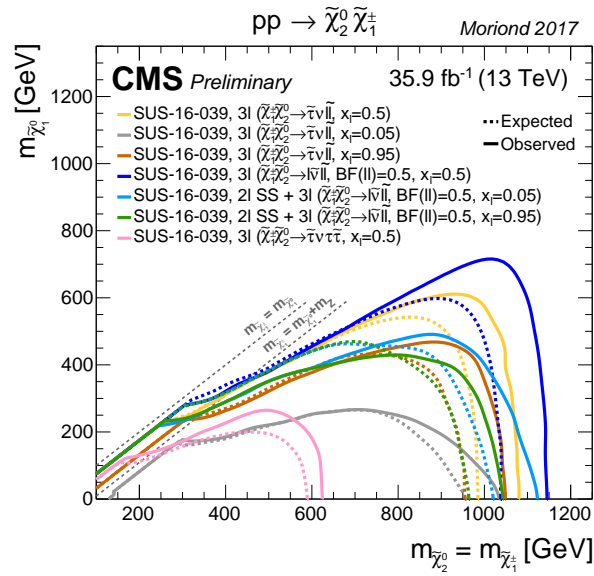
## 5.5 Experimental searches

No direct or indirect evidence for supersymmetry has been found so far. Indirect evidence could come from contributions at loop level to some rare processes in the SM, like  $b \rightarrow s\gamma$  or electron dipole moments, for instance. Direct production could happen through quark-antiquark annihilation, gluino fusion, gluino-quark interaction, and quark-quark scattering, which would lead to pairs of sparticles. The sparticles then would decay into SM particles and neutralinos, the latter escape detection carrying with them some missing energy. The LHC experiments, CMS and ATLAS, have put bounds on the SUSY particles, and excluded some regions of the parameter space of the mSUSM. In general, these analyses are done in simplified models that exhibit generic features, since it is difficult to test simultaneously all of the free parameters of the MSSM. We present two examples of how such searches are presented, for a more detailed description the reader should go to the experiments public pages [ ] and papers therein.

The results for the searches for gluinos and first and second generation squarks from proton-proton collisions at 8 TeV are presented in [31]. Limits in simplified models with gluinos and squarks of the first and second generations are derived for direct and one- or two-step decays of squarks and gluinos, and gluino decays via third-generation squarks are derived. They considered simplified models with R-parity conservation, and in all of them the limit on the gluino mass exceeds 1150 GeV at 95% CL, for an LSP mass smaller than 100 GeV. They also derive exclusion limits for a number of simplified



**Fig. 6:** Exclusion limits for MSUGRA /cMSSM models in the  $(m_0, m_{1/2})$  plane, from [31, 32].



**Fig. 7:** CMS exclusion limits in the chargino-lightest neutralino plane are show, from [33, 34]

models, like the cMSSM and a model with non-universal Higgs mass model with gaugino mediation, among others. Fig. 6 the exclusion limits at 95% CL for 8 TeV analyses are shown in the  $(m_0, m_{1/2})$  plane for the MSUGRA/cMSSM model. The other three parameters have been set to  $\tan(\beta) = 30$ ,  $A_0 = -2m_0$ ,  $\mu > 0$  [31]. The Higgs mass is  $\sim 125$  GeV in a large part of the parameter space. The search excludes gluino masses  $< 1280$  GeV.

The results of the searches for chargino-neutralino pair production in proton-proton collisions with sleptons as final decay modes were performed in [33]. Chargino-neutralino pair production is expected to have the largest cross section from all the electroweak processes. Higgsino pair production in a gauge mediated SUSY breaking scenario, was also studied. Different simplified models lead to different final states, where limits on the charginos and neutralino masses can be set, these range from 180 GeV to 1150 GeV at 95% C.L. For instance, models with light left-handed leptons lead to the stringest bounds, with the mass limit for charginos and neutralinos up to 1150 GeV at 95% C.L. In 7 exclusion limits for this study in the chargino/lightest neutralino plane are shown.

## 5.6 The Next to Minimal Supersymmetric Standard Model

Another widely studied extension of the SM is the Next to Minimal Standard Model (NMSSM). The model has the same matter content than the MSSM plus a chiral singlet superfield  $\hat{S}$ , i.e. an extra Higgs singlet. It provides for a solution of the so-called “ $\mu$  problem” in the MSSM, which refers to the fact that the  $\mu$  term in the Higgs potential is of the order of the SUSY breaking scale, while the fundamental scale is high (see for instance [35, 36] and references therein).

The NMSSM solves the  $\mu$  problem by generating this term dynamically through the new chiral singlet. The superpotential adds two terms to the MSSM usual one

$$W_{NMSSM} = W_{MSSM} + \lambda \hat{S} H^u H^d + \frac{\kappa}{3} \hat{S}^3 \quad (26)$$

and the Higgs potential is given by

$$V_{NMSSM} = m_{H_u}^2 |H_u|^2 + m_{H_d}^2 |H_d|^2 + m_S^2 |S|^2 + (\lambda A_\lambda S H_u H_d + \frac{\kappa}{3} A_\kappa \hat{S}^3 + h.c.). \quad (27)$$

When the new singlet field acquires a vev  $\langle S \rangle = v_S$ , it generates an effective  $\mu$  term  $\mu_{eff} = \lambda v_S$ , which appears then naturally at the electroweak scale. The NMSSM shares also the nice features of the MSSM in terms of candidates to dark matter and unification of couplings, besides solving the hierarchy problem.

The NMSSM has two more degrees of freedom than the MSSM, and after electroweak symmetry breaking there are seven physical massive scalars. If there is no extra CP breaking these states are three neutral CP even scalars  $H_{1,2,3}$ , two charged ones  $H^\pm$ , and two neutral pseudoscalars  $A_{[1,2]}$ . The SM-like Higgs boson can be either the lightest scalar, or the second lightest scalar. In the latter it is possible to have a very light neutral Higgs that has escaped detection so far, although this situation is very constrained by LHC searches of scalars decaying into  $\tau$  pairs, as well as by flavour observables. The heavier one is very similar to the heavy neutral Higgs of the MSSM.

In the NMSSM there is no upper bound for the tree-level mass, as in the MSSM, where large stop masses or large mass splittings are necessary to lift the Higgs mass through radiative corrections. Thus, the NMSSM can achieve more naturally a Higgs mass of 125 GeV, and still retain the good features of the MSSM: solution to the hierarchy problem, good unification of couplings, and good dark matter candidates. Also, in the NMSSM, the extended Higgs sector allows for Higgs to Higgs decays that are not present in the MSSM. The results for LHC searches for light Higgses, i.e. with mass below 125 GeV in the diphoton channel are presented in refs. [37, 38], and a review on the LHC NMSSM Higgs boson searches with emphasis on the mono-Higgs signature in [39].

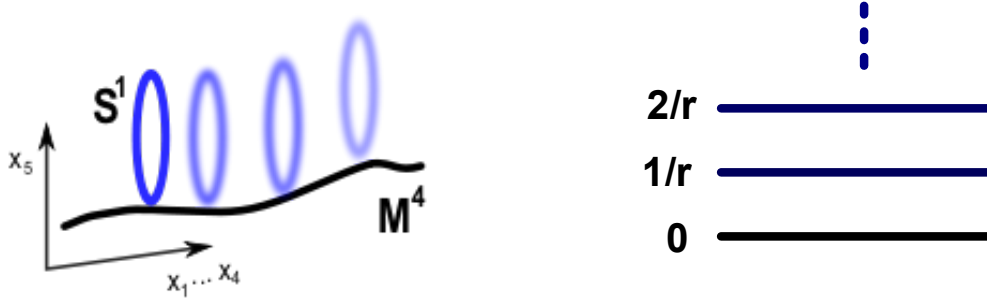
## 6 Extra Dimensions

Another way to go beyond the SM is to add extra space-time dimensions, namely space ones. The idea of having extra space dimensions predates the SM by around half a century. It was proposed first by Niels Bohr in 1914 to unify electromagnetism and scalar gravity. Independently T. Kaluza proposed to extend general relativity in five dimensions (1921), and then O. Klein proposed to compactify the extra dimension in 1926. Thus it is possible to describe four-dimensional gravity and electromagnetism. Although this original proposal does not work, it is the basis for modern superstring theories and the brane world scenario. For reviews on extra dimensions see for instance [13, 40].

Consider first that the fifth dimension is compactified in a circle of radius R, and the other dimensions are extended, see Fig. fig:KK-circle

The action of a massless field in five dimensions is

$$S_5 = \int d^5x \frac{1}{2} \partial_M \phi(x^\mu, y) \partial^M \phi(x^\mu, y), \quad (28)$$



**Fig. 8:** A string compactification of one dimension in a circle  $S^1$  times four dimensional Minkowski space  $M^4$  is shown to the left, to the right is shown the associated tower of massive Kaluza-Klein modes  $m_n = |n|/r$ , both figures from ref. [11].

where  $x^M = (x^\mu, y)$ ,  $M = 1, \dots, 5$ ,  $\mu = 0, \dots, 4$  and  $y$  is the coordinate in the fifth dimension, which is compact. This means that it has periodic boundary conditions of the form

$$\phi(x^\mu, y) = \phi(x^\mu, y + 2\pi R).$$

Since it is periodic in  $y$  it is possible to make a Fourier expansion

$$\phi(x^\mu, y) = \phi_n(x^\mu) \exp(iny/R), \quad (29)$$

upon substituting the Fourier expansion in the 5-D Eq. (28) action we obtain

$$S_4 = \int d^4x \left( \partial_\mu \phi^0 \partial^\mu \phi^0 + \sum_n \left( \partial_\mu \phi^{n\dagger} \partial^\mu \phi^n - \frac{n^2}{R^2} \phi^{n\dagger} \phi^n \right) \right). \quad (30)$$

The last term in the action represents an infinite tower of massive particles  $m_n = n/R$ . These are the so-called Kaluza-Klein modes, and a signature of theories with compactified extra dimensions.

The 5-dim action is given by

$$S_{5D} = \int d^5x \frac{1}{g_{5D}^2} F_{MN} F^{MN}, \quad (31)$$

where

$$F_{MN} = d_M A_N - d_N A_M. \quad (32)$$

Then, decomposing again in four dimensional fields we have the usual vector field  $A_\mu$ , and a scalar  $\rho = A_4$ . In a similar way as the scalar one, it is possible to expand in a Fourier series along the compact dimension the gauge fields and find an expression for the 4-dim action,

$$A_\mu = \sum_{n=-\infty}^{\infty} A_\mu^n \exp\left(\frac{iny}{R}\right); \quad \rho = \sum_{n=-\infty}^{\infty} \rho_n \exp\left(\frac{iny}{R}\right). \quad (33)$$

Choosing a traverse gauge to remove the mixed terms

$$\partial^M A_M = 0, \quad A_0 = 0 \quad \rightarrow \quad \partial^M \partial_M A_N = 0$$

leads to

$$S_4^{gauge} = \int d^4x \left( \frac{2\pi R^2}{g_{5D}} F_0^{\mu\nu} F_{0\mu\nu} + \frac{2\pi R}{g_{5D}^2} \partial_\mu \rho_0 \partial^\mu \rho_0 + \dots \right) \quad (34)$$



This action corresponds to a massless gauge particle, a massless scalar, plus infinite towers of massive scalar and vector fields. The gauge couplings of the 5-dim and the 4-dim theory are related through

$$\frac{1}{g_4^2} = \frac{2\pi R}{g_5^2}.$$

In the gravity sector, we denote the graviton as  $G_{MN}$ , where  $G_{\mu\nu}$  is the graviton,  $G_{\mu n}$  gravivectors, and  $G_{mn}$  graviscalars. After a similar treatment to the gauge and vector fields, the 4-dim Einstein-Hilbert action reads

$$S_{4D}^{EH} = \int d^4x \sqrt{|g|} \left( M_1^2 {}^{(4)}R - \frac{1}{4} \phi^{(0)} F_{\mu\nu}^{(0)} F^{(0)\mu\nu} + \frac{1}{6} \frac{\partial^\mu \phi^{(0)} \partial_\mu \phi^{(0)}}{(\phi^{(0)})^2} + \dots \right) \quad (35)$$

This way gravity, scalar fields, and electromagnetism get unified, but the Planck mass is not fundamental, it is a derived quantity

$$M_{Pl}^2 \sim M_D^{D-2} V_{D-4} \sim M_D^{D-2} R^{D-4}. \quad (36)$$

In order to include non-Abelian gauge fields in this kind of formalism one needs to go to higher dimensions, it is not possible to include strong and weak interactions only in five dimensions. Kaluza-Klein theories are the inspiration for more modern theories like superstrings and brane worlds, as well as universal extra dimensions models.

## 6.1 Brane Worlds

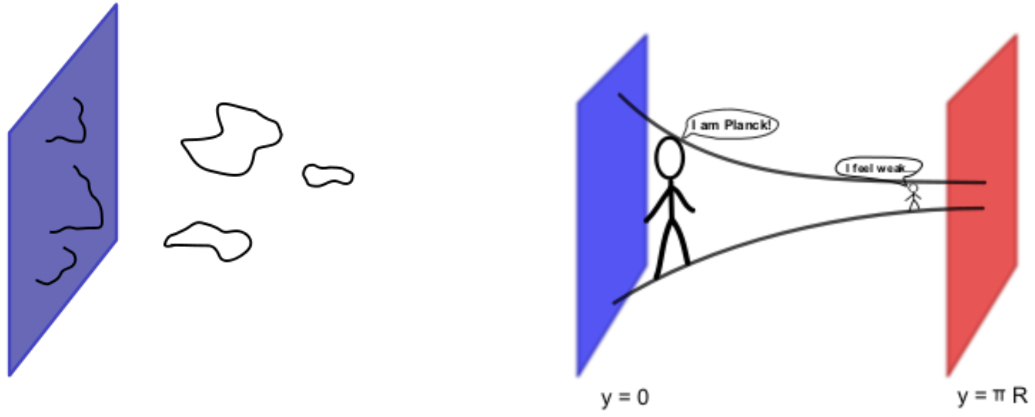
String theory was formulated originally as a theory of the strong interaction, but it turned out to provide a description of quantum gravity, with the realisation that it contains a massless spin 2 state which can be identified with the graviton. There is actually an infinite number of excitations of the strings, and the massless ones are interpreted as non-Abelian gauge bosons and matter fermions. At large distances these excitations appear like point-like objects, whose properties, like mass and charge depend on the vibrational modes of the string.

In superstring theory, instead of point particles the fundamental objects are one-dimensional: either open or closed strings. The mathematical consistency requires that the theory is supersymmetric and formulated in six extra space dimensions. Usually, these extra dimensions are assumed to be compactified with extremely small compactification radii, of the order of the Planck scale, and that is why they would be unobservable.

String theory is not formulated as a quantum field theory, it is a geometrical theory where particles and interactions, upon compactification of the extra dimensions it may lead to a quantum field theory, many different vacua appear as possibilities, the number can be as large as  $10^{500}$ . This makes it very difficult to find the vacuum where the SM could come from, since the fundamental scale of string theory is the Planck scale, and we cannot make experiments that reach that scale in order to probe the different possibilities.

Another possibility, inspired in string theory and extra dimensions, was the realization that different fields might live in different dimensions. Then, the SM fields would be described by open strings whose ends are attached to a brane, which is a 3-dimensional hypersurface, whereas the gravitons are described by closed strings that can travel through all D dimensions. For reviews on string theory, branes, and extra dimensions see for instance [41, 42].

The following scenarios have been proposed as possible solutions to the hierarchy problem, where the hierarchy problem has been now turned into the problem of finding the number and size of the extra dimensions.



**Fig. 9:** The left figure depicts a brane, where the SM fields are open strings that start and end on the brane, whereas gravitons are closed strings moving in the bulk. The right figure depicts the Randall-SUndrum setup, with the Planck brane on the left, warped down to the weak brane in the right, both figures from [11].

## 6.2 Large extra dimensions

Large extra dimensions have been proposed as a solution, or a way to go around, the hierarchy problem. Rather than asking why the electroweak scale is so much smaller than the Planck scale, the question is why is gravity so much weaker than the other interactions. The idea is that the Planck scale is not the fundamental one, but it appears so because of the existence of extra dimensions, it is a derived quantity. The actual fundamental scale is around  $\mathcal{O}(1) TeV$ . From Eq. (36) we can figure out the number and size of the extra dimensions needed. From experimental tests the radius has to be smaller than  $R \geq 1 TeV$ , since no K-K tower or its effects have been found. For  $d$  extra dimensions,

$$d = 1 \quad R \sim 10^9 km \quad (37)$$

$$d = 2 \quad R \sim 0.5mm \quad (38)$$

$$d = 3 \quad R \sim 1^{-6}cm \quad (39)$$

the first possibility,  $d = 1$  is clearly ruled out,  $d = 2$  is currently being tested, and larger values of  $d$  are still allowed.

## 6.3 Warped extra dimensions

The warped extra dimension or Randall-Sundrum (RS) simplest scenario consists also of a five dimensional theory, which is an interval bounded by two three dimensional branes. This interval has a warped geometry, that is the metric is exponentially warped along the  $y$  direction

$$ds^2 = \exp(\kappa y) \eta_{\mu\nu} dx^\mu dx^\nu + dy^2. \quad (40)$$

In this scenario the Planck scale is the fundamental one, and sits in one of the branes at  $y = 0$ , the other is the SM brane at  $y = \pi R$ . So, from Eq. (40), we can see that the metric changes exponentially as  $\eta_{\mu\nu} \rightarrow \exp(-\kappa\pi R) \eta_{\mu\nu}$  from the SM to the Planck brane. The change in the metric implies a change in energy and length scales, for the electroweak scale this implies

$$\Lambda_{Ew} \sim M_{Pl} \exp(\pi\kappa R) \sim 1 TeV, \quad (41)$$

with a small dimension  $R \gtrsim 50l_{Pl}$ , just slightly bigger than the Planck length. In this scenario  $d = 1$ , i.e. only one extra dimension, is still allowed.

A very attractive feature of the Randall-Sundrum scenario is that it can be tested experimentally. The interaction Lagrangian is

$$\mathcal{L}_I = -G^{\mu\nu} T_{\mu\nu} / \Lambda_{Ew} , \quad (42)$$

where  $T_{\mu\nu}$  is the energy-stress tensor, which involves the SM fields. These interactions are at the order of the electroweak scale, and in principle can be produced at the LHC. If only gravity travels through the warped extra dimension the first resonance will be the RS graviton. In more modern variants of the RS scenario, it is necessary to allow more particles to travel in the bulk, but leaving the Higgs boson attached at the SM brane, in order to avoid FCNCs.

## 7 Multi-Higgs Models

It is possible to extend the SM by just adding more particles, but preserving the gauge group of the SM. This is the case of some multi-Higgs models (see for instance [43, 44] and references therein). In general, adding  $N$  Higgs doublets to a Lagrangian leads to the Higgs potential of an  $N$ -Higgs doublet model (NHDM),

$$V(\phi) = Y_{ij} \phi_i^\dagger \phi_j + Z_{ijkl} (\phi_i^\dagger \phi_j) (\phi_k^\dagger \phi_l) , \quad (43)$$

which is Hermitian, thus  $Y_{ij} = Y_{ji}^*$ ,  $Z_{ijkl} = Z_{jilk}^*$ ,  $Z_{ijkl} = Z_{kijl}$ . For  $N$  Higgs complex doublets, each with four degrees of freedom, the potential Eq. (43) has  $N^2 + N^2(N^2 + 1)/2$  real parameters. So, for a general two Higgs doublet model there are 14 real parameters, as compared to the Higgs potential of the MSSM, which has only four. For three Higgs doublet models there are in principle 54 real parameters. Because of the increase in the number of parameters, phenomenological studies of NHDM have mainly focused on two and three Higgs doublet models.

By far the most widely studied NHDM are two Higgs doublet models (2HDM) [43]. The Higgs potential for a general 2HDM is

$$\begin{aligned} V = & m_{11} \phi_1^\dagger \phi_1 + m_{22} \phi_2^\dagger \phi_2 - (m_{12} \phi_1^\dagger \phi_2 + h.c.) + \\ & \frac{\lambda_1}{2} (\phi_1^\dagger \phi_1)^2 + \frac{\lambda_2}{2} (\phi_2^\dagger \phi_2)^2 + \lambda_3 (\phi_1^\dagger \phi_1) (\phi_2^\dagger \phi_2) + \lambda_4 (\phi_1^\dagger \phi_2) (\phi_2^\dagger \phi_1) \\ & + \left[ \frac{\lambda_5}{2} (\phi_1^\dagger \phi_2)^2 + \lambda_6 (\phi_1^\dagger \phi_1) (\phi_1^\dagger \phi_2) + \lambda_7 (\phi_2^\dagger \phi_2) (\phi_1^\dagger \phi_2) + h.c. \right] , \end{aligned} \quad (44)$$

where  $m_{11}, m_{22}, \lambda_{1234}$  are real, and the rest  $m_{12}, \lambda_{5,6,7}$  are complex. These are 14 parameters, but only eleven of which are physical. as can be seen by a change of basis. This general potential allows for charge breaking minima, which is usually avoided, as well as CP conserving and CP violating minima. The possibility of CP violating minima is quite interesting, since it allows for new sources of CP violation to address baryogenesis, and many models with this feature have been studied.

One of the main challenges of any NHDM is to avoid FCNCs. In the case of 2HDM there are two models that naturally avoid this problem. Type I 2HDM couples only one of the Higgs fields to the quarks, by convention it is taken to be  $\phi_2$ . This can be achieved by requiring a  $Z_2$  symmetry that acts like  $\phi_1 \rightarrow -\phi_1$ . The type II 2HDM is similar to the MSSM in that the right-handed up-type quarks couple to one of the Higgs fields, and the right-handed bottom-type quark to the other. In this case, besides the  $\phi_1 \rightarrow -\phi_1$  discrete symmetry, it is required that  $d_R^i \rightarrow -d_R^i$ . In both models it is assumed the right-handed leptons couple to the Higgs field in the same way as the right-handed down quarks. From these two models, type II has been more studied, due to its similarities with the MSSM. There are variations of these two models, like the Lepton-specific or the flipped one, which will not be discussed here. In the Higgs sector, the  $Z_2$  symmetry means there is no CP violation. Thus, in some models a term like  $m_{12}^2 (\phi_1^\dagger \phi_2)$  is added to allow for CP violating terms. In type III models both Higgs fields are allowed to couple to the matter sector, and in this case care has to be taken to avoid FCNCs.

In general, phenomenological studies of 2HDMs usually make a number of simplifying assumptions, to get rid of some of the parameters. The Higgs sector is assumed to be CP conserving, no

explicit and no spontaneous breaking is allowed, quartic terms odd in any of the doublets are zero (can be achieved with a discrete symmetry). To exemplify this, let's look at the following potential, which is gauge invariant, CP conserving, and that includes a term that breaks softly the symmetries, to allow for more diverse possibilities

$$\begin{aligned}
V = & m_{11}\phi_1^\dagger\phi_1 + m_{22}\phi_2^\dagger\phi_2 - m_{12}(\phi_1^\dagger\phi_2 + \phi_2^\dagger\phi_1) + \\
& \frac{\lambda_1}{2}(\phi_1^\dagger\phi_1)^2 + \frac{\lambda_2}{2}(\phi_2^\dagger\phi_2)^2 + \lambda_3(\phi_1^\dagger\phi_1)(\phi_2^\dagger\phi_2) + \lambda_4(\phi_1^\dagger\phi_2)(\phi_2^\dagger\phi_1) \\
& + \frac{\lambda_5}{2} \left[ (\phi_1^\dagger\phi_2)^2 + (\phi_2^\dagger\phi_1)^2 \right] .
\end{aligned} \tag{45}$$

The procedure to minimise the potential follows the one of the MSSM, where the fields have to be expanded in terms of their neutral and charged parts, and requiring that the minimisation gives the correct electroweak breaking. In a similar way, one has to check that the potential minimum is really bounded from below, and that perturbative unitarity is maintained.

The Higgs fields are denoted by

$$\langle \phi_i \rangle = \begin{pmatrix} \varphi_i^+ \\ \frac{v_i + \rho_i + \eta_i}{\sqrt{2}} \end{pmatrix} \tag{46}$$

with  $j = 1, 2$ .

At the minimum the Higgs fields are

$$\langle \phi_1 \rangle_0 = \begin{pmatrix} 0 \\ \frac{v_1}{\sqrt{2}} \end{pmatrix} \quad \langle \phi_2 \rangle_0 = \begin{pmatrix} 0 \\ \frac{v_2}{\sqrt{2}} \end{pmatrix} \tag{47}$$

In the CP conserving model, CP-odd and CP-even states do not mix, and the term which will become the pseudoscalar decouples from the mass matrix. The CP neutral states mix to give two neutral states, the mixing angle is denoted by  $\alpha$

$$h = \sin \alpha \rho_1 + \cos \alpha \rho_2 , \tag{48}$$

$$H = -\cos \alpha \rho_1 + \sin \alpha \rho_2 . \tag{49}$$

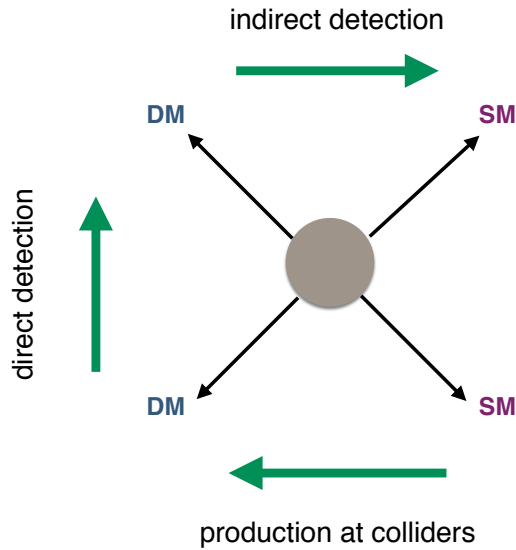
As in the MSSM the ratio of the vevs  $\tan \beta = v_2/v_1$  is an important parameter, since it is the mixing angle for the charged scalars, which are given by

$$H^\pm = -\sin \beta \varphi_1^\pm + \cos \beta \varphi_2^\pm , \tag{50}$$

where  $v_1 = v \cos \beta$  and  $v_2 = v \sin \beta$ . After electroweak symmetry breaking there are 5 massive scalar bosons,  $h, H, H^\pm$ , and  $A$ . From the two neutral ones,  $h$  is assumed to be the lightest. In a general 2HDM, the five masses, together with the two angles  $\alpha$  and  $\beta$  are free parameters.

In general, it is hard to minimise potential Eq. (45), so often the strategy is to fix the vevs, and then fix the mass parameters  $m_{ij}$  through the tadpole equations. In this kind of models, one has to be aware that there might be more than one neutral minima, and it is difficult to find the deeper one. It is possible to find whether the studied minimum is the deeper one by calculating a discriminant. This quantity is built from a combination of the potential parameters and vevs, a positive sign means that the potential is stable at tree-level, and a negative sign that there is a deeper minimum.

Because of the increased number of free parameters it is common to add more symmetries to multi-Higgs models in general, and to the 2HDMs in particular, to be able to do phenomenological studies.



**Fig. 10:** The figure shows possible ways of detecting dark matter.

### 7.1 Searches for exotic Higgs scalars

To test multi-Higgs models one can look directly to the production of the extra scalars, decays that are not possible in the SM or deviations from processes of the SM.

Again, searches have been focused mainly on two Higgs doublet models. The 2HDM has a natural decoupling limit, where the massive exotic Higgs states are much heavier than the SM one. This makes the type I 2HDM very difficult to distinguish phenomenologically from the SM, unless we go to high energies.

The main way to distinguish different 2HDMs from the SM and among themselves, is through the branching ratios of their Higgs decays (see for instance [43, 44] and references therein). These studies are usually done assuming the decoupling limit, where the extra Higgs bosons are very heavy and thus decoupled from the SM, although there are also searches for low mass Higgs scalars and pseudoscalars [45–47]. What can also be measured, besides the branching ratios, are corrections to the different SM couplings due to the heavy states. Higgs production has also been extensively studied in 2HDMs.

There are experimental bounds which are generic to 2HDM, for instance the couplings to gauge bosons like  $ZHA$  or  $\gamma H^+ H_-$ , which appear already at tree level and would have been already produced at the LHC, imply that the charged scalars are heavy. The same as with SUSY models, flavour observables can place stringent bounds on 2HDMs. The decay  $b \rightarrow s\gamma$  puts a bound on the charged Higgs mass  $m_{H^{\pm\pm}} > 480 \text{ GeV}$  at 95% C.L. [48], and in type II models flavour physics puts a bound of  $m_{H^{\pm\pm}} \gtrsim 600 \text{ GeV}$  [49].

### 7.2 Scalar Dark Matter

Extra Higgs scalars can provide also with good candidates to dark matter. They appear in multi-Higgs models with global symmetries, under which the SM particles do not transform. Of particular interest are the inert models, in which at least one extra Higgs doublet is added, with no couplings to the matter fields and with zero vacuum expectation value. This is achieved by adding an extra discrete symmetry, usually  $\mathbb{Z}_2$ . The combination of the symmetry and the zero vev guarantee the stability of the dark matter

scalar. The new heavier scalars decay into the lightest one, which is the DM candidate, which cannot decay further. This type of models are referred to as inert doublet models [50, 51].

The inert 2HDM has been extensively studied, since it is very predictive and can be tested in colliders, as well as in direct and indirect DM searches. The extra scalars can still have pair interactions among themselves, and also pair interactions with the gauge and Higgs bosons of the SM are possible. It also has cosmological consequences, in particular a sequence of strong first order phase transitions in the early Universe, which restricts considerably the allowed parameter space (see for instance [52] and references therein).

The searches for dark matter, be it dark scalars, neutralinos or other candidates, will go on via direct production from DM particles by collision of SM ones in colliders, indirect detection which can be DM annihilation or decays, or direct detection through collision with a target nucleus, as depicted in Fig. 10 (see for instance [4, 5] and references therein).

A review on searches for new particles excluding supersymmetry, for instance leptoquarks, higgses, heavy leptons, can be found in ref [53].

## 8 Conclusions

There is experimental and observational evidence that there is physics beyond the Standard Model in the neutrino masses and the existence of dark matter. The need to go beyond the SM is backed also by sound theoretical motivations. We have not found yet any other evidence of new, or rather unknown to us, physics beyond the Standard Model, and it is not for want of models or experimental searches. Since physics BSM impacts particle physics and cosmology, as well as some astrophysical processes, the search has to continue in all three fronts, both from the theoretical as from the experimental point of view. The continuous feedback between these three areas will guide our future searches.

## Acknowledgements

I thank the organizers of CLASHEP 2017 for the warm hospitality and the students for their attention. This work was partly supported by DGAPA-UNAM project PAPIIT IN111518.

## References

- [1] E. Noether, *Gott. Nachr.* **1918**, 235 (1918), arXiv:physics/0503066, [Transp. Theory Statist. Phys.1,186(1971)].
- [2] R. N. Mohapatra, Seesaw mechanism and its implications, in *Seesaw mechanism. Proceedings, International Conference, SEESAW25, Paris, France, June 10-11, 2004*, pp. 29–44, 2004, arXiv:hep-ph/0412379.
- [3] P. Hernandez, Neutrino Physics, in *Proceedings, 8th CERN-Latin-American School of High-Energy Physics (CLASHEP2015): Ibarra, Ecuador, March 05-17, 2015*, pp. 85–142, 2016, arXiv:1708.01046.
- [4] K. Garrett and G. Duda, *Adv. Astron.* **2011**, 968283 (2011), arXiv:1006.2483.
- [5] M. Lisanti, Lectures on Dark Matter Physics, in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings (TASI 2015): Boulder, CO, USA, June 1-26, 2015*, pp. 399–446, 2017, arXiv:1603.03797.
- [6] C. Csikı and P. Tanedo, Beyond the Standard Model, in *Proceedings, 2013 European School of High-Energy Physics (ESHEP 2013): Paradfurdo, Hungary, June 5-18, 2013*, pp. 169–268, 2015, arXiv:1602.04228.
- [7] F. Ellinghaus, K. Hoepfner, and T. Ohl, Searches for Physics Beyond the Standard Model, in *The Large Hadron Collider: Harvest of Run 1*, edited by T. Schıner-Sadenius, pp. 463–513, 2015.
- [8] M. Dine, *Supersymmetry and String Theory* (Cambridge University Press, 2016).

- [9] R. Rosenfeld, Physics Beyond the Standard Model, in *Proceedings, 8th CERN-Latin-American School of High-Energy Physics (CLASHEP2015): Ibarra, Ecuador, March 05-17, 2015*, pp. 159–164, 2016, arXiv:1708.00800.
- [10] S. Willenbrock and C. Zhang, *Ann. Rev. Nucl. Part. Sci.* **64**, 83 (2014), arXiv:1401.0470.
- [11] B. C. Allanach, *Beyond the Standard Model Lectures for the 2016 European School of High-Energy Physics*, PhD thesis, Cambridge U., DAMTP, 2016, arXiv:1609.02015.
- [12] H. E. Logan, (2014), arXiv:1406.1786.
- [13] F. Quevedo, S. Krippendorff, and O. Schlotterer, (2010), arXiv:1011.1491.
- [14] G. G. Ross, *GRAND UNIFIED THEORIES* (, 1985).
- [15] R. N. Mohapatra, *UNIFICATION AND SUPERSYMMETRY. THE FRONTIERS OF QUARK - LEPTON PHYSICS* (Springer, Berlin, 1986).
- [16] M. F. Sohnius, *Phys. Rept.* **128**, 39 (1985).
- [17] S. P. Martin, (1997), arXiv:hep-ph/9709356, [Adv. Ser. Direct. High Energy Phys.18,1(1998)].
- [18] H. E. Haber and L. Stephenson Haskins, Supersymmetric Theory and Models, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Anticipating the Next Discoveries in Particle Physics (TASI 2016) Boulder, CO, USA, June 6-July 1, 2016*, 2017, arXiv:1712.05926.
- [19] S. R. Coleman and J. Mandula, *Phys. Rev.* **159**, 1251 (1967).
- [20] R. Haag, J. T. Lopuszanski, and M. Sohnius, *Nucl. Phys.* **B88**, 257 (1975), [,257(1974)].
- [21] D. I. Kazakov, Beyond the standard model: In search of supersymmetry, in *2000 European School of high-energy physics, Caramulo, Portugal, 20 Aug-2 Sep 2000: Proceedings*, pp. 125–199, 2000, arXiv:hep-ph/0012288.
- [22] G. Jungman, M. Kamionkowski, and K. Griest, *Phys. Rept.* **267**, 195 (1996), arXiv:hep-ph/9506380.
- [23] E. A. Bagnaschi *et al.*, *Eur. Phys. J.* **C75**, 500 (2015), arXiv:1508.01173.
- [24] F. Jegerlehner, *Springer Tracts Mod. Phys.* **274**, pp.1 (2017).
- [25] F. V. Flores-Baez, M. Gümez Bock, and M. Mondragün, *Eur. Phys. J.* **C76**, 561 (2016), arXiv:1512.00902.
- [26] M. Lindner, M. Platscher, and F. S. Queiroz, *Phys. Rept.* **731**, 1 (2018), arXiv:1610.06587.
- [27] Particle Data Group, C. Patrignani *et al.*, *Chin. Phys.* **C40**, 100001 (2016).
- [28] S. Raby, *Lect. Notes Phys.* **939**, 1 (2017).
- [29] J. L. Hewett *et al.*, Planning the Future of U.S. Particle Physics (Snowmass 2013): Chapter 2: Intensity Frontier, in *Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013*, 2014, arXiv:1401.6077.
- [30] S. Heinemeyer, M. Mondragün, G. Patellis, N. Tracas, and G. Zoupanos, *Symmetry* **10**, 62 (2018), arXiv:1802.04666.
- [31] ATLAS, G. Aad *et al.*, *JHEP* **10**, 054 (2015), arXiv:1507.05525.
- [32] ATLAS, <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/SupersymmetryPublicResults>.
- [33] CMS, A. M. Sirunyan *et al.*, *JHEP* **03**, 166 (2018), arXiv:1709.05406.
- [34] CMS, <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS>.
- [35] U. Ellwanger, C. Hugonie, and A. M. Teixeira, *Phys. Rept.* **496**, 1 (2010), arXiv:0910.1785.
- [36] M. Maniatis, *Int. J. Mod. Phys.* **A25**, 3505 (2010), arXiv:0906.0777.
- [37] CMS, C. Collaboration, (2015).
- [38] ATLAS, G. Aad *et al.*, *Phys. Rev. Lett.* **113**, 171801 (2014), arXiv:1407.6583.
- [39] S. Baum, K. Freese, N. R. Shah, and B. Shakya, *Phys. Rev.* **D95**, 115036 (2017), arXiv:1703.07800.



- [40] H.-C. Cheng, Introduction to Extra Dimensions, in *Physics of the large and the small, TASI 09, proceedings of the Theoretical Advanced Study Institute in Elementary Particle Physics, Boulder, Colorado, USA, 1-26 June 2009*, pp. 125–162, 2011, arXiv:1003.1162.
- [41] M. J. Duff, Supermembranes, in *26th British Universities Summer School in Theoretical Elementary Particle Physics (BUSSTEPP 1996) Swansea, Wales, September 3-18, 1996*, 1996, arXiv:hep-th/9611203.
- [42] S. Forste, Fortsch. Phys. **50**, 221 (2002), arXiv:hep-th/0110055.
- [43] G. C. Branco *et al.*, Phys. Rept. **516**, 1 (2012), arXiv:1106.0034.
- [44] I. P. Ivanov, Prog. Part. Nucl. Phys. **95**, 160 (2017), arXiv:1702.03776.
- [45] R. Aggleton, D. Barducci, N.-E. Bomark, S. Moretti, and C. Shepherd-Themistocleous, JHEP **02**, 035 (2017), arXiv:1609.06089.
- [46] CMS, V. Khachatryan *et al.*, JHEP **10**, 076 (2017), arXiv:1701.02032.
- [47] ATLAS, M. Aaboud *et al.*, Submitted to: Phys. Lett. (2018), arXiv:1803.11145.
- [48] M. Misiak *et al.*, Phys. Rev. Lett. **114**, 221801 (2015), arXiv:1503.01789.
- [49] A. Arbey, F. Mahmoudi, O. Stal, and T. Stefaniak, Eur. Phys. J. **C78**, 182 (2018), arXiv:1706.07414.
- [50] L. Lopez Honorez, E. Nezri, J. F. Oliver, and M. H. G. Tytgat, JCAP **0702**, 028 (2007), arXiv:hep-ph/0612275.
- [51] L. Lopez Honorez and C. E. Yaguna, JCAP **1101**, 002 (2011), arXiv:1011.1411.
- [52] A. Belyaev, G. Cacciapaglia, I. P. Ivanov, F. Rojas-Abatte, and M. Thomas, Phys. Rev. **D97**, 035011 (2018), arXiv:1612.00511.
- [53] T. Golling, Prog. Part. Nucl. Phys. **90**, 156 (2016).

## Heavy-Ion Physics

A. Ayala

Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, CdMx 04510, México, and Centre for Theoretical and Mathematical Physics, University of Cape Town, South Africa.

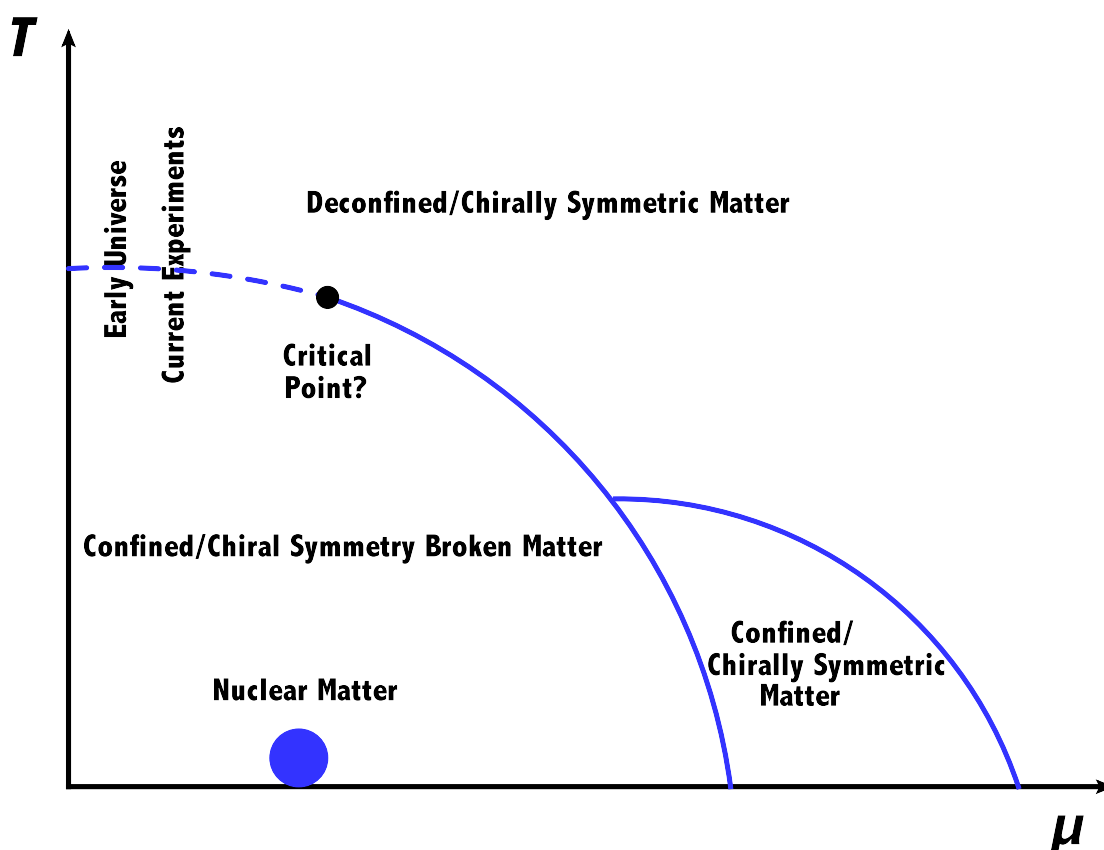
### Abstract

The field of relativistic heavy-ion physics has become an important testing ground for our current understanding of the properties of strongly interacting matter under extreme conditions. Strong interactions are described by Quantum Chromodynamics which, in addition to its defining gauge symmetry, possesses additional symmetries and properties that can be put to test in collisions of heavy-nuclei at different beam energies. The exploration of these properties is currently being conducted and will still continue for years to come when new facilities enter into operation. This is the so called study of the QCD phase diagram. Of particular relevance is the search of a possible critical end point (CEP). In these lectures I make a brief survey of QCD properties and of its symmetries. Since the phase diagram refers to the study of phase transitions, I also give the main generalities of such and mention some of the statistical tools than can be used to study the fluctuations in conserved charges to identify the CEP location. I also give a brief summary of some of the experimental signals used to study the properties of the kind of matter created in heavy-ion collisions at the highest available energy, the so called Quark Gluon Plasma.

### 1 Introduction

The field of relativistic heavy-ion physics has become the main playground to explore the properties of hadron matter under extreme conditions of temperature and density. Experimentally, the relativistic heavy-ion program started out at modest energies in the mid-1980s, when collisions of heavy-nuclei, in fixed target experiments with center-of-mass energy of 5 and 17 GeV per nucleon pair, were carried out at the Alternating Gradient Synchrotron (AGS) in Brookhaven and the Super Proton Synchrotron (SPS) at CERN, respectively. Shortly before the turn of the century, the relativistic heavy-ion program got a new boost when the Relativistic Heavy-Ion Collider (RHIC) in Brookhaven started to collide gold nuclei at a, by then unprecedented, center-of-mass energy of 200 GeV per nucleon pair. At the end of 2010, the first collisions of lead nuclei were delivered by the Large Hadron Collider at CERN, at a center-of-mass energy of 2.76 TeV per nucleon pair. By the end of 2015, the LHC was able to increase this energy to 5.02 TeV. The exciting results obtained from experiments measuring the properties of the hot and dense matter created in these collisions offer a better shaped but still incomplete picture of the properties of strongly interacting matter under extreme conditions. The program promises to keep producing new and equally exciting results when facilities such as NICA, FAIR, J-PARC and KEK, designed to explore the properties of this kind of matter at higher baryon densities, enter into operation.

The purpose of these lecture notes is to provide a theoretical framework, at the introductory level, to put into context the aim and meaning of some of the results of the experimental program. I focus on the description of the different phases that strongly interacting matter can reach when its temperature and baryon density are varied by varying the center-of-mass energy in the collisions. These phases can be better described in terms of an idealized picture based on the so called *QCD Phase Diagram*, where



**Figure 1:** Representation of the QCD phase diagram in the temperature ( $T$ ) and quark chemical potential ( $\mu$ ) plane.

the transition lines correspond to the boundaries between one and another phase. Figure 1 shows a representation of this phase diagram. Close to the phase boundaries, the relevant quark species are the light quarks  $u$ ,  $d$  and  $s$ . A complete description, accounting for the abundance of these species, should in principle be given in terms of the chemical potentials associated to each of these quarks. Nevertheless, when one assumes equilibrium, these chemical potentials are not independent from each other. They are related by the requirements of beta equilibrium and charge neutrality. Therefore, out of the three chemical potentials only one is independent. Any one of them can be chosen and the usual choice is the baryon chemical potential  $\mu_B$ , related to the quark chemical potential  $\mu$  by  $\mu = \mu_B/3$ .

The transit through the phase boundaries is related to the restoration/breaking of QCD symmetries and thus an account of such symmetries is the main unifying concept along the text. The lectures are organized as follows: In Sec. 2, I focus on the description of QCD flavor and chiral symmetries. In Sec. 3 I discuss the main features of QCD confinement and asymptotic freedom. In Sec. 4 I describe how chiral symmetry/deconfinement are restored in heavy-ion collisions at high energy. Since the description is made in terms of thermodynamical quantities, I recall the concept of phase transitions and provide a survey of our current knowledge of the QCD phase structure obtained from lattice QCD (LQCD). I also discuss how the search for the QCD critical end point is a central subject in the field and some of the theoretical tools to try to identify this point from experimental measurements. In Sec. 5 I discuss some

of the experimental signals that give an account of our current understanding of the properties of the deconfined state of matter produced in heavy-ion reactions, generally known as the quark-gluon plasma (QGP). I finally summarize in Sec. 6.

## 2 Flavor and chiral symmetries [1]

The quantum field theory that describes strong interactions is Quantum Chromodynamics (QCD). This is a gauge theory based on the local symmetry group  $SU(N_c)$ , where  $N_c$  is the number of colors. In Nature  $N_c = 3$ . The fundamental fields are the quarks (matter fields) and the gluons (gauge fields). Each one of the  $N_f$  quark fields belong to the fundamental representation of the color group which is  $N_c$ -dimensional. The antiquark fields belong to the complex conjugate of the fundamental representation, also  $N_c$ -dimensional and the gluon fields to the adjoint representation, which is  $N_c^2 - 1$ -dimensional.

The QCD Lagrangian at the classical level is written as

$$\mathcal{L}_{QCD} = \sum_{i=1}^{N_f} \bar{\psi}_i^a \left( i\gamma^\mu (\partial_\mu \delta^{ab} + ig_s A_\mu^{ab}) - m_i \delta^{ab} \right) \psi_i^b - \frac{1}{4} G_{\mu\nu}^\alpha G_{\alpha}^{\mu\nu} \quad (1)$$

where  $G_{\mu\nu}^\alpha = \partial_\mu A_\nu^\alpha - \partial_\nu A_\mu^\alpha + g_s f^{\alpha\beta\sigma} A_\mu^\beta A_\nu^\sigma$ ,  $A_\mu^{ab} = A_\mu^\sigma (\tau_\sigma)^{ab}$  with  $a, b = 1, \dots, N_c$  and  $\alpha, \beta, \sigma = 1, \dots, N_c^2 - 1$ .  $g_s$  is the coupling strength both between matter and gauge fields and between gauge fields themselves,  $m_i$  is the (bare) mass for each quark flavor and  $\tau_\sigma$  are the generators of the  $SU(N_c)$  algebra. For  $N_c = 3$ , these are usually taken as the Gell-Mann matrices.

Quarks are distinguished from one another by a quantum number called *flavor*. There are six flavors:  $u, d, s, c, b, t$ . and we refer to the number of flavors generically as  $N_f$ . Consider the ideal case in which all  $N_f$  flavors have the same mass. Quark and antiquark fields are assigned to the fundamental and complex conjugate representations (each  $N_f$ -dimensional), respectively, of the  $SU(N_f)$  group. The Lagrangian corresponding to the quark sector becomes (we omit color indices for quarks and gluons)

$$\mathcal{L}_q = \sum_{i=1}^N \bar{\psi}_i (i\gamma^\mu (\partial_\mu + ig_s A_\mu) - m) \psi_i \quad (2)$$

$\mathcal{L}_q$  is invariant under continuous *global transformations* of  $SU(N_f)$

$$\begin{aligned} \psi_i &\longrightarrow \psi'_i &= e^{-i\alpha^A (T^A)_i^j} \psi_j \\ \bar{\psi}_i &\longrightarrow \bar{\psi}'_i &= \bar{\psi}_j e^{i\alpha^A (T^A)_i^j} \\ A_\mu &\longrightarrow A'_\mu &= A_\mu, \end{aligned} \quad (3)$$

$A = 1, \dots, N_f^2 - 1$ ,  $T^A$  are  $N_f \times N_f$  matrices.

In infinitesimal form

$$\begin{aligned} \delta\psi_i &= -i\delta\alpha^A (T^A)_i^j \psi_j \\ \delta\bar{\psi}_i &= i\delta\alpha^A \bar{\psi}_j (T^A)_i^j \\ \delta A_\mu &= 0. \end{aligned} \quad (4)$$

Using Noether's theorem, one finds  $N_f^2 - 1$  conserved currents

$$\begin{aligned} j_\mu^A(x) &= \bar{\psi}(x) \gamma_\mu (T^A)_i^j \psi^i(x) \\ \partial^\mu j_\mu^A &= 0. \end{aligned} \quad (5)$$

The generators of the group (the charges) are obtained from  $j_0^A$  by space integration

$$Q^A = \int d^3x j_0^A(x). \quad (6)$$

$Q^A$ 's satisfy the  $SU(N_f)$  algebra

$$[Q^A, Q^B] = if^{ABC} Q^C, \quad (7)$$

$A, B, C = 1, \dots, N_f^2 - 1$ . Current conservation implies that the generators are independent of time and therefore they commute with the Hamiltonian  $H$

$$[H, Q^A] = 0. \quad (8)$$

The transformation properties of the fields can be translated to the states. Introduce one-particle quark states (omit spin labels)  $|\mathbf{p}, i\rangle$ . If the vacuum state is invariant under the group transformations, then

$$Q^A |\mathbf{p}, i\rangle = (T^A)_i^j |\mathbf{p}, j\rangle. \quad (9)$$

Since these states are also eigenstates of the Hamiltonian, the above means that the various one-particle states of the fundamental representation multiplet have equal masses  $m$ . This mode to realize the symmetry is called the *Wigner-Weyl* mode.

Consider however the real case where flavors have different masses

$$\mathcal{L}_q = \sum_{i=1}^N \bar{\psi}_i (i\gamma^\mu (\partial_\mu + ig_s A_\mu) - m_i) \psi_i. \quad (10)$$

The mass term spoils  $SU(N_f)$  invariance, therefore currents are not conserved and we find

$$\partial^\mu j_\mu^A = -i \sum_{i,j=1}^{N_f} (m_i - m_j) \bar{\psi}_i (T^A)^i_j \psi^j \neq 0 \quad (11)$$

In Nature, quarks are divided into two groups: *Light quarks*  $u, d, s$  and *heavy quarks*  $c, b, t$ . The mass difference between each group is large ( $> 1$  GeV). An approximate symmetry can be expected only for the light quarks:  $SU(2)$  for  $u, d$  or  $SU(3)$  for  $u, d, s$ .

Notice that quarks can also be transformed by means of unitary transformations that include the  $\gamma_5$  matrix. The transformations are called *axial flavor transformations*. In infinitesimal form

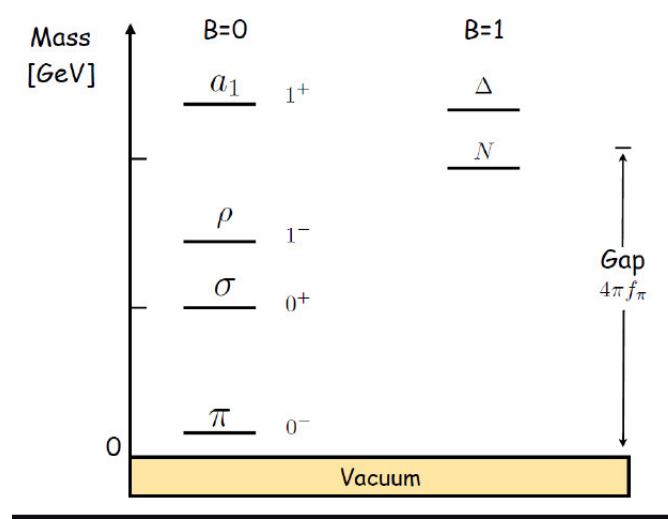
$$\begin{aligned} \delta \psi_i &= -i \delta \alpha^A (T^A)_i^j \gamma_5 \psi_j \\ \delta \bar{\psi}_i &= i \delta \alpha^A \bar{\psi}_j \gamma_5 (T^A)^j_i \\ \delta A_\mu &= 0. \end{aligned} \quad (12)$$

Consider the quark part of the Lagrangian  $\mathcal{L}_q$  with equal masses. Under these transformations  $\mathcal{L}_q$  is not invariant because of the mass

$$\delta \mathcal{L}_q = 2im \delta \alpha^A \bar{\psi}^i (T^A)_i^j \gamma_5 \psi_j. \quad (13)$$

Invariance under axial flavor transformations requires vanishing of the quark mass. Contrary to flavor symmetry transformations, equality of masses is not sufficient for invariance under axial flavor transformations. In the general case of different masses, we introduce the *mass matrix*  $\mathcal{M} = \text{diag}(m_1, m_2, \dots, m_{N_f})$ . The variation of the Lagrangian is

$$\delta \mathcal{L}_q = i \delta \alpha^A \bar{\psi}^i \{ \mathcal{M}, T^A \}_i^j \gamma_5 \psi_j, \quad (14)$$



**Figure 2:** Schematic representation of the parity doublets for the lightest meson and baryon states.

where  $\{, \}$  is the anticommutator.

Consider the massless case. The Lagrangian is invariant under both the flavor and the axial flavor transformations. The conserved currents are

$$\begin{aligned} j_\mu^A(x) &= \bar{\psi}_i(x) \gamma_\mu (T^A)^i_j \psi^j(x); & \partial^\mu j_\mu^A &= 0 \\ j_{5\mu}^A(x) &= \bar{\psi}_i(x) \gamma_\mu \gamma_5 (T^A)^i_j \psi^j(x); & \partial^\mu j_{5\mu}^A &= 0. \end{aligned} \quad (15)$$

The corresponding charges are

$$Q^A = \int d^3x j_0^A(x), \quad Q_5^A = \int d^3x j_{50}^A(x). \quad (16)$$

Together the flavor and axial flavor transformations form the *chiral transformations*. The corresponding charges satisfy the commutation relations

$$[Q^A, Q^B] = if^{ABC} Q^C, \quad [Q^A, Q_5^B] = if^{ABC} Q_5^C, \quad [Q_5^A, Q_5^B] = if^{ABC} Q^C. \quad (17)$$

The axial charges do not form an algebra, however, if we define

$$Q_L^A = \frac{1}{2}(Q^A - Q_5^A), \quad Q_R^A = \frac{1}{2}(Q^A + Q_5^A) \quad (18)$$

we obtain

$$[Q_L^A, Q_L^B] = if^{ABC} Q_L^C, \quad [Q_R^A, Q_R^B] = if^{ABC} Q_R^C, \quad [Q_L^A, Q_R^B] = 0. \quad (19)$$

The result can be summarized as follows: The left-handed and right-handed charges decouple and operate separately. Each generate an  $SU(N_f)$  group. The chiral group is decomposed into the direct product of two  $SU(N_f)$  groups, labeled with the subscripts  $L$  and  $R$ , *i.e.* the chiral group =  $SU(N_f)_L \otimes SU(N_f)_R$ . In Nature, chiral symmetry is not exact, *quark masses break it explicitly*. In the light quark sector, the breaking can be treated as a perturbation, the symmetry is approximate. What is the signature of this approximate symmetry?

Suppose that the symmetry is realized in the Wigner-Weyl mode. Thus, in the massless quark limit a chiral transformation acting on a massive state gives

$$Q_5^A |M, s, \mathbf{p}, +, i\rangle = (T^A)^j_i |M, s, \mathbf{p}, -, j\rangle. \quad (20)$$

This means that we should find parity partners for the massive states. When we consider the light quark masses, the degeneracy within parity doublets is lifted, but the masses should remain close to each other. No degenerate parity doublets are observed, thus the Wigner-Weyl mode realization of chiral symmetry does not happen under ordinary conditions. The alternative is spontaneous symmetry breaking, also known as Nambu-Goldstone mode. What happens if the generators of some transformations do not annihilate the vacuum?

$$Q_5^A |0\rangle \neq 0. \quad (21)$$

In this case we say that the symmetry has been spontaneously broken. The axial charges when applied to the vacuum state produce new states

$$Q_5^A |0\rangle = |A, -\rangle, \quad A = 1, \dots, N_f^2 - 1 \quad (22)$$

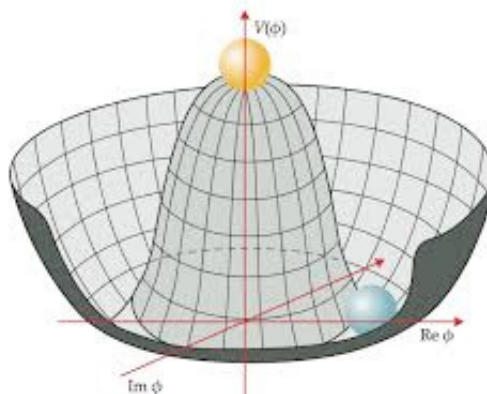
The states have the same properties as the axial charges that generate it, in particular they are *pseudoscalar* states.

In the massless limit, the charges commute with the Hamiltonian therefore, these states are massless (Goldstone theorem). Spontaneous chiral symmetry breaking is manifested by means of the existence of  $N_f^2 - 1$  pseudoscalar massless particles called Nambu-Goldstone bosons. The breaking of the symmetry involves only the axial sector. The ordinary flavor symmetry is still realized in the Wigner-Weyl mode.

$$SU(N_f)_L \otimes SU(N_f)_R \longrightarrow SU(N_f)_V. \quad (23)$$

In Nature  $N_f = 3$ , this corresponds to eight pseudoscalar bosons ( $\pi, K, \eta$ ) with small masses.

We have thus seen that chiral symmetry is broken due to a finite mass of the quarks. Under ordinary circumstances, quarks are confined within hadrons. The quark mass is thus an *effective* mass which corresponds to about one third of the nucleon mass  $\sim 300$  MeV. The origin of this effective, or dynamically generated mass, is the strength of the interaction and can be explained in terms of non-perturbative methods [2]. It seems intuitively clear that if we could somehow overcome the confining of quarks, their mass would decrease and chiral symmetry would be *restored*. Let us now study how the strong interaction produces this confinement to later ask how a collision of relativistic nuclei could help to study the situation where chiral symmetry and deconfinement are achieved.



**Figure 3:** Schematic representation of spontaneous symmetry breaking. Goldstone bosons correspond to the directions where the potential is flat.



### 3 Confinement and asymptotic freedom

To grasp why under ordinary circumstances quarks are confined within hadrons let us recall the characteristics of the strong interaction: The potential between two quarks at “long” distances,  $\mathcal{O}(1 \text{ fm})$ , is linear, this means that the separation of quarks requires “infinite” amount of energy. Confinement is a direct consequence of the gluon self-interaction. Quarks and gluons confined inside the QCD potential must combine into zero net color charge particles called hadrons.

The strength of strong interaction is characterized by a coupling “constant”  $\alpha_s = g_s^2/4\pi$  whose strength decreases with distance since the bare color charge is *antiscreened* due to gluon self-interaction. To study confinement and the *running* of the strong coupling with distance (or equivalently, with the energy involved in the physical process), one resorts to a renormalization group analysis of the gluon polarization tensor

$$\Pi^{\mu\nu}(q) = \Pi(q^2, \alpha_s) \left( g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2} \right), \quad (24)$$

where we refer to  $\Pi$  as the gluon self-energy. Gauge invariance dictates that the gluon polarization tensor be transverse and thus the tensor structure in Eq. (24). Let  $\Pi(q^2)$  represent the un-renormalized gluon self-energy. Let us scale each factor of the momentum  $q$  appearing in  $\Pi$  by the renormalization ultraviolet energy scale  $\tilde{\mu}$ , writing

$$q^2 = \tilde{\mu}^2(q^2/\tilde{\mu}^2). \quad (25)$$

Therefore, we have

$$\Pi(q^2; \alpha_s) = \tilde{\mu}^D \Pi(q^2/\tilde{\mu}^2; \alpha_s), \quad (26)$$

where  $D = 2$  is the energy dimension of  $\Pi$ . Since  $\tilde{\mu}$  is arbitrary, the statement that  $\Pi$  should be independent of this scale is provided by the Renormalization Group Equation (RGE) [3]

$$\left( \tilde{\mu} \frac{\partial}{\partial \tilde{\mu}} + \alpha_s \beta(\alpha_s) \frac{\partial}{\partial \alpha_s} - \gamma \right) \Pi(q^2; \alpha_s) = 0, \quad (27)$$

where  $\beta(\alpha_s)$  is the QCD beta function defined by

$$\alpha_s \beta(\alpha_s) = \tilde{\mu} \frac{\partial \alpha_s}{\partial \tilde{\mu}}, \quad (28)$$

and

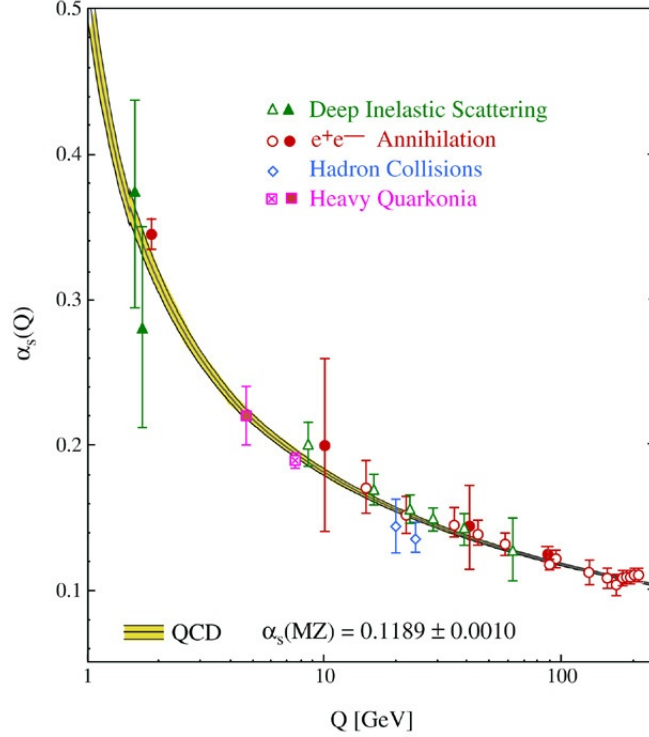
$$\gamma = \tilde{\mu} \frac{\partial}{\partial \tilde{\mu}} \ln Z^{-1}, \quad (29)$$

where  $Z$  is the gluon self-energy renormalization. The beta function represents the rate of change of the renormalized coupling as the renormalization scale  $\tilde{\mu}$  is increased. The dependence of a given Green’s function on  $\tilde{\mu}$  happens through the counter-terms that subtract ultraviolet divergences. Therefore, the beta function can be computed from the counter-terms that enter a properly chosen Green’s function. In QCD, to lowest order, the beta function can be computed as

$$\beta = g_s \tilde{\mu} \frac{\partial}{\partial \tilde{\mu}} \left( -\delta_1 + \delta_2 + \frac{1}{2} \delta_3 \right) \quad (30)$$

where  $\delta_1$ ,  $\delta_2$  and  $\delta_3$  are the counter-terms for the quark-gluon vertex, the quark self-energy and the gluon self-energy [4]. The QCD beta function is negative and to one-loop level it is given by

$$\beta(\alpha_s) = -b_1 \alpha_s, \quad b_1 = \frac{1}{12\pi} (11N_c - 2N_f) > 0. \quad (31)$$



**Figure 4:** Running of  $\alpha_s$  as a function of the momentum transferred  $Q$  in the corresponding process.

To find the evolution of the strong coupling with the momentum scale, we start from Eq. (27) and introduce the variable

$$t = \ln(Q^2/\tilde{\mu}^2), \quad (32)$$

where  $Q^2$  is the momentum transferred in a given process. Notice that the reference scale  $\tilde{\mu}$  is usually large enough, so as to make sure that the calculation is well within the perturbative domain, therefore  $Q^2 < \tilde{\mu}^2$ . After this change of variable, the RGE becomes

$$\left( -\frac{\partial}{\partial t} + \alpha_s \beta(\alpha_s) \frac{\partial}{\partial \alpha_s} - \gamma \right) \Pi(q^2; \alpha_s) = 0. \quad (33)$$

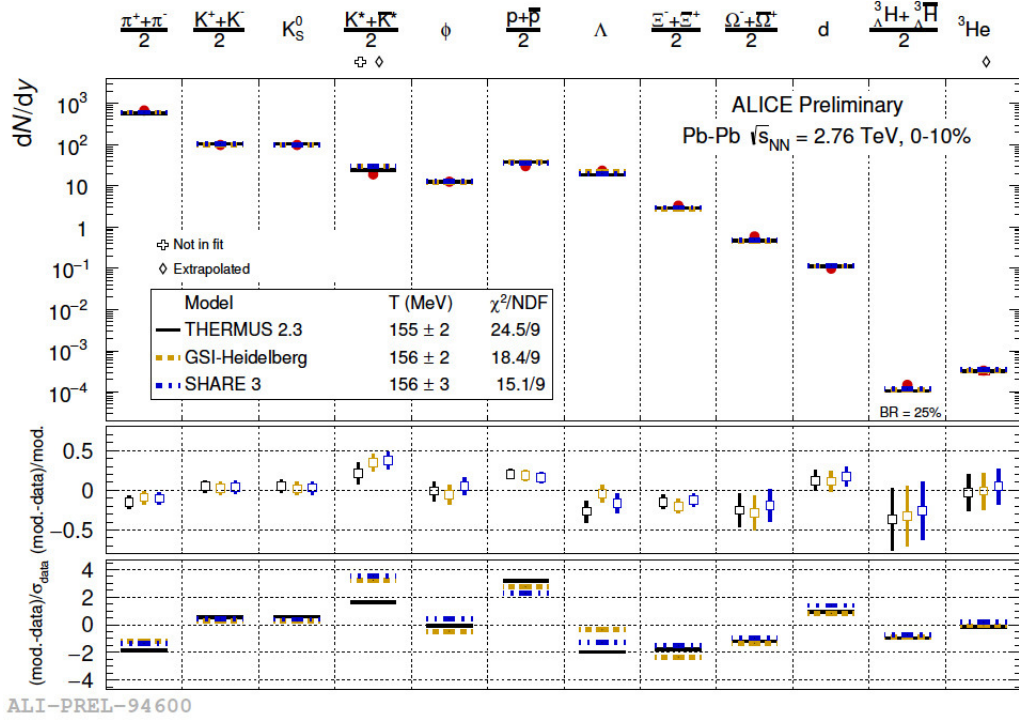
Using the method of the characteristics [5], one obtains the relation between the coupling values evaluated at  $Q^2$  and the reference scale  $\mu^2$  as

$$\int_{t(Q^2=\tilde{\mu}^2)}^{t(Q^2)} dt = -\frac{1}{b_1} \int_{\alpha_s(Q^2=\tilde{\mu}^2)}^{\alpha_s(Q^2)} \frac{d\alpha_s}{\alpha_s^2}. \quad (34)$$

Solving for  $\alpha_s(Q^2)$ , we obtain

$$\alpha_s(Q^2) = \frac{\alpha_s(\tilde{\mu}^2)}{1 + b_1 \alpha_s(\tilde{\mu}^2) \ln(Q^2/\tilde{\mu}^2)}, \quad (35)$$

from where it is seen that as  $Q$  increases, the coupling decreases. This behavior is known as *asymptotic freedom* and it is responsible for the fact that strong interaction processes can be computed in perturbation theory when the transferred momentum is large. Conversely, when this momentum is small, the coupling is so large that perturbative calculations become meaningless. This is the so called *non-perturbative*



**Figure 5:** Particle ratios obtained in the statistical model compared to experimentally measured ratios in central Pb + Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV, obtained by the ALICE Collaboration.

*regime*. Processes where perturbation theory can be applied are usually those where the transferred momentum satisfies  $Q^2 \gtrsim 1 \text{ GeV}^2$ . To quantify this statement, notice that from Eq. (35) we can define a transferred momentum value  $\Lambda_{QCD}$  small enough such that the denominator vanishes and thus the coupling blows up, namely

$$1 + b_1 \alpha_s(\mu^2) \ln(\Lambda_{QCD}^2 / \tilde{\mu}^2) = 0, \quad \Lambda_{QCD}^2 = \tilde{\mu}^2 e^{-\frac{1}{b_1 \alpha_s(\tilde{\mu})}} \quad (36)$$

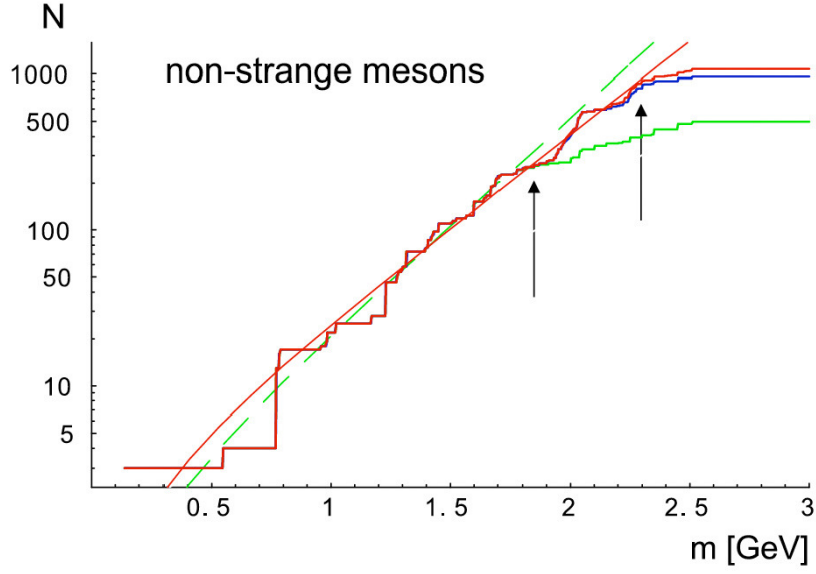
$\Lambda_{QCD}$  is a renormalization scheme dependent quantity. It also depends on the number of active flavors. For instance, in the  $\overline{\text{MS}}$  scheme and for three active flavors, its value is of order  $\Lambda_{QCD} \sim 300$  MeV. All dimensionful QCD results where the transferred momentum is small, scale with  $\Lambda_{QCD}$ . The existence of this scale is the reason for the existence of the mass of baryons and thus of the mass of the visible universe.

The question we now set up to address concerns whether and how heavy-ion reactions allow us to explore the situation where the two above discussed properties of QCD, namely, the breaking of chiral symmetry by the effective quark mass and confinement, can be overcome.

#### 4 Chiral symmetry/deconfinement transitions at high temperature and baryon density

When nuclear matter is subject to extreme conditions, interesting phenomena take place. There are two important parameters when this matter can be described as being in equilibrium: The temperature  $T$  and the baryon number density  $n_B$  (or equivalently its conjugate variable  $\mu_B = 3\mu$ ). Since the intrinsic QCD scale is of order  $\Lambda_{QCD} \sim 200 - 300$  MeV, one expects a transition around  $T \sim \Lambda_{QCD}$  and/or  $n_B \sim \Lambda_{QCD}^3 \sim 1 \text{ fm}^{-3}$ . The temperature and baryon density in a heavy-ion reaction are functions of the center-of-mass energy in the collision.

To estimate the possible values that these quantities achieve, one usually resorts to the *statistical model*. This model assumes that hadron matter is in thermal equilibrium during chemical freeze-out,



**Figure 6:** Schematic representation of the meson density of states as a function of their mass. Notice that the density of states increases exponentially. Figure taken from Ref. [8].

that is to say, when hadron abundances are established during the reaction. In this manner, it is possible to extract the values for the temperature and baryon chemical potential from fits to particle abundances, given in the model by

$$n_j = \frac{g_j}{2\pi^2} \int_0^\infty p^2 dp \left[ \exp \left\{ \sqrt{p^2 + M_j^2} / T_{ch} - \mu_{ch} \right\} \pm 1 \right]^{-1}, \quad (37)$$

where  $\pm$  refer to fermions and bosons, respectively. Figure 5 illustrates a comparison between the statistical model calculations and experimentally measured particle ratios. The main claim of the model is that since multi-particle scattering rates fall-off rapidly, the experimentally determined chemical freeze-out temperature  $T_{ch}$  and chemical potential  $\mu_{ch}$  are good measures of the phase transition temperature and baryon chemical potential [6].

Using this model, it is found that for central collisions, the baryon chemical potential decreases roughly as the inverse of the center-of-mass energy per nucleon pair in the collision [7]. Therefore, for current collider based experiments where the highest energies are reached, such as the LHC and RHIC, the baryon chemical potential associated to the reaction is the smallest. This can be understood in terms of an increase of the degree of transparency of colliding nuclear matter with collision energy, whereby in the interaction zone, the energy deposited produces roughly an equal number of particles and antiparticles. On the contrary, when the collision energy decreases, the transparency decreases and the reaction zone becomes baryon richer.

A systematic exploration over a wider range of  $n_B$  values, up to several times the normal nuclear matter density  $n_0 \sim 0.16 \text{ fm}^{-3}$ , can be carried out currently by the BES-RHIC and in the future by other facilities such as FAIR, NICA, J-PARC and KEK. In Nature, the interior of compact stellar objects is the relevant system where dense and low temperature QCD matter is realized.

#### 4.1 Phase transitions

A phase transition is a transformation of a given substance from one state of matter to another. During the phase transition some quantities change, often in a discontinuous manner. Changes result from variations of external conditions such as pressure, temperature, etc. In technical terms, a phase transition occurs when the free energy is non-analytic (one of its derivatives diverges) for some values of the thermodynamical variables. Phase transitions result from the interaction of a large number of particles and in general do not occur when the system is very small or has a small number of particles. On the phase transition lines the free energies in both phases coincide. Sometimes it is possible to change the state of a substance without crossing a phase transition line. Under these conditions one talks about a crossover transition.

Phase transitions are classified according to behavior of the free energy as a function of a given thermodynamical variable (Ehrenfest). They are named according to the derivative of lowest order that becomes discontinuous during the transition: First order; the first derivative of free energy is discontinuous. A prototypical example is boiling water. During this process there appears a discontinuity in the density, *i.e.* the derivative of free energy with respect to chemical potential. Second order; the first derivative is continuous but the second derivative is discontinuous. A prototypical example is ferromagnetism. The magnetization, *i.e.* the derivative of the free energy with respect to the external field is continuous. The susceptibility, *i.e.* the derivative of the magnetization with respect to the external field, is discontinuous.

In a modern classification, a first order phase transition involves latent heat. The system absorbs or releases heat at a constant temperature. Phases coexist, although some parts have completed the transition whereas some others have not. A second order phase transition is a continuous transitions. Susceptibilities diverge, correlation lengths become large.

#### 4.2 Deconfinement transition form hadron thermodynamics

Consider an ideal gas of identical neutral scalar particles of mass  $m_0$  contained in a box volume  $V$ . Assume Boltzmann statistics. The partition function and related thermodynamical quantities are given by

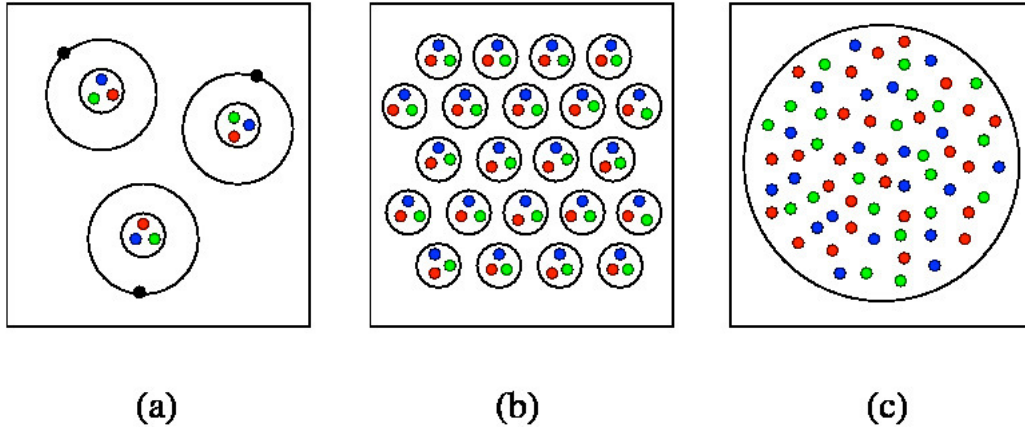
$$\begin{aligned}
 \mathcal{Z}(T, V) &= \sum_N \frac{1}{N!} \left[ \frac{V}{(2\pi)^3} \int d^3 p \exp \left\{ -\frac{\sqrt{p^2 + m^2}}{T} \right\} \right]^N \\
 \ln \mathcal{Z}(T, V) &= \frac{VTm_0^2}{2\pi^2} K_2(m_0/T) \\
 \varepsilon(T) &= -\frac{1}{V} \frac{\partial \ln \mathcal{Z}(T, V)}{\partial (1/T)} \xrightarrow{T \gg m_0} \frac{3}{\pi^2} T^4 \quad \text{energy density} \\
 n(T) &= -\frac{1}{V} \frac{\partial \ln \mathcal{Z}(T, V)}{\partial (V)} \xrightarrow{T \gg m_0} \frac{1}{\pi^2} T^3 \quad \text{particle density} \\
 \omega(T) &= \varepsilon(T)/n(T) \simeq 3T \quad \text{average energy per particle} \tag{38}
 \end{aligned}$$

Notice that the above relations imply that an increase of system's energy has three consequences: A higher temperature, more constituents, more energetic constituents.

Let us now include in the analysis hadron resonances whose mass is  $m_i$ . It is easy to show that the partition function is given by

$$\ln \mathcal{Z}(T, V) = \sum_i \frac{VTm_i^2}{2\pi^2} \rho(m_i) K_2(m_i/T), \tag{39}$$

where  $i$  starts with the ground state ( $m_0$ ) and then includes the possible resonances with masses  $m_i$  and  $\rho(m_i)$  is the weight (density of states) corresponding to the state  $m_i$ . It is thus crucial to determine  $\rho(m_i)$ , *i.e.*, how many states there are having mass  $m_i$ .



**Figure 7:** Schematic representation of the change of effective degrees of freedom for the description of a set of hadrons when the density and/or temperature increases, as a consequence of a heavy-ion reaction. At low temperatures/densities, the description is given in terms of a collection of hadron resonances. As the temperature/density increases, limiting values of the baryon density and/or temperature are reached above which the descriptions is made in terms of the fundamental QCD degrees of freedom. This signals the onset of a phase transition.

Figure 4 shows a schematic representation of the density of states, albeit for non-strange mesons. In any case, it serves to illustrate that the density of states grows exponentially with the mass of the species,  $\rho(m) \propto \exp\{m/T^H\}$ , where  $T^H \simeq 0.19$  GeV. This exponential growth should be balanced by the Boltzmann factor

$$\exp\left\{\frac{m}{T^H} - \frac{m}{T}\right\}, \quad (40)$$

such that when  $T > T^H$ , the integration over  $m$  becomes singular.  $T^H$  plays the role of a limiting temperature known as the *Hagedorn* temperature above which the hadronic description breaks down.

Applying a similar argument, we can also estimate the critical line at finite  $\mu_B$ . The density of baryon states  $\rho(m_B) \propto \exp\{m_B/T^H\}$ , where  $m_B$  is the typical baryon mass (of order 1 GeV) is balanced by the Boltzmann factor  $\exp\{-(m_B - \mu_B)/T\}$ . The limiting temperature becomes

$$T = \left(1 - \frac{\mu_B}{m_B}\right) T^H. \quad (41)$$

Notice that the treatment of the system as a gas made out of resonances leads to three consequences: More and more species of ever heavier particles appear; a fixed temperature limit  $T \rightarrow T^H$  above which the resonance picture does not hold; the momentum of the constituents do not continue to increase. All in all, these observations imply that above and to the right of the limiting curve, Eq. (41), a different description of hadron matter, in terms of degrees of freedom other than hadron resonances, is called for. The situation is illustrated in Fig. 6. As the system becomes hotter/denser, the boundaries between individual hadrons disappear and the description should be made instead in terms of the fundamental QCD degrees of freedom. A deconfinement phase transition takes place.

### 4.3 Chiral symmetry transition

The QCD vacuum within hadrons should be regarded as a medium responsible for the non-perturbative quark mass. In hot and/or dense matter, quarks turn bare due to asymptotic freedom. We expect a phase transition from a state with heavy constituent quarks to another with light current quarks. This transition is called *chiral phase transition*.

At finite  $T$  and  $\mu$  the QCD phase diagram can also be studied from the point of view of chiral symmetry restoration. In the chiral limit ( $m = 0$ ), a true order parameter for the transition is the *quark-antiquark condensate*  $\langle \bar{\psi}\psi \rangle$ , since it is finite in the chirally broken (Nambu-Goldstone) phase and vanishes in the chirally symmetric (Wigner-Weyl) phase. In this limit the (true) critical temperature can be obtained from the *chiral susceptibility*

$$\chi_m = \frac{\partial}{\partial m} \langle \bar{\psi}\psi \rangle. \quad (42)$$

In vacuum  $\langle \bar{\psi}\psi \rangle_0 = -(0.24 \text{ GeV})^3$ . This value sets the scale for the critical temperature of chiral restoration.

From  $\chi$ PT at low  $T$  and low  $n_B$  one knows that

$$\begin{aligned} \langle \bar{\psi}\psi \rangle_T / \langle \bar{\psi}\psi \rangle_0 &= 1 - T^2 / (8f_\pi^2) - T^4 / (384f_\pi^4) \\ \langle \bar{\psi}\psi \rangle_{n_B} / \langle \bar{\psi}\psi \rangle_0 &= 1 - \sigma_{\pi N} n_B / (f_\pi^2 m_\pi^2) - \dots \end{aligned} \quad (43)$$

where  $f_\pi \simeq 93 \text{ MeV}$ , is the pion decay constant and  $\sigma_{\pi N} = 40 \text{ MeV}$ , is the  $\pi - N$  sigma term. Equations (43) indicate that the quark-antiquark condensate melts at finite  $T$  and  $n_B$ .

For physical quark masses neither  $\langle \bar{\psi}\psi \rangle$  vanishes nor  $\chi_m$  diverges at the pseudocritical temperature. Nevertheless these quantities retain a behavior reminiscent of the corresponding one in the chiral limit. In particular  $\chi_m$  has a peaked structure as a function of  $T$ . It is customary to define  $T_c$  as the temperature for which  $\chi_m$  reaches its peak. Other susceptibilities, such as  $\chi_T = \frac{\partial}{\partial T} \langle \bar{\psi}\psi \rangle$ , can also be used to define  $T_c$ . It has been shown [9] that the critical temperatures thus obtained coincide within a narrow band and therefore using any of these susceptibilities gives basically the same  $T_c$ .

Lattice calculations have provided values for  $T_c$  extracted from the peak of  $\chi_m$  for 2+1 flavors using different types of improved staggered fermions [10]. These values show some discrepancies. The MILC collaboration [11] obtained  $T_c = 169(12)(4) \text{ MeV}$ . The BNL-RBC-Bielefeld collaboration [12] reported  $T_c = 193(7)(4) \text{ MeV}$ . The Wuppertal-Budapest collaboration [13] has consistently obtained smaller values, the latest being  $T_c = 147(2)(3) \text{ MeV}$ . The HotQCD collaboration [14] has reported  $T_c = 154(9) \text{ MeV}$ . The differences could perhaps be attributed to different lattice spacings.

The unambiguous picture presented by lattice QCD for  $T \geq 0$ ,  $\mu = 0$  cannot be easily extended to the case  $\mu \neq 0$ , given that standard Monte Carlo simulations can only be applied to the case where  $\mu$  is either 0 or purely imaginary. Simulations with  $\mu \neq 0$  are hindered by the *sign problem* [15]. Recall that in the computation of the QCD partition function with finite  $\mu$ , integration over each fermion field produces a *fermion determinant*, i.e. a factor  $\text{Det}M \equiv \text{Det}(\not{D} + m + \mu\gamma_0)$ , where  $M$  is the fermion matrix. Let us consider in general a complex  $\mu$ . Taking the determinant on both sides of the identity

$$\gamma_5 (\not{D} + m + \mu\gamma_0) \gamma_5 = (\not{D} + m - \mu^* \gamma_0)^\dagger, \quad (44)$$

we get

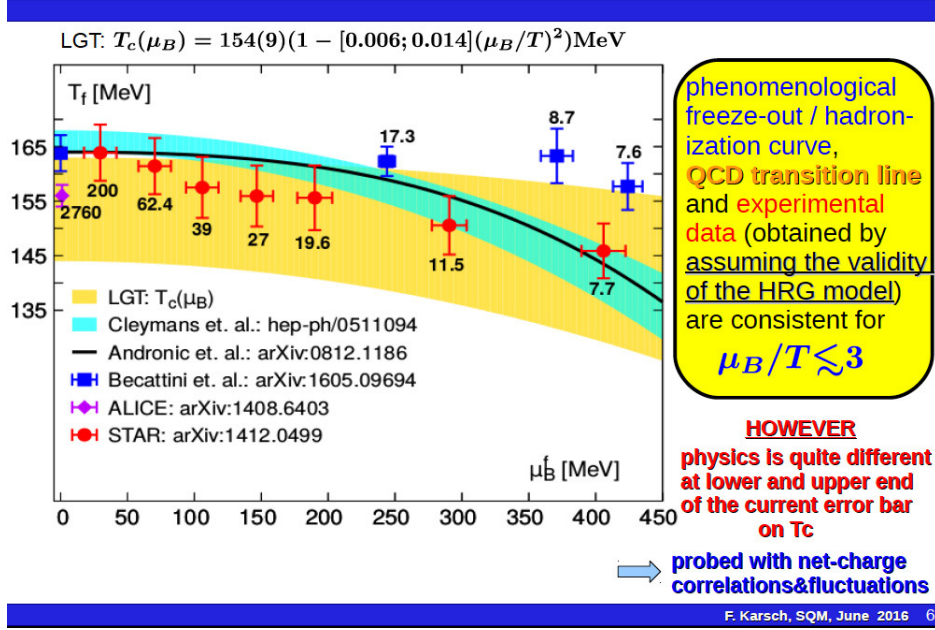
$$\text{Det}(\not{D} + m + \mu\gamma_0) = [\text{Det}(\not{D} + m - \mu^* \gamma_0)]^*, \quad (45)$$

which shows that the fermion determinant is not real unless  $\mu = 0$  or purely imaginary. For real  $\mu$  the direct sampling of a finite density ensemble by Monte Carlo methods is not possible, since the sampling requires a real non-negative measure. This problem is referred to as the sign problem, although a more appropriate name would seem to be the *phase problem*.

That the integrand of the partition function becomes complex would seem to be only a minor inconvenience. A naive approach to deal with the sign problem would be to write [16]

$$\text{Det}M = |\text{Det}M| e^{i\theta}. \quad (46)$$

## Chiral transition, hadronization and freeze-out



**Figure 8:** Comparison between phenomenological freeze-out curves with the experimentally measured values for the temperature and baryon chemical potential obtained from particle abundances using the statistical model. Shown is also the analytical expression for the chiral phase transition curve obtained by means of LQCD techniques.

To compute the thermal average of an observable  $O$  in QCD one writes

$$\langle O \rangle = \frac{\int DU e^{S_{YM}} \text{Det}M O}{\int DU e^{S_{YM}} \text{Det}M} = \frac{\int DU e^{S_{YM}} |\text{Det}M| e^{i\theta} O}{\int DU e^{S_{YM}} |\text{Det}M| e^{i\theta}}, \quad (47)$$

where  $S_{YM}$  is the Yang-Mills action. Notice that written in this manner, the simulations could be implemented in terms of the phase-quenched (pq) theory where the measure involves  $|\text{Det}M|$  and the thermal average could be written as

$$\langle O \rangle = \frac{\langle O e^{i\theta} \rangle_{\text{pq}}}{\langle e^{i\theta} \rangle_{\text{pq}}}. \quad (48)$$

The average phase factor (also referred to as the average sign) in the phase-quenched theory can be written as

$$\langle e^{i\theta} \rangle_{\text{pq}} = e^{-V(f-f_{\text{pq}})/T}, \quad (49)$$

where  $f$  and  $f_{\text{pq}}$  are the free-energy densities in the full and phase-quenched theories, respectively and  $V$  is the three-dimensional volume. If  $f - f_{\text{pq}} \neq 0$ , the average phase factor decreases exponentially when  $V$  goes to infinity (the thermodynamic limit) and/or  $T$  goes to zero. Under these circumstances the signal to noise ratio worsens. This is referred to as the *severe sign problem*.

To circumvent the sign problem, a possibility is to determine the first Taylor coefficients in the expansion of a given observable in powers of  $\mu_B/T$ . The coefficients of the expansion can be expressed as expectation values of traces of matrix polynomials taken in the  $\mu_B = 0$  ensemble. Although care has to be taken regarding the growth of the statistical errors with the order of the expansion, this strategy has



provided a very important result: the curvature of the transition line at  $\mu_B = 0$ . The curvature  $\kappa$  is defined as the dimensionless coefficient of the quadratic term in the Taylor expansion of the pseudocritical line

$$T_c(\mu_B) = T_c \left( 1 - \kappa \frac{\mu_B^2}{T_c^2} \right), \quad \kappa \equiv - \left( T_c \frac{dT_c(\mu_B)}{d\mu_B^2} \right) \Big|_{\mu_B=0}. \quad (50)$$

Values for  $\kappa$  between 0.006 – 0.02 have been reported [17]. It should be noted that since the phases are separated by a crossover, the curvature should depend in principle on the observable considered. Nevertheless, these curvatures give considerable smaller values than that of the chemical freeze-out curve obtained from statistical models [18]. This observation could be of potential importance for if the pseudo critical line is flatter than the chemical freeze-out line, the distance between the possible QCD critical end point (CEP) and the freeze-out curve increases. If this happens then possible experimental evidences for criticality may be washed out by the moment when particle abundances are established after a heavy-ion collision.

Figure 8 shows a comparison between phenomenological freeze-out curves with the experimentally measured values for the temperature and baryon chemical potential obtained from particle abundances using the statistical model. Shown is also the analytical expression for the chiral phase transition curve obtained by means of LQCD techniques. Although the physics portrayed in the statistical model is quite different from the physics of chiral symmetry restoration, the agreement of the descriptions is remarkable, a similarity that is worth exploring and understanding.

#### 4.4 The critical end point

In the study of the QCD phase diagram, the location of a possible Critical End Point (CEP) is of particular relevance. This point marks the end of a first order phase transition line. There are strong arguments based on effective models suggesting that close to the  $\mu_B$ -axis, the transition is first order [19]. Since, on the other hand, close to the  $T$ -axis, the transition is a smooth crossover, a CEP should be located somewhere in the middle of the phase diagram. To locate its position, the STAR BES-I program has recently analyzed collisions of heavy-nuclei in the energy range  $200 \text{ GeV} > \sqrt{s_{NN}} > 7.7 \text{ GeV}$  [20]. Future experiments [21–23] will keep on conducting an exploration to locate the CEP changing the collision energy down to about  $\sqrt{s_{NN}} \simeq 5 \text{ GeV}$  and the system size in heavy-ion reactions.

From the theoretical side, efforts to locate the CEP make use of a variety of techniques. These involve Schwinger-Dyson equations, finite energy sum rules, functional renormalization methods, holography, and effective models [24–32, 36–38]. These studies have produced a wealth of results for the CEP location that range from low to large values  $\mu_B$  and  $T$ . Recent LQCD analyses [39] have resorted to the imaginary baryon chemical potential technique, extrapolating afterwards to real values, to study the chiral transition near the  $T$ -axis. The method has still large uncertainties, however this technique has shown that the transition keeps being a smooth crossover [40]. The Taylor expansion LQCD technique has also been employed to restrict the CEP's location to values  $\mu_B/T > 2$  for the temperature range  $135 \text{ MeV} < T < 155 \text{ MeV}$ . Its location for temperatures larger than  $0.9 T^c(\mu_B = 0)$  seems to also be highly disfavored [41] (see also [42]).

Table I summarizes the CEP location found in some recent works.

A powerful tool to experimentally locate the CEP is the study of event-by-event fluctuations in relativistic heavy-ion collisions [44]. These are sensitive to the early thermal properties of the created medium. In particular, the possibility to detect non Gaussian fluctuations in conserved charges is one of the central topics in this field.

Let  $n(x)$  be the density of a given charge  $Q$  in the phase space described by the set of variables  $x$ . These quantities are related by

$$Q = \int_V dx n(x), \quad (51)$$

Reference	$T_{CEP}$	$\mu_{CEP}$
C. Shi, <i>et al.</i> [29]	$0.85 T_c$	$1.11 T_c$
G. A. Contrera, <i>et al.</i> [30]	69.9 MeV	319.1 MeV
T. Yokota, <i>et al.</i> [33]	5.1 MeV	286.7 MeV
S. Sharma [34]	145-155 MeV	$> 2 T_{CEP}$
J. Knaute, <i>et al.</i> [36]	112 MeV	204 MeV
N. G. Antoniou, <i>et al.</i> [37]	119-162 MeV	84-86 MeV
Z. F. Cui, <i>et al.</i> [31]	38 MeV	245 MeV
P. Kovács and G. Wolf [35]		$> 133.3$ MeV
R. Rougemont, <i>et al.</i> [38]	$< 130$ MeV	$> 133.3$ MeV
A. Ayala, <i>et al.</i> [43]	18-45 MeV	315-349 MeV

**Table 1:** Summary of some recent results for the CEP location taken from Ref. [43].

where  $V$  is the total phase space volume available. When the measurement of  $Q$  is performed over the volume  $V$  in a thermal system, we speak of a thermal fluctuation. For instance, the variance of  $Q$  is given by

$$\langle \delta Q^2 \rangle_V = \langle (Q - \langle Q \rangle_V)^2 \rangle_V = \int_V dx_1 dx_2 \langle \delta n(x_1) \delta n(x_2) \rangle \quad (52)$$

The integrand on the right hand side of Eq. (52) is called a *correlation function*, whereas the right hand side is called a (second order) *fluctuation*. With this example, we see that fluctuations are closely related to correlation functions. In relativistic heavy-ion collisions, fluctuations are measured on an event-by-event basis in which the number of some charge or particle species is counted in each event. Although these fluctuations are not necessarily equal to the thermal fluctuations, there are good reasons to expect that, with an appropriate treatment these two can be taken as equivalent.

For a probability distribution function  $\mathcal{P}(x)$  of an stochastic variable  $x$ , the moments are defined as

$$\langle x^n \rangle = \int dx x^n \mathcal{P}(x). \quad (53)$$

We can define the *moment generating function*  $G(\theta)$  as

$$G(\theta) = \int dx e^{x\theta} \mathcal{P}(x), \quad (54)$$

from where

$$\langle x^n \rangle = \left. \frac{d^n}{d\theta^n} G(\theta) \right|_{\theta=0} \quad (55)$$

To define the cumulants, it is convenient to define the *cumulant generating function*

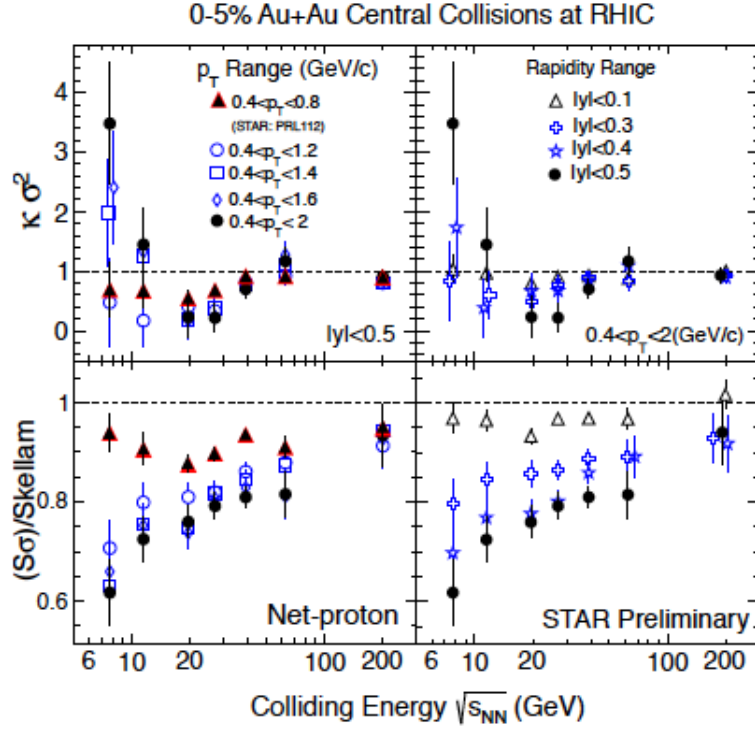
$$K(\theta) = \ln G(\theta). \quad (56)$$

The cumulants of  $\mathcal{P}(x)$  are defined by

$$\langle x^n \rangle_c = \left. \frac{d^n}{d\theta^n} K(\theta) \right|_{\theta=0}. \quad (57)$$

It is then possible to write the cumulants in terms of the moments. For instance

$$\langle x \rangle_c = \langle x \rangle,$$



**Figure 9:** Net proton number cumulants measured by STAR. The figure is from Ref. [45].

$$\begin{aligned}
 \langle x^2 \rangle_c &= \langle x^2 \rangle - \langle x \rangle^2 = \langle \delta x^2 \rangle, \\
 \langle x^3 \rangle_c &= \langle \delta x^3 \rangle, \\
 \langle x^4 \rangle_c &= \langle \delta x^4 \rangle - 3 \langle \delta x^2 \rangle^2.
 \end{aligned} \tag{58}$$

The relation with thermodynamics comes through the partition function  $\mathcal{Z}$ , which is the fundamental object. The partition function is also the moment generating function and therefore the cumulant generating function is given by  $\ln \mathcal{Z}$ .

Cumulants are extensive quantities. Consider the number  $N$  of a conserved quantity in a volume  $V$  in a grand canonical ensemble. It can be shown that its cumulant of order  $n$  can be written as

$$\langle N^n \rangle_{c,V} = \chi_n V, \tag{59}$$

where the quantities  $\chi_n$  are called the *generalized susceptibilities*. From the thermodynamical side, the derivatives of  $\ln \mathcal{Z}$  with respect to the chemical potentials give the susceptibilities. For instance

$$\chi_{BQS}^{ijk} = \frac{\partial^{i+k+j}(P/T^4)}{\partial^i(\mu_B/T) \partial^j(\mu_Q/T) \partial^k(\mu_S/T)}; \quad P = \frac{T}{V} \ln \mathcal{Z}. \tag{60}$$

Also, since cumulants higher than second order vanish for a Gaussian probability distribution, non-Gaussian fluctuations are signaled by non-vanishing higher order cumulants.

Two important higher order moments are the *skewness*  $S$  and the *curtosis*  $\kappa$ . The former measures the asymmetry of the distribution function whereas the latter measures its sharpness. When the stochastic

variable  $x$  is normalized to the square root of the variance,  $\sigma$ , such that  $x \rightarrow \tilde{x} = x/\sigma$ , the skewness and the kurtosis are given as the third and fourth-order cumulants, namely

$$S = \langle \tilde{x}^3 \rangle_c, \quad \kappa = \langle \tilde{x}^4 \rangle_c. \quad (61)$$

When fluctuations of conserved charges in relativistic heavy-ion collisions are well described by hadron degrees of freedom in equilibrium, their cumulants should be consistent with models that describe these degrees of freedom, such as the Hadron Resonance Gas (HRG) model. On the other hand, when fluctuations deviate from those in the HRG model, they can be used as experimental signals of non-hadron and/or non-thermal physics. Figure 9 shows that experimental ratios of some cumulants present statistically significant deviations from the HRG model, though within large errors. Near the CEP, higher order cumulants of conserved charges also behave anomalously. In particular, they change sign in the vicinity of the critical point. They are also sensitive to the increase of correlation lengths [46].

## 5 Experimental signatures of deconfinement

The results from relativistic heavy-ion experiments carried out at the largest available energies have produced a picture of the properties of the created matter at high temperatures, though for small baryon chemical potentials. These properties are extracted from observables optimized to probe the evolution of the system during the different stages of the collision. There are many reviews devoted to the detailed description of such observables and their meaning [47]. Here I content myself with a brief review of some of the main characteristics that have been reported and on how these help us providing a coherent picture of the kind of matter produced in these reactions.

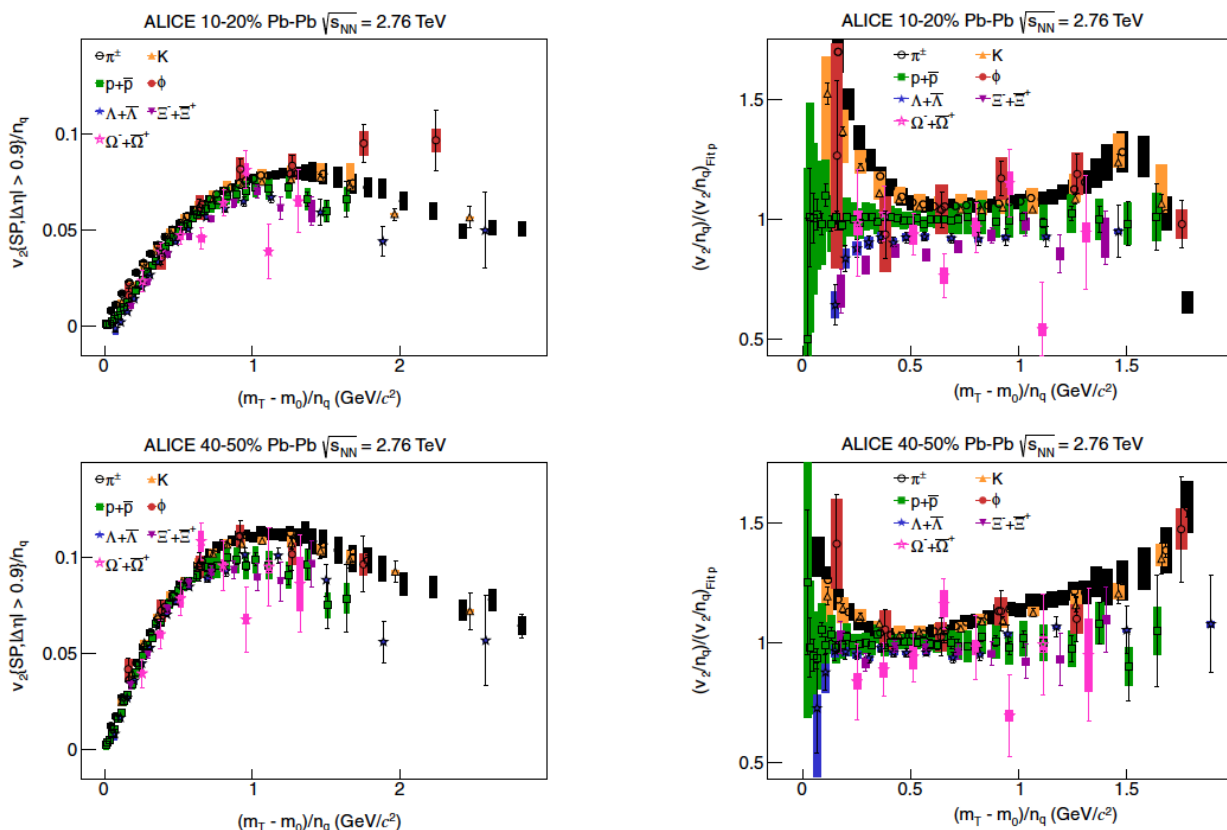
During the initial stages of the collision, dense gluon fields create a strongly interacting medium, the initial state is described by the Color Glass Condensate. This Medium rapidly expands and thermalizes; a Quark Gluon Plasma (QGP) is produced which continues to expand and eventually cools down below  $T_c \simeq 155$  MeV where it hadronizes and becomes a hadron-resonance gas. At a very similar temperature (known as chemical freeze-out temperature  $T^{\text{chem}}$ ), the particle composition is fixed. After chemical freeze-out, particles continue to interact. Only their momentum distributions are affected since their energy is below the inelastic reaction threshold. Hadrons then cease to interact at a kinetic freeze out temperature  $T^{\text{kin}} \simeq 95$  MeV, instant when they have developed a radial flow velocity  $\langle \beta_T \rangle \simeq 0.65$ . A summary of the parameters that characterize the produced medium is as follows:

- Temperature: 100 – 500 MeV.
- Volume:  $1 - 5 \times 10^3$  fm<sup>3</sup>.
- Lifetime: 10 – 20 fm/c.
- Pressure: 100 – 300 MeV/fm<sup>3</sup>.
- ensity:  $1 - 10 \rho_0$  ( $\rho_0 = 0.17$  fm<sup>-3</sup> normal nuclear density).

At low  $p_T < 2$  GeV/c, hydrodynamics provides a good description of this bulk properties. Notice that a large fraction of all particles is produced in this  $p_T$  regime. The produced bulk medium behaves like an almost perfect fluid with a value of shear viscosity to entropy ratio  $\eta/s$  close to its lower theoretical value. The medium is opaque to hard probes, quenching their energy. Radial flow tends to deplete the particle spectrum at low values (blue shift), which increases with increasing particle mass and transverse velocity.

In peripheral collisions, an elliptic flow, characterized by the coefficient  $v_2$  of the the azimuthal angle particle distribution's harmonic expansion, develops. The azimuthal distribution can be expressed as

$$\frac{dN}{d(\phi - \Psi_R)} = N_0 [1 + 2v_1(p_T) \cos(\phi - \Psi_R) + 2v_2(p_T) \cos(2(\phi - \Psi_R)) \dots] \quad (62)$$

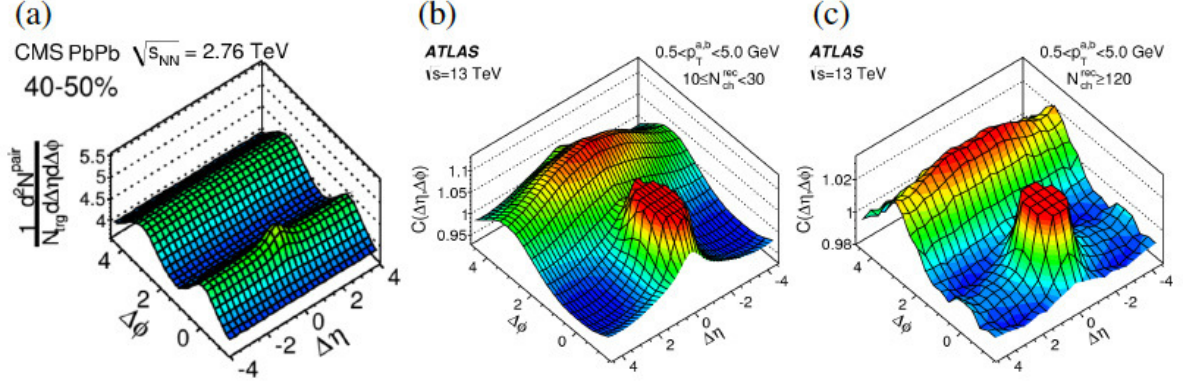


**Figure 10:** Deviations from  $v_2$  scaling with hadron's quark content measured by the ALICE Collaboration. The figure is taken from Ref. [48]

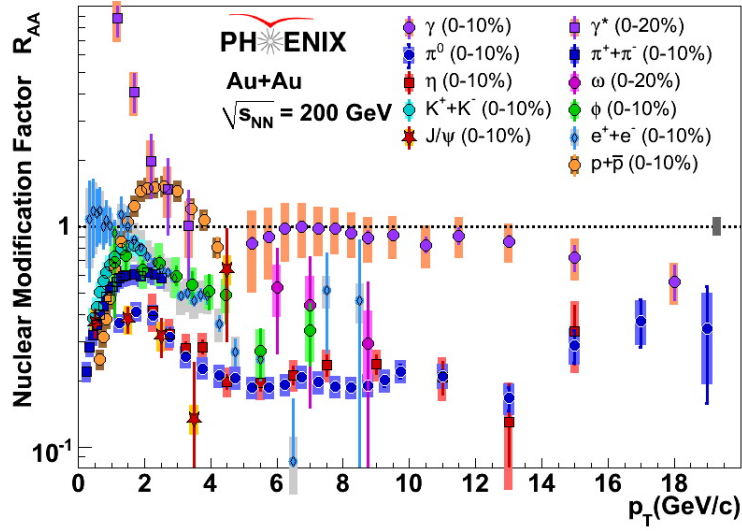
It is believed that  $v_2$  originates from the asymmetric pressure gradients of the initial ellipsoidal overlap region. The strength of this coefficient increases, as expected, with the initial geometric asymmetry from central to peripheral collisions, with maximal value for the centrality range 40%–50%. It is interesting to notice that particle depletion becomes larger in-plane than out-of-plane, thereby reducing  $v_2$ . The net result is that at a fixed value of  $p_T$  heavier particles have smaller  $v_2$  than lighter ones.

At RHIC energies, it was reported that at intermediate  $p_T$  if both  $v_2$  and  $p_T$  were scaled by the number of constituent quarks  $n_q$ , the various identified hadrons follow an approximate common behavior. Scaling was interpreted as a signature that quark coalescence was a dominant hadronization mechanism in this momentum domain and also as the onset of the quark degrees of freedom importance during the early stages of heavy-ion collisions, when collective flow develops. However, recent ALICE data shows that scaling, if any, is only approximate, for all centrality intervals. This is illustrated in Fig. 10.

Angular correlations also provide with an important experimental tool to explore collective phenomena. It has been observed that  $\Delta\eta\Delta\phi$  distributions contain two important features: (1) A peak around  $(\Delta\eta, \Delta\phi) = (0, 0)$  (near side peak from jets) and (2) Long-range correlations called ridges (collective phenomena). This is illustrated in Fig. 11. Similar structures are observed in small reference systems. Possible explanations for these observations are either hydrodynamical behavior and/or gluon saturation of the initial state (CGC).



**Figure 11:**  $\Delta\eta\Delta\phi$  distributions in Pb+Pb collisions measured by the CMS and ATLAS Collaboration. The main features of these correlations are visible: A peak around  $(\Delta\eta, \Delta\phi) = (0, 0)$  coming from the near side jet and long-range correlations called ridges, associate to collective phenomena. The figure is from Ref. [49].

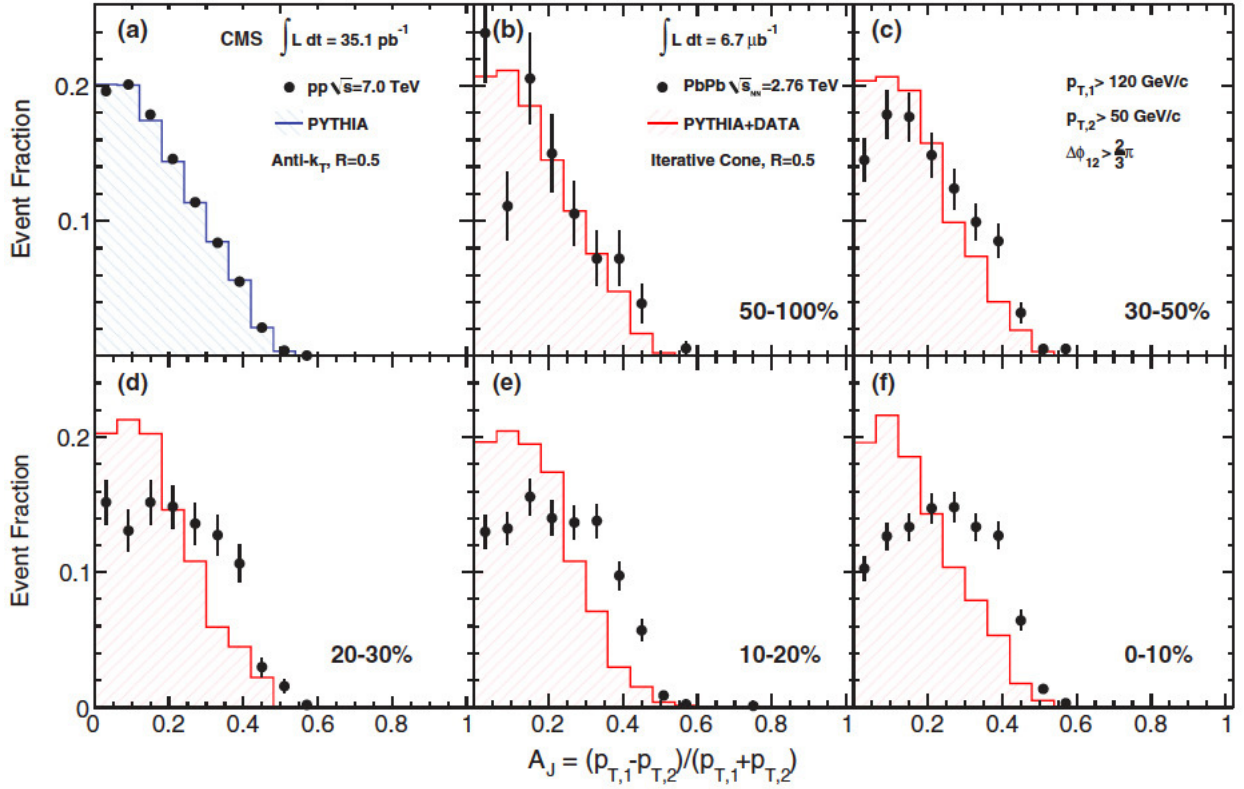


**Figure 12:** Nuclear Modification Factor, for all identified particles measured by PHENIX in central Au+Au collisions at  $\sqrt{s_{NN}}=200$  GeV.

Hard probes such as highly energetic jets of hadrons with heavy flavor content are also a powerful tool to explore the response properties of the medium. An important finding in this context is the observation of the *jet quenching* phenomenon. This is quantified in terms of the *nuclear modification factor*

$$R_{AA}(p_T) = \frac{dN^{AA}(p_T)/dp_T}{\langle N_{col} \rangle dN^{PP}(p_T)/dp_T}. \quad (63)$$

The prototypical behavior of  $R_{AA}$  for several species is illustrated in Fig. 12. Notice that hadrons species show quenching. This quenching is usually attributed to the energy loss suffered by the propagating parton that later on hadronizes. It is important to emphasize that energy loss and elliptic flow are interconnected, that is to say that a clear relationship between jet suppression  $R_{AA}$  and initial nuclear geometry  $v_2$  is observed. This relationship confirms not only the existence of the medium but also the expectation that jet suppression is strongest in the out-of-plane direction where partons traverse the largest amount of hot matter.



**Figure 13:** Asymmetry distribution as a function of jet asymmetry for p+p and Pb+Pb collisions from CMS data. Notice that for the case of Pb+Pb collisions the peak of the distribution moves to larger jet asymmetries as the centrality of the collision increases. The figure is from Ref. [50].

A pertinent question is, where does the quenched jet energy go? The answer can be quantified in terms of the asymmetry distribution as a function of the jet asymmetry variable

$$A_J = \frac{P_{T,1} - P_{T,2}}{P_{T,1} + P_{T,2}}. \quad (64)$$

When no medium modification is present, near and away-side jets carry similar momentum and the asymmetry distribution peaks for  $A_J \sim 0$ . However, when near and away-side jets carry different amounts of energy, the asymmetry distribution does not peak for  $A_J \sim 0$  anymore. This is illustrated in Fig. 13 where data from CMS on p+p and Pb+Pb data are compared to PYTHIA (no medium present) simulations. The data on Pb+Pb collisions clearly show that the peak of the distribution moves to larger jet asymmetries as the centrality of the collision increases.

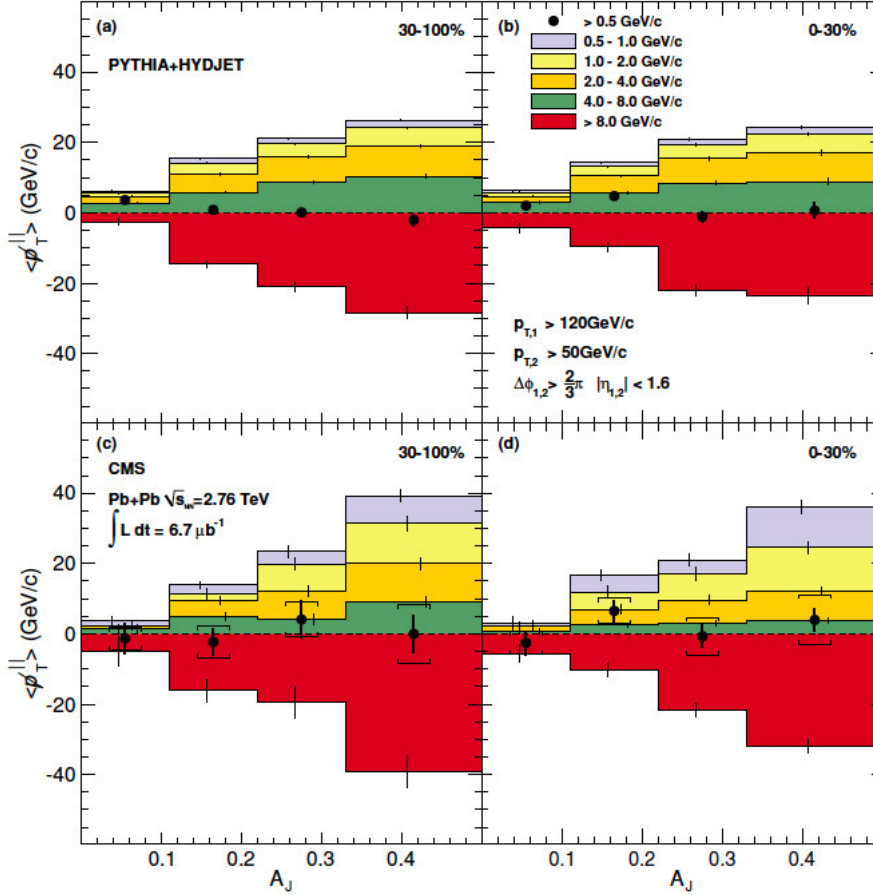
To quantify the amount of missing momentum inside away jet, one defines the average missing  $p_T$

$$\langle p_T^{\parallel} \rangle \equiv \frac{1}{N} \sum_{i \in \text{all } N \text{ tracks}} -p_T^i \cos(\phi_i - \phi_L). \quad (65)$$

The momentum in the away-side is obtained for tracks around the sub-leading jet within a cone aperture larger than the jet cone. Data show that the contribution to the momentum around the leading cone comes mostly from tracks with  $p_T > 8$  GeV. This momentum is balanced by the combined contributions from tracks with  $0.5 < p_T < 8$  GeV outside the away-side jet cone with  $\Delta\phi < \pi/6$ . This is shown in Fig. 14

Finally, we mention that another important tool to probe medium properties is the study of heavy flavors. Heavy flavors are produced by initial hard-scattering processes at time scales of order  $\tau \sim 1/2m_H$  (0.07 fm for charm and 0.02 fm for bottom), which are short compared to QGP formation ( $\tau_0 \sim 0.1 - 1$





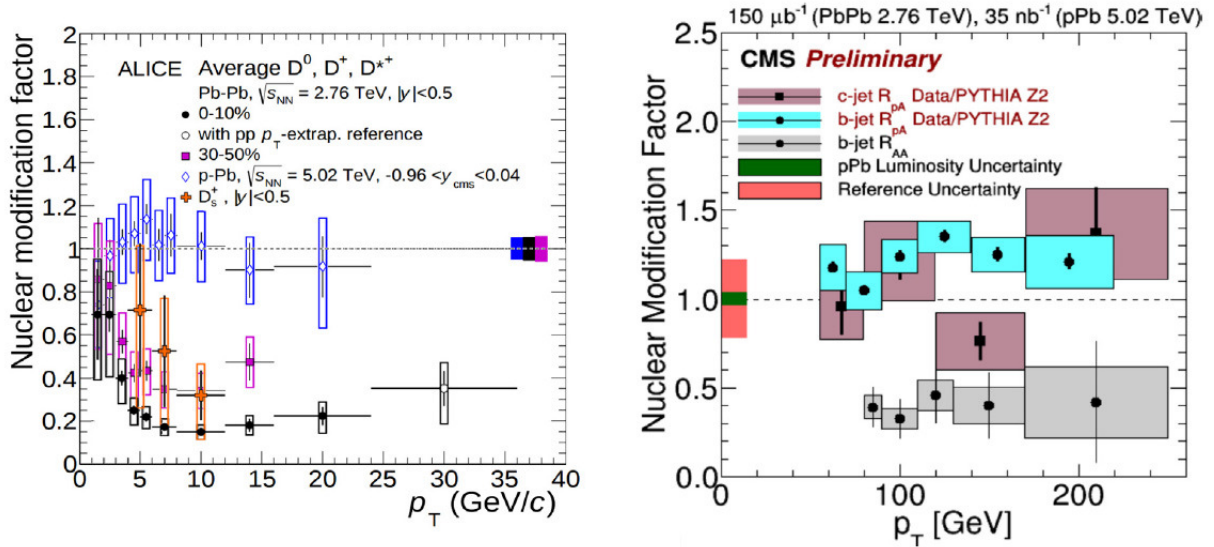
**Figure 14:** Missing  $p_T$ , that is, the sum of the momentum in tracks in the away side with aperture larger than the jet cone. The contribution to the momentum around the leading cone comes mostly from tracks with  $p_T > 8$  GeV and this is balanced by the combined contributions from tracks with  $0.5 < p_T < 8$  GeV outside the away-side jet cone. The figure is from Ref. [50].

fm). Therefore heavy flavors witness the entire medium evolution. Their annihilation rate in the QGP is small, although interaction with the medium may redistribute their momentum. These characteristics make heavy flavors a good probe for medium properties, that is, for the study of transport coefficients. At the LHC, the production cross section is much larger than at RHIC, thus heavy flavors can be studied more systematically.

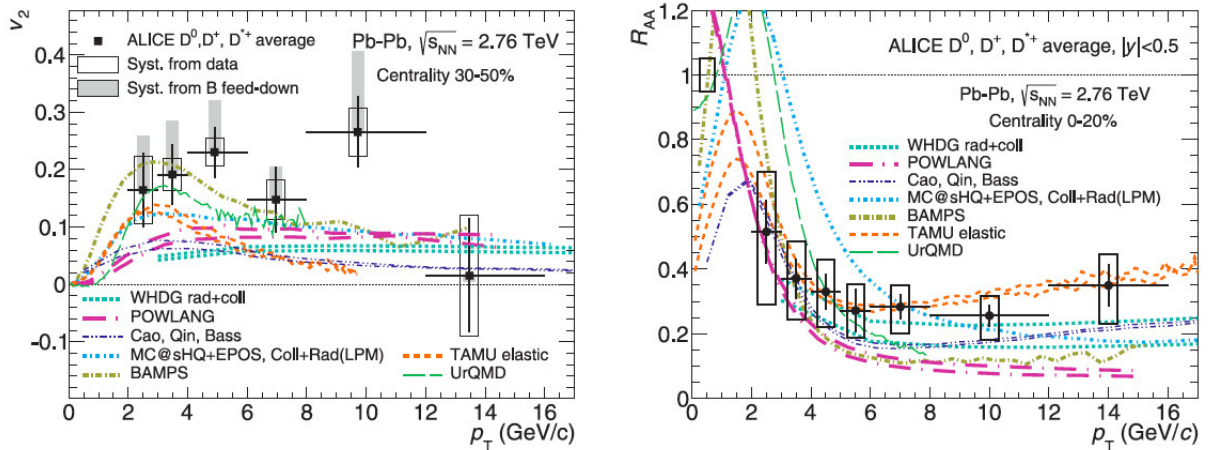
Figure 15 shows the quenching pattern for open-flavor heavy-mesons. Notice that in nuclear collisions, the suppression is as strong as for the case of light flavors. This is contrary to the case of p+Pb collisions. A large suppression indicates a strong heavy-flavor coupling with medium. Pure energy loss predicts effects predict a hierarchy in the suppression pattern,  $R_{AA}^{\text{light}} < R_{AA}^D < R_{AA}^B$ . However, one should be cautious since there are a number of effects that may alter such suppression pattern, for instance the differences between primordial spectral shapes of produced partons and their fragmentation functions; the differences between the kinds of processes of flavor production (recall that light flavors are mainly produced by soft processes, whereas heavies are produced by hard processes), etc. The observed agreement  $R_{AA}(D) \simeq R_{AA}(\pi)$  is reproduced by models that include different fragmentation functions and shapes of the primordial  $p_T$  distributions, in addition to the expected energy loss hierarchy. On the other hand, a comparison of  $R_{AA}(D)$  and  $R_{AA}(J/\psi)$  shows the expected suppression pattern.

Heavy-flavor hadrons can share the medium azimuthal anisotropy quantified by  $v_2$ . Data show large  $v_2$  of charm (same magnitude as  $v_2$  of light-hadrons) which implies that charm thermalizes in





**Figure 15:**  $R_{AA}$  for prompt  $D$  mesons as a function of  $p_T$  for Pb+Pb collisions compared to p+p collisions obtained by ALICE (left) and CMS (right) Collaborations. Notice that in nuclear collisions, the heavy-flavor  $R_{AA}$  is comparable to the light flavor one and that in p+Pb collisions it is comparable to 1. The figures are from Refs. [51, 52].



**Figure 16:** Prompt  $D$  mesons  $v_2$  for Pb+Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV measured by the ALICE Collaboration. Figures are taken from Ref. [53]

medium. This is illustrated in Fig.16.

As a final remark, we notice that the simultaneous measurements of  $R_{AA}$  and  $v_2$  help to disentangle the interplay of different energy loss scenarios and imposes constraints on theoretical models.

## 6 Conclusions

The heavy-Ion Standard Model is being developed as we speak. For this purpose there is a strong synergy between experiment and theory. Experimental measurements pose many theoretical challenges and rise questions stimulating progress. The field represents a rich and diversity field of approaches. Semi-classical gauge theory for initial conditions, LQCD for static thermodynamic properties, perturbative QCD in vacuum and in-medium, transport theory and particularly viscous hydrodynamics for the evolution of bulk matter and even holographic methods can be employed to describe the dynamics of

thermalization. There is also a large variety of open problems in different fronts: Thermal photon puzzle, extraction of transport coefficients, interplay between hard and soft modes, limit of applicability of hydro approach and inclusion of bulk viscosity in 3D calculations, role of magnetic fields in peripheral collisions, critical point of phase diagram, etc., are only some of the questions that need attention. Overall this is an exciting field with many opportunities to continue exploring the properties of QCD matter under extreme conditions.

## References

### Bibliography

- [1] This section is largely based on H. Sazdjian, Introduction to chiral symmetry in QCD, arXiv:1612.04078.
- [2] M. S. Bhagwat, M. A. Pichowsky, C. D. Roberts, P. C. Tandy, Phys. Rev. C **68**, 015203 (2003).
- [3] P. Pascual and R. Tarrach (1984), QCD: Renormalization for the practitioner. Springer Berlin Heidelberg.
- [4] Michael E. Peskin and Daniel V. Schroeder (1995), An Introduction to Quantum Field Theory, Addison-Wesley Publishing Company, Massachusetts.
- [5] Fritz John (1971). Partial Differential Equations. Springer-Verlag New York.
- [6] See for example: A. Andronic, P. Braun-Munzinger, J. Stachel, Nucl. Phys. A **772**, 167-199 (2006).
- [7] S. K. Tiwari, C. P. Singh, Adv. High Energy Phys. **2013**, 805413 (2013).
- [8] W. Broniowski, W. Florkowski, L. Ya. Glozman, Phys. Rev. D **70**, 117503 (2004).
- [9] X. Shu-Sheng, S. Yuan-Mei, Y. You-Chang, C. Zhu-Fang and Z. Hong-Shi, Chin. Phys. Lett. **32**, 121203 (2015).
- [10] For a review of these results see L. Levkova, PoS Lattice **2011**, 011 (2011).
- [11] C. Bernard *et al.* (MILC Collaboration), Phys. Rev. D **71**, 034504 (2005).
- [12] M. Cheng *et al.*, Phys. Rev. D **74**, 054507 (2006).
- [13] S. Borsányi *et al.*, J. High Energy Phys. 1009.073 (2010); Y. Aoki *et al.*, J. High Energy Phys. 0906.088 (2009); Y. Aoki *et al.*, Phys. Lett. B **643**, 46 (2006).
- [14] A. Bazavov, PoS Lattice **2011**, 182 (2011); A. Bazavov *et al.*, Phys. Rev. D **85**, 054503 (2012).
- [15] For a recent review see P. de Forcrand, PoS Lattice **2009**, 010 (2009).
- [16] G. Aarts, PoS Lattice **2012**, 017 (2012).
- [17] Z. Fodor, C. Guse, S. D. Katz, K. K. Szabo, PoS Lattice **2007**, 189 (2007); O. Kaczmarek, F. Karsch, E. Laermann, C. Miao, S. Mukherjee, P. Petreczky, C. Schmidt, W. Soeldner, W. Unger, Phys. Rev. D **83**, 014504 (2011); G. Endrödi, Z. Fodor, S. D. Katz, K. K. Szabo, J. High Energy Phys. **1104**, 001 (2011); E. Laermann, F. Meyer, M. P. Lombardo, J. Phys. Conf. Ser. **432**, 012016 (2013); P. Cea, L. Cosmai, A. Papa, Phys. Rev. D **89**, 074512 (2014); C. Bonati, M. D'Elia, M. Mariti, M. Mesiti, F. Negro, F. Sanfilippo, Phys. Rev. D **90**, 114025 (2014); C. Bonati, M. D'Elia, M. Mariti, M. Mesiti, F. Negro, F. Sanfilippo, Phys. Rev. D **92**, 054503 (2015); R. Bellwied, S. Borsanyi, Z. Fodor, J. Günther, S. D. Katz, C. Ratti, K. K. Szabo, Phys. Lett. B **751**, 559-564 (2015); P. Cea, L. Cosmai, A. Papa, Phys. Rev. D **93**, 014507 (2016); P. Hegde and H.-T. Ding (for the Bielefeld-BNL-CCNU Collaboration), PoS Lattice **2015**, 141 (2016).
- [18] J. Cleymans, H. Oeschler, R. Redlich, S. Weaton, J. Phys. G **32**, S165 (2006).
- [19] M. Asakawa and K. Yazaki, Nucl. Phys. A **504**, 668 (1989); A. Barducci, R. Casalbuoni, S. De Curtis, R. Gatto and G. Pettini, Phys. Lett. B **231**, 463 (1989); Phys. Rev. D **41**, 1610 (1990); A. Barducci, R. Casalbuoni, G. Pettini and R. Gatto, Phys. Rev. D **49**, 426 (1994); J. Berges and K. Rajagopal, Nucl. Phys. B **538**, 215 (1999); M. A. Halasz, A. D. Jackson, R. E. Shrock, M. A. Stephanov and J. J. M. Verbaarschot, Phys. Rev. D **58**, 096007 (1998); O. Scavenius, A. Mocsy,

- I. N. Mishustin and D. H. Rischke, Phys. Rev. C **64**, 045202 (2001); N. G. Antoniou and A. S. Kapoyannis, Phys. Lett. B **563**, 165 (2003); Y. Hatta and T. Ikeda, Phys. Rev. D **67**, 014028 (2003).
- [20] L. Adamczyk *et al.* [STAR Collaboration], Phys. Rev. Lett. **112**, 032302 (2014); Phys. Rev. Lett. **113**, 092301 (2014).
- [21] C. Yang [for the STAR Collaboration], Nucl. Phys. A **967**, 800-803 (2017).
- [22] P. Senger, J. Phys. Conf. Ser. **798**, 012062 (2017).
- [23] V. Kekelidze, A. Kovalenko, R. Lednicky, V. Matveev, I. Meshkov, A. Sorin, G. Trubnikov, Nucl. Phys. A **967**, 884-887 (2017).
- [24] P. Costa, M. C. Ruivo, and C. A. de Sousa, Phys. Rev. D **77**, 096001 (2008).
- [25] A. Ayala, A. Bashir, C. A. Dominguez, E. Gutierrez, M. Loewe, A. Raya, Phys. Rev. D **84**, 056004Z (2011).
- [26] X.-Y. Xin, S.-X. Qin, Y.-X. Liu, Phys. Rev. D **90**, 076006 (2014)
- [27] C. S. Fischer, J. Luecker and C. A. Welzbacher, Phys. Rev. D **90**, 034022 (2014).
- [28] Y. Lu, Y.-L. Du, Z.-F. Cui, H.-S. Zong, Eur. Phys. J. C **75**, 495 (2015).
- [29] C. Shi, Y.-L. Du, S.-S. Xu, X.-J. Liu, H.-S. Zong, Phys. Rev. D **93**, 036006 (2016).
- [30] G. A. Contrera, A. G. Grunfeld, D. Blaschke, Eur. Phys. J. A **52**, 231 (2016).
- [31] Z.-F. Cui, J.-L. Zhang, H.-S. Zong, Sci. Rep. **7**, 45937 (2017).
- [32] S. Datta, R. V. Gavai, S. Gupta, Phys. Rev. D **95**, 054512 (2017).
- [33] T. Yokota, T. Kunihiko and K. Morita, arXiv:1611.06669 [hep-ph].
- [34] S. Sharma [Bielefeld-BNL-CCNU Collaboration], Nucl. Phys. A **967**, 728 (2017).
- [35] P. Kovács and G. Wolf, Acta Phys. Polon. Supp. **10**, 1107 (2017).
- [36] J. Knaute, R. Yaresko and B. Kämpfer, arXiv:1702.06731 [hep-ph].
- [37] N. G. Antoniou, F. K. Diakonov, X. N. Maintas and C. E. Tsagkarakis, arXiv:1705.09124 [hep-ph].
- [38] R. Rougemont, R. Critelli, J. Noronha-Hostler, J. Noronha and C. Ratti, Phys. Rev. D **96**, no. 1, 014032 (2017).
- [39] P. Cea, L. Cosmai, A. Papa, Phys. Rev. D **93**, 014507 (2016); C. Bonati, M. D'Elia, M. Mariti, M. Mesiti, F. Negro, F. Sanfilippo, Phys. Rev. D **92**, 054503 (2015).
- [40] R. Bellwiede, S. Borsanyi, Z. Fodor, J. Gäjnthner, S. D. Katz, C. Ratti, K. K. Szabo, Phys. Lett. B **751**, 559-564 (2015).
- [41] A. Bazavov, *et al.*, Phys. Rev. D **95**, 054504 (2017).
- [42] C. Schmidt and S. Sharma, J. Phys. G **44**, 104002 (2017).
- [43] A. Ayala, L. Hernandez, S. Hernandez-Ortiz, Rev. Mex. Fis. **64**, 302-313 (2018).
- [44] M. Asakawa, M. Kitazawa, Prog. Part. Nucl. Phys. **90**, 299-342 (2016).
- [45] X. Luo [STAR Collaboration], PoS CPOD 2014, 019 (2015).
- [46] M. A. Stephanov, Phys. Rev. Lett. **102**, 032301 (2009).
- [47] P. Foka and M. A. Janik, Rev. Phys. **1**, 154 (2016); *ibid* **1**, 172 (2016).
- [48] F. Noferini (for the ALICE Collaboration), Eur. Phys. Jour. Web of Conf. **90**, 08005 (2015).
- [49] S. Chatrchyan, *et al.*, Eur. Phys. J. C **72** (2012) 2012.
- [50] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Rev. C **84**, 024906 (2011).
- [51] J. Adam, D. Adamová *et al.* (ALICE Collaboration), J. High Energy Phys. **2016**, 81 (2016).
- [52] QM2015, Nucl. Phys. A **956**, 1âŠš974 (2015).
- [53] N. Armesto, E. Scapparini, Eur. Phys. J. Plus **131** (3), 52 (2016).



## A Sketchy Introduction to Cosmology

*Rogério Rosenfeld*

ICTP South American Institute for Fundamental Research & Instituto de Física Teórica,  
Universidade Estadual Paulista, São Paulo, Brazil &  
Laboratório Interinstitucional de e-Astronomia, Rio de Janeiro, Brazil

### 1 Introduction

Why should a particle physicist learn cosmology? I can think of several reasons:

- the main evidence from new physics beyond the Standard Model comes from cosmology: eg dark matter, dark energy and inflation;
- particle physics affect cosmology: eg origin of matter- anti-matter asymmetry (needs baryon number violation, CP violation, non-equilibrium processes), the Higgs field can have a role in inflation, neutrinos have a role in the formation of structures (such as galaxies and galaxy clusters) in the Universe, cosmological phase transitions due to particle physics (electroweak breaking, QCD confinement, etc) can occur in the early Universe with observable consequences (such as the generation of gravitational waves, topological defects, etc) ;
- cosmology affects particle physics: eg evolution of the Universe may be responsible for electroweak symmetry breaking (eg relaxion idea);
- early Universe is a testbed for SM and BSM: eg stability or metastability of SM vacuum;
- gravity (geometry) may play an important role in particle physics: eg models with warped extra dimensions
- new particles from geometry: KK excitations, radion, etc;
- models with extra dimensions can change the evolution of the Universe (and hence be tested).

Now that I hope to have convinced you of the importance to study cosmology with all these very interesting topics I'll unavoidably disappoint you since I will not be able to cover in much detail any of them - but my intention is to give a short start so that you can continue on your own.

I must say right away that in my opinion a full-fledged write-up of these lectures, which still need much improvement, is not essential at this point. In fact there are several excellent books and lectures on Cosmology among which I list:

- My favourite cosmology book, even if outdated, still is “The Early Universe” by Kolb and Turner [1]. I guess this is for sentimental reasons, since I learned the subject from it. Dodelson’s “Modern Cosmology” is more recent and highly recommended [2];
- Lectures by Daniel Baumann at Cambridge [3] - I borrowed extensively from these lectures;
- Lectures by Pierre Binétruy at the 2012 European CERN School [4];
- Lectures by Toni Riotto at the 2010 European CERN School [5];
- Lectures on Dark Matter by Graciela Gelmini [6] and Mariangela Lisanti [7].

Hence these notes are intended as just a brief guide to what was discussed in the lectures. In the following I will present a simple sketch of my lectures pointing to some references where more details can be found. The slides of my lectures, as well as for the other lectures of the School can be found in [8].

Cosmology has become a precision, data-driven science in the last 20 years or so. I remember that when I started to read about it (admittedly a long time ago) the age of the universe was written as  $t_0 = 10^{9\pm 1}$  years. The uncertainty was in the exponent! It is fantastic how measurements coming from different observables now determine  $t_0 = (13.799 \pm 0.021) \times 10^9$  years.

There are several observational probes that are used to find out what is the best cosmological model that describes our universe. These include:

- the cosmic microwave background (CMB);
- the abundance of light elements, as described by Big Bang Nucleosynthesis (BBN);
- the use of supernovae of type Ia as standard candles to measure distances in the universe;
- the study of the Large Scale Structure (LSS) of the Universe, particularly in the use of a feature in the distribution of galaxies called Baryon Acoustic Oscillation (BAO) as a standard ruler;
- the use of weak gravitational lensing, small distortions in the shape of galaxies, for the determination of the distribution of matter in the universe;
- the use of counts of galaxy clusters as a measure of the growth of perturbations in the universe.

The picture that has emerged from all these observations is consistent and somewhat disturbing: we know that we don't know what 95% of the universe is made of. Of the cosmic energy-density budget, roughly only 5% is matter that we know and love - atoms. The rest we believe is dark matter (roughly 25%) and dark energy (roughly 70%). The best model that describes our universe is a model where dark matter is made of nonrelativistic particles (called cold dark matter) and dark energy is described by a simple cosmological constant (denoted by  $\Lambda$ ). The so-called  $\Lambda$ CDM model became the Standard Cosmological Model.

As you all know, particle physics also has a Standard Model (SM) - the  $SU(3)_C \times SU(2)_L \times U(1)_Y$  Glashow-Weinberg-Salam model, where the  $SU(2)_L \times U(1)_Y$  symmetry is spontaneously broken to electromagnetism  $U(1)_{EM}$  through the Higgs mechanism. The last missing piece of the model, the Higgs boson was finally detected at the LHC in 2012. The SM of particle physics has been tremendously successful - it can actually explain all the data measured at accelerators given some input parameters. However, it is not satisfactory since it does not describe neutrino masses, it has no candidate for cold dark matter, it suffers from the so-called hierarchy problem (simply stated, why is the Higgs mass much smaller than a high energy scale where new physics reside when quantum corrections are taken into account), among other issues. This has led to the development of new physics models, extensions of the SM called Beyond SM (BSM), where these problems can be addressed.

The Standard Cosmological Model is also unsatisfactory. We do not know what dark matter and dark energy are. There are several alternatives to the cold dark matter paradigm: warm dark matter, fuzzy dark matter, self-interacting dark matter, modified newtonian dynamics, TeVeS, etc. The cosmological constant suffers from the same hierarchy problem as the Higgs mass, only worse since the sensitivity to a new physics scale is quartic instead of quadratic. For dark energy there are even more exotic alternatives that go by names such as quintessence, Horndesky, massive gravity, clustered dark energy, interacting dark energy, etc.

One important development that happened after the School was the observation by LIGO and LISA on August 17, 2017 of the gravitational waves produced by the fusion of a pair of neutron stars, with an electromagnetic counterpart identified by several observatories in different wavelengths. The time difference between the gravitational and electromagnetic waves from the event was less than 1.7 seconds. Given the distance of 130 million of light-years this measurement puts a very stringent bound on the difference of the velocities of the two types of waves of one part in  $10^{15}$  - we can say that gravitational waves propagate with the speed of light. This simple new observational fact has eliminated several models of dark energy (see, eg [9]). I also should mention that the 2017 Nobel prize was awarded to Weiss, Barish and Thorne "for decisive contributions to the LIGO detector and the observation of gravitational waves".

For these lectures I mostly concentrate on the Standard Cosmological Model.

## 2 First lecture: The averaged Universe

### 2.1 The basics

The Universe is ruled primordially by gravitation. Gravity is described by Einstein's General Relativity, which relates a geometry determined by a metric to the content of the Universe as described by an energy-momentum tensor. In 1917 Einstein applied his new theory to describe the whole Universe. It's complexity can be domesticated under the assumption that the Universe is on average homogeneous and isotropic. This assumption, sometimes called the "cosmological principle", leads to a considerable simplification since the metric in this case is the so-called Friedmann-Lemaitre-Robertson-Walker (FLRW) metric:

$$ds^2 = dt^2 - a(t)^2 d\vec{x}^2. \quad (1)$$

The FLRW metric is determined by the scale factor  $a(t)$  which controls the evolution of the averaged Universe. By definition the scale factor is set to one today ( $a(t_0) = 1$ ). In fact the FLRW allows for the introduction of a spatial curvature that also impacts the evolution of the Universe. It is remarkable that observations in the last 20 years or so have measured that the curvature is consistent with zero with small errors: our Universe is on average spatially flat to a very good degree and hence I'm not considering curvature in these lectures.

The expansion rate of the Universe is defined as the Hubble factor:

$$H(t) = \frac{\dot{a}}{a}, \quad (2)$$

where the dot denotes time derivatives. The Hubble constant is the Hubble factor today ( $H_0 = H(t_0)$ ). The acceleration of the Universe is set by  $\ddot{a}$ . The redshift  $z$ , given by the relative change in the position of spectral lines due to the expansion of the Universe, is related to the scale factor by:

$$a = \frac{1}{1+z}. \quad (3)$$

One describes the matter-energy content of the average Universe by a perfect homogeneous fluid energy-momentum tensor which in a rest-frame is characterized only by its energy density ( $\rho$ ) and the pressure ( $p$ ) can be written as

$$T_{\nu}^{\mu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & -p & 0 & 0 \\ 0 & 0 & -p & 0 \\ 0 & 0 & 0 & -p \end{pmatrix} \quad (4)$$

then Einstein's equation reduces to the two well-known Friedmann's equations:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho \quad (5)$$

and

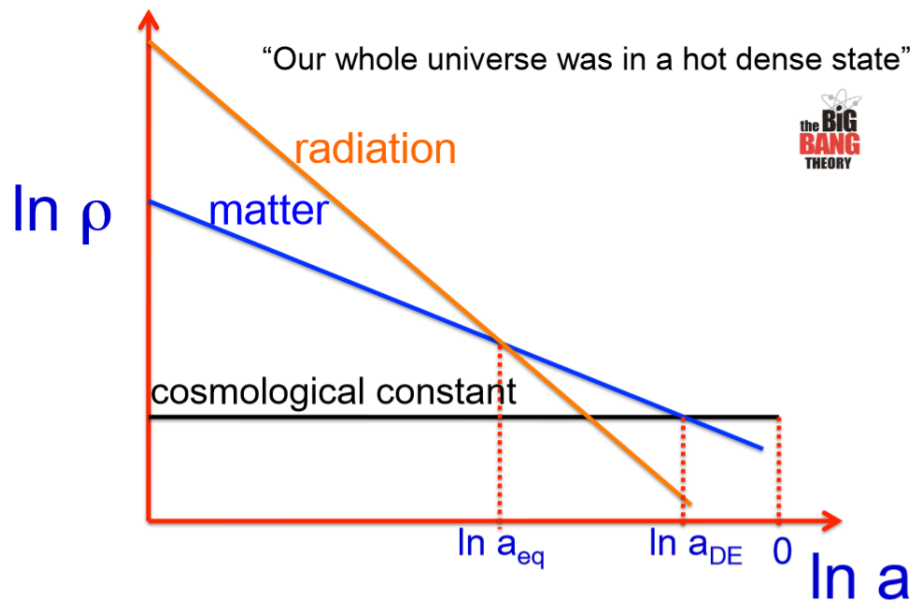
$$\left(\frac{\ddot{a}}{a}\right) = -\frac{4\pi G}{3}(\rho + 3p) \quad (6)$$

One can obtain a continuity equation by taking the time derivative of the first equation and substituting in the second one:

$$\dot{\rho} + 3H(\rho + p) = 0 \quad (7)$$

In order to study the evolution of a fluid in the Universe we must assume an *equation of state*: a relation between its pressure and energy density. The simplest form is

$$p = \omega\rho, \quad (8)$$



**Fig. 1:** Evolution of densities in the Universe as a function of the scale factor for different components.

where  $\omega$  goes by the name of equation of state parameter. For a constant  $\omega$  it is easy to show that the energy density changes as

$$\rho(t) = \rho(t_i) \left( \frac{a(t)}{a(t_i)} \right)^{-3(1+\omega)} \quad (9)$$

and therefore for nonrelativistic fluid ( $\omega = 0$ ), relativistic fluid ( $\omega = 1/3$ ) and for a cosmological constant ( $\omega = -1$ ) one has  $\rho \propto a^{-3}$ ,  $\rho \propto a^{-4}$  and  $\rho \propto a^0$  respectively.

These simple considerations lead to the conclusion that the early universe is very dense and hot (dominated by radiation) and the late universe is dominated by the cosmological constant. A sketch of this behaviour is shown in Figure 1.

It's also easy to show that since  $\dot{a}/a \propto \sqrt{\rho}$  the scale factor grows with time as:

$$a(t) \propto \begin{cases} t^{2/3} & \text{for matter} \\ t^{1/2} & \text{for radiation} \\ e^{Ht} & \text{for a cosmological constant} \end{cases} \quad (10)$$

The Universe is now dominated by a cosmological constant and hence it is expanding exponentially. We don't feel it for the same reason that we don't feel our money growing exponentially in a savings account.

The average Universe being spatially flat is a consequence of its density being equal to the so-called critical density:

$$\rho_c = \frac{3H_0^2}{8\pi G}. \quad (11)$$

The Hubble constant has been measured with a few percent accuracy recently and there is a  $3\sigma$  tension between the value measured in the local Universe and the value extracted from CMB measurements [10]:

$$\begin{aligned} H_0 &= 73.24 \pm 1.74 \text{ km/s/Mpc (local)} \\ H_0 &= 69.3 \pm 0.7 \text{ km/s/Mpc (Planck)}. \end{aligned} \quad (12)$$



The jury is still out whether this tension merits some explanation from new physics. In any case, the critical density amounts to something equivalent to only five hydrogen atoms per cubic meter. The average Universe is a pretty empty place.

One quantifies the composition of the Universe by the ratios  $\Omega_i$  between the energy density in a given component  $i$  to the total critical density:  $\Omega_i = \rho_i/\rho_c$ . For a spatially flat Universe:

$$\sum_i \Omega_i = 1 \quad (13)$$

These quantities are time-dependent and one can re-write the first Friedmann equation as:

$$\left(\frac{H}{H_0}\right)^2 = \sum_i \Omega_i^{(0)} a^{-3(1+w_i)}. \quad (14)$$

It's worth remarking again that the equations derived so far are valid for constant  $\omega$ 's - it is not difficult to generalize them for time-varying equation of state. This is the case for the so-called models of Dynamical Dark Energy (DDE), where for instance Dark Energy can be modelled by a dynamical scalar field. The energy density and pressure for a spatially homogeneous, time-dependent, canonically normalized scalar field  $\phi$  with a potential  $V(\phi)$  is easily computed from its energy-momentum tensor and the equation of state is given by:

$$\omega = \frac{p}{\rho} = \frac{\dot{\phi}^2/2 - V(\phi)}{\dot{\phi}^2/2 + V(\phi)} \quad (15)$$

and hence  $-1 \leq \omega \leq 1$ . Given initial conditions and a potential one can devise DDE models that can easily mimick the background expansion.

## 2.2 Distances in the Universe

In order to get observational information about the evolution of the Universe it is of fundamental importance to be able to measure distances. There are basically two ways to measure distances in the Universe: using standard candles or using standard rulers. Standard candles are objects of known intrinsic brightness, or luminosity. Cepheid stars and supernovas of type Ia are objects that can be calibrated into standard candles. Standard rulers are features in the Universe that have a well-defined physical length, such as the scale of the so-called Baryon Acoustic Oscillation (BAO), which I will discuss later on.

Let's introduce some different distances that will be useful in the following.

(a) Comoving distance between us and an object at redshift  $z$  ( $\chi(z)$ ):

This is defined as a distance  $\chi(z)$  that a light ray would travel

$$ds^2 = 0 \implies dt^2 = a(t)^2 d\chi^2 \quad (16)$$

and with some change of variables it is easy to show that (we are setting  $c = 1$ )

$$\chi(z) = \int_0^z \frac{dz}{H(z)} \quad (17)$$

(b) Comoving particle horizon ( $\chi(z)_{hor}$ ):

Largest region in causal contact since the Big Bang - it's given by

$$\chi(z)_{hor} = \int_0^t \frac{dt'}{a(t')} \quad (18)$$

(c) Luminosity distance ( $d_L$ ):

This is the distance measured using standard candles and is given in terms of the flux of photons received in a detector and the intrinsic luminosity of the source

$$F = \frac{L}{4\pi d_L^2} \quad (19)$$

and it can be shown that it is related to the comoving distance by

$$d_L = (1+z)\chi(z) \quad (20)$$

The luminosity distance is larger for a given redshift for larger values of the dark energy density. This can be traced back to the accelerated expansion of the Universe due to dark energy. Therefore SNIa's look dimmer in a Universe with dark energy. This is how the accelerated expansion of the Universe was discovered in 1998, which was a big surprise (actually a shock) for the community. The 2011 Nobel prize in Physics was awarded for "the discovery of the accelerated expansion of the Universe through observations of distant supernovae".

(d) Angular diameter distance ( $d_A$ ):

This distance is related to the angle  $\delta\theta$  subtended by a physical length  $l$  ( $d_A = l/\delta\theta$ ) and it can be shown that

$$d_A = \frac{1}{1+z}\chi(z). \quad (21)$$

(e) Comoving Hubble radius ( $r_H$ ):

Comoving distance that particles can travel in a Hubble time, sometimes known as the Hubble comoving horizon

$$r_H = \frac{1}{aH} \quad (22)$$

and therefore

$$r_H(t) \propto \begin{cases} a^{1/2} & \text{for matter} \\ a & \text{for radiation} \\ 1/a & \text{for a cosmological constant} \end{cases} \quad (23)$$

It's interesting to study how a comoving scale compares with the comoving Hubble radius - we sketch in Figure 2 how a given scale can enter and exit the Hubble horizon during the evolution of the Universe.

### 2.3 The thermal history of the Universe

I showed above that for radiation the energy density goes as  $\rho_{rad} \propto a^{-4}$ . However, from the Stefan-Boltzmann law  $\rho_{rad} \propto T^4$ , where  $T$  is the temperature of the radiation. Therefore one finds a relation between the scale factor and the temperature of the radiation:  $a \propto T^{-1}$  - the Universe cools down as it expands. Since we also know that  $a \propto t^{1/2}$  in a radiation dominated Universe, one finds that

$$T(\text{MeV}) \approx 1.5g_*^{-1/4}t(s)^{-1/2} \quad (24)$$

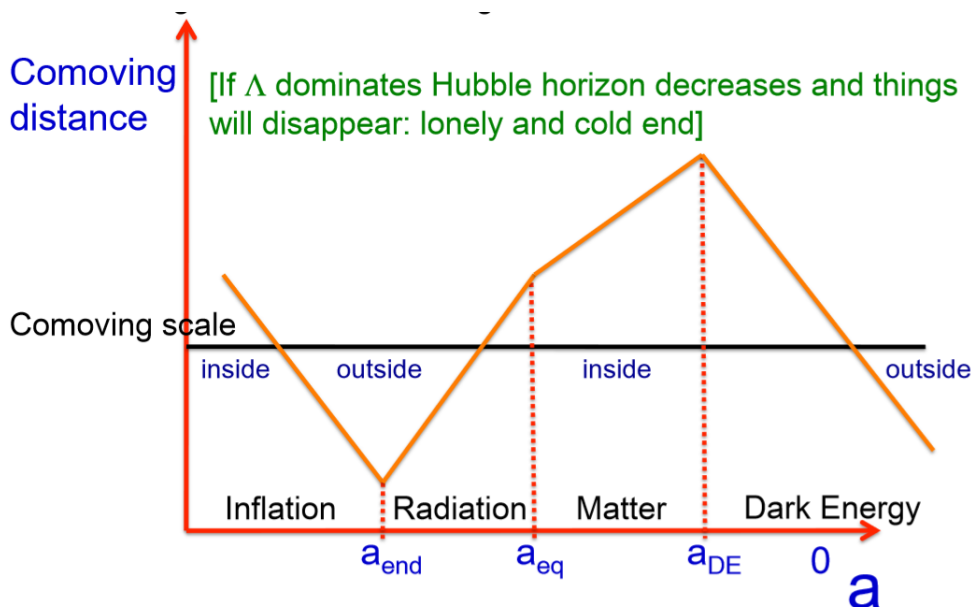
where  $g_*$  measures the number of relativistic degrees of freedom in thermal equilibrium, the temperature is given in MeV and the time is given in seconds.

For a gas of nonrelativistic matter with mass  $m$  and temperature  $T$  ( $T \ll m$ ) the energy density is exponentially suppressed (we are using unity for the Boltzmann constant):

$$\rho \propto m(mT)^{3/2}e^{-m/T}. \quad (25)$$

When the Universe cools down and particles become nonrelativistic the abundance of these particles is exponentially suppressed and they drop out of thermal equilibrium. The reason is the rate of interactions ( $\Gamma$ ) which is given by

$$\Gamma(T) = n(T)\langle\sigma v\rangle, \quad (26)$$



**Fig. 2:** Comoving Hubble radius during the evolution of the Universe compared to a given comoving scale.

where  $n(T) = \rho(t)/m$  is the number density of the particles at temperature  $T$  and  $\langle\sigma v\rangle$  is the thermally averaged cross section times velocity. This interaction rate should be compared with the expansion rate of the Universe given by the Hubble parameter  $H(T)$ . When the interaction rate is much larger than the expansion rate the interactions have time to bring the particles to thermal equilibrium. But as the Universe cools down, the interaction rate drops and eventually the particles will get out of thermal equilibrium and their number will then freeze since the interactions are no longer efficient. This happens at the so-called freeze-out or decoupling temperature. A rough estimate of the freeze-out temperature  $T_{f.o.}$  is obtained from equating  $\Gamma(T_{f.o.}) = H(T_{f.o.})$ . A more precise estimate follows from an explicit solution of the Boltzmann equation.

A simple example is the decoupling of neutrinos in the early Universe. Neutrinos are kept in equilibrium by the weak interactions, eg  $\nu_e \bar{\nu}_e \leftrightarrow e^+ e^-$  and the cross section is roughly given by  $\sigma \approx G_F T^2$  from dimensional analysis ( $G_F \approx 10^{-5} \text{ GeV}^{-2}$  is the Fermi constant). Their number density decreases as  $n_\nu \propto a^{-3} \propto T^3$  and therefore  $\Gamma \approx G_F T^5$ . On the other hand, the Hubble parameter is given by

$$H(T) \approx T^2/M_{Pl}, \quad (27)$$

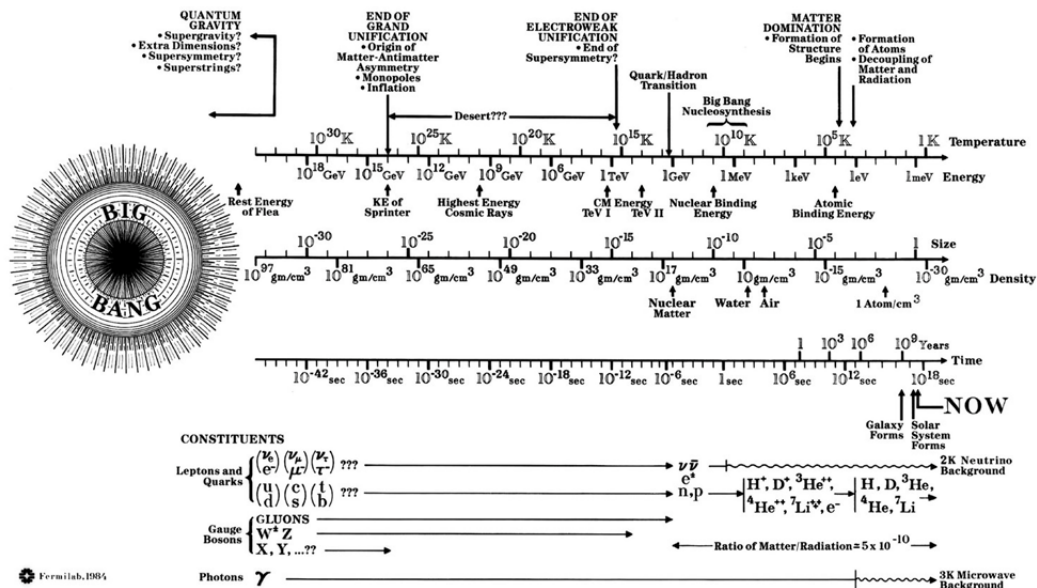
where we used Friedmann's equation with  $\rho \propto T^4$  and that  $G = 1/M_{Pl}^2$ . Therefore the freeze-out temperature for neutrinos can be estimated as

$$T_{f.o.} \approx \left( \frac{1}{G_F^2 M_{Pl}} \right)^{1/3} \approx 1 \text{ MeV} \quad (28)$$

After decoupling neutrinos cool down as  $T \propto 1/a$ . They would have the same temperature as photons except for the fact that photons get heated up by the annihilation of  $e^+ e^-$  at around  $T \approx 0.5 \text{ MeV}$ . Hence neutrinos are a bit cooler than photons today ( $T_\nu = 1.95 \text{ K}$ ).

This relic low energy neutrino background has not been directly detected yet. There is a planned experiment called PTOLOMY designed to detect it [11].

In the SM there are only left-handed neutrinos and they are massless. However, we now know that neutrinos oscillate among different flavors and must be massive. Therefore the SM should be augmented, most plausibly with the inclusion of a new degree of freedom, the right-handed neutrino, which



Fermilab Photograph 85-138CN

Fig. 3: Thermal history of the Universe. Figure from Kolb & Turner “The Early Universe”.

is a gauge singlet. Right-handed neutrinos may have cosmological consequences, acting as warm dark matter. Neutrinos could be dark matter but because they are light they would be relativistic at the time of decoupling - this is the definition of hot dark matter. Hot dark matter is already ruled out by cosmology, more specifically due to the fact that it would erase small scale structure in the Universe because of its free streaming.

We now understand that neutrinos contribute today a very small amount to the energy density budget of the Universe, just like photons, but bounds on their masses can still be derived from cosmological observations related to the formation of structures in the Universe and from the CMB (for a recent review, see [12]). Using a recent combination of CMB and LSS observables results in the 95% upper bound  $\sum_i m_{\nu_i} \leq 0.16$  eV [13].

The thermal history of the Universe is a very rich and dramatic one, with several events taking place during its evolution, such as inflation, phase transitions (electroweak symmetry breaking, QCD confinement, etc), the changes in the dominant component of the Universe, decoupling and annihilation of whole species and finally the take over of Dark Energy. There are many illustrative figures of this history but my favourite one is of course from the book by Kolb and Turner, which I reproduce in Figure 3.

## 2.4 The cosmological constant: the elephant in the room

The cosmological constant, first introduced by Einstein in 1917 in order to allow for a static Universe that was then consistent with observations, has been rejected by him as his biggest blunder after the discovery of the expansion of the Universe in 1929. However, as George Gamow puts it in his 1970

autobiography [14], “But this “blunder”, rejected by Einstein, is still sometimes used by cosmologists even today and the cosmological constant rears its ugly head again, again and again.”

The cosmological constant was always something that we, theoretical physicists, wanted to put aside. But the discovery of the accelerated expansion of the Universe blew the cosmological constant into our faces. It is too much of an embarrassment. The main reason for this embarrassment is the fact that its value is uncontrollable: it suffers from quadratic divergences when quantum corrections are taken into account, much worse than the quadratic divergences that afflicts the mass of the Higgs boson in the SM. The easiest way to see the origin of this divergence is to consider the contribution of the zero-point (vacuum) energy to a scalar quantum field of mass  $m$  (the  $\hbar\omega/2$  factor, where  $\omega^2 = k^2 + m^2$ ), which can be thought of as an infinite number of harmonic oscillators:

$$\rho_{vac} = \int \frac{d^3k}{(2\pi)^3} \frac{1}{2} \sqrt{k^2 + m^2}. \quad (29)$$

This integral diverges as an energy cut-off scale  $E_{cut}^4$ . If one uses the Planck scale as the cut-off the well known discrepancy of  $10^{120}$  with respect to observations is obtained. This may very well be largest discrepancy in the history of Physics and it is referred to as the cosmological constant problem [15]. However, it only reflects the fact that we do not understand what is going on. To this date there is not a good solution to this problem. Therefore, most of us chose to continue ignoring it.

### 3 Second lecture: Origins

In this second lecture I briefly discuss how some things came to existence in the Universe:

- Origin of light elements;
- Origin of baryons;
- Origin of dark matter;
- Origin of inhomogeneities.

#### 3.1 Origin of light elements: BBN

Big Bang Nucleosynthesis (BBN) is one of the pillars of the Standard Cosmological Model (for a recent review see [16])<sup>1</sup>. It is the earliest cosmological probe that we have so far (couple of minutes after the Bang). The idea goes back to George Gamow and his students in the 1940’s.

The details involve a complicated set of nuclear reactions which can be studied with sophisticated codes but here I will present a very simplified picture of BBN developed in five easy steps:

(1) When  $T \gg 1$  MeV ( $t \ll 1$  s), the Universe is made out of neutrons, protons, electrons and photons. Neutrons and protons are in thermal equilibrium due to the weak force. When  $T = \mathcal{O}(1$  MeV), protons and neutrons are non-relativistic with the ratio of their number density (denoted by  $n_n/n_p$ ) given by

$$\frac{n_n}{n_p} = e^{-Q/T}, \quad (30)$$

where  $Q = (m_n - m_p)$  is the mass difference between them. Notice that the neutron-proton mass difference ( $Q = 1.3$  MeV) is very small compared to their masses ( $m_n \approx m_p \approx 1000$  MeV).

(2) Neutrons and protons freeze out at  $T = 0.8$  MeV. Their number remains constant afterwards (number densities are  $\propto a^{-3}$ ), except for neutron decay which we will take into account in step (3). The neutron fraction ( $X_n$ ) then becomes

$$X_n = \frac{n_n}{n_n + n_p} = \frac{e^{-Q/T}}{1 + e^{-Q/T}} \Bigg|_{T=0.8 \text{ MeV}} \approx \frac{1}{6}. \quad (31)$$

---

<sup>1</sup>I learned BBN many years ago from an excellent review by Gary Steigman, a pioneer in the development of the connection between cosmology and particle physics and an expert on BBN [17]. Gary used to come to Brazil frequently and we became good friends. I was very saddened to learn about his untimely passing last year a bit after the School took place.

(3) The neutron fraction is almost frozen except for the fact that free neutrons decay with a lifetime  $\tau_n \approx 900$ s. Hence, after freeze-out

$$X_n(t) = e^{-t/\tau_n} X_n. \quad (32)$$

(4) Helium can only be formed by nuclear reactions when deuterium is present. Hence the temperature of the Universe has to be smaller than the deuterium binding energy -  $E_D \approx 0.06$  MeV ( $t_D \approx 330$ s). At this time

$$X_n(330 \text{ s}) \approx 1/8. \quad (33)$$

(5) At this point we can approximate that all neutrons present are used to form helium-4 ( $n_{He} = 2n_n$ ). Then we can compute the fraction in mass of the Universe in  ${}^4\text{He}$  as:

$$Y_{He} = \frac{4n_{He}}{n_p} \approx 2X_n \approx 1/4. \quad (34)$$

Therefore after the first few minutes of the Universe approximately 25% of the mass of atoms are in the form of  ${}^4\text{He}$ . There are also small quantities of D,  ${}^3\text{He}$ , and  ${}^6\text{Li}$  that can be computed by dedicated computer codes that take into account the dynamics of the many nuclear reactions involved in the background of an expanding Universe. The results from these computations depend crucially on the amount of photons in the Universe or, more precisely, on the ratio between the number of protons (or baryons, more generally) and photons. This ratio, denoted by  $\eta$ , determined from BBN is in good agreement with an independent measurement from the CMB:

$$\eta = \frac{n_b}{n_\gamma} = (6.10 \pm 0.04) \times 10^{-10}. \quad (35)$$

The determination of  $\eta$  coupled to the measurement of CMB provides an estimate of the baryonic content of the Universe,  $\Omega_b$ .

The agreement between the measurements of the abundances of the light elements (with some extrapolation to their primordial abundance) and the predictions from BBN is one of the great successes of the Standard Cosmological Model.

### 3.2 Origin of baryons: baryogenesis

One can try to estimate what would be the relic amount of baryons using arguments similar to the ones discussed above: baryons can be kept in thermal equilibrium by the strong interactions until their density drop sufficiently when they become nonrelativistic so that they leave thermal equilibrium and the corresponding baryon number gets frozen. A simple calculation that I sketched in the lectures lead to the disturbing result of  $\eta \approx 10^{-19}$ . It is not a typo: there is a disagreement of 9 orders of magnitude with the observed value!

What is the catch? We started with the same number of baryons and anti-baryons, a very reasonable assumption given that they are in thermal equilibrium. But that can't work. It turns out that we need to generate a tiny asymmetry between baryons and anti-baryons. In fact, a difference of one extra baryon in 1 billion baryons would be sufficient.

How can such an asymmetry be generated in particle physics? Andrei Sakharov laid the conditions for this to happen in 1967 [18]. These are:

- presence of baryon number violating processes;
- presence of C and CP violation in these processes;
- these processes to be out-of-equilibrium.

It is fair to say that there is no standard model of baryogenesis. All these conditions are met in the SM (even baryon number violation which occurs non-perturbatively) but the amount of CP violation turns out to be too small. This is one of the motivations for searching new sources of CP violation

(eg, neutrino sector). Models with Grand Unified Theories, baryogenesis through leptogenesis and other BSM possibilities are also being currently considered. This is certainly one of the hottest topics of research today and the jury is still out.

### 3.3 Origin of dark matter

#### 3.3.1 Evidences

Evidence for dark matter (DM) arises from different observations at different scales. So far they are unfortunately all based on observations in the heavens and not in laboratories. Among them I can list:

- Dynamics of clusters of galaxies
- Rotational curves of galaxies
- Gravitational lensing
- Cosmic microwave background
- Big bang nucleosynthesis
- Structure formation in the universe
- Baryon acoustic oscillations
- Bullet cluster

I'll not have time to describe them in any detail here. For a review see the lectures [6, 7].

DM is most possible a neutral, long-lived particle. I don't know any viable alternative. In the SM the neutrino could have the role of DM but they are ruled out since their contribution to the energy budget of the Universe is very small. Hence, DM implies *new physics beyond the SM*.<sup>2</sup> Structure formation tells us that dark matter must be cold (ie, nonrelativistic at decoupling) or a most warm (there are bounds from the so-called Lyman- $\alpha$  absorption lines from distant quasars setting  $m \geq 5.3$  keV at  $2\sigma$  for the mass of the warm dark matter particle [21]).

There are several candidates for dark matter: weakly interacting massive particles (WIMPs), new scalars (phion, inert Higgs models),  $\nu_R$ , axions, primordial black holes, lightest KK particle, self-interacting dark matter, etc. Usually one needs a symmetry (most times discrete, such as  $Z_2$ ) to ensure the stability of the lightest particle that is odd under it. WIMPs in particular are well-motivated candidates since they are predicted in SUSY extensions of the SM (with  $R$  parity conservation): the lightest supersymmetric particle (LSP), usually a neutralino (a given combination of gauginos and higgsinos). Candidates must pass several observational constraints. In [22] a ten-point test is proposed and 16 candidates are scrutinized.

#### 3.3.2 Thermal production of dark matter: the "miracle"

Dark matter can have other interactions in addition to the gravitational one. Hence it could have been in thermal equilibrium in the early Universe - in this case it is called thermal DM.<sup>3</sup> In the lectures I presented a very simple way to estimate the relic abundance of thermal DM particles. For accurate estimates one must solve the appropriate Boltzmann equation with the correct thermally averaged cross section. There are specialized and sophisticated codes such as MicroOMEGAs<sup>4</sup> and others that compute DM relic densities.

We already saw in the first lecture that particles get out of thermal equilibrium and their abundance freeze-out roughly when the interaction rate becomes of the order of the expansion rate. Using Eqs. (26)

<sup>2</sup>The lack of signals for new physics and the detection of gravitational waves have led people to think harder in DM within the SM, either in the form of primordial black holes or new quark states, see eg [19, 20].

<sup>3</sup>There is also the possibility that DM is non-thermally produced, eg from the non-equilibrium decay of other particles or by coherent field oscillations (as in the case of axions).

<sup>4</sup><https://lapth.cnrs.fr/micromegas/>



and (27) one obtains for its number density at freeze-out:

$$n_{f.o.}^{\chi} \approx \frac{T_{f.o.}}{\langle \sigma v \rangle M_{Pl}}. \quad (36)$$

On the other hand, if the DM is cold it is nonrelativistic at freeze-out and its number density is given by a Boltzmann distribution:

$$n_{f.o.}^{\chi} = (m_{\chi} T_{f.o.})^{3/2} e^{-m_{\chi}/T_{f.o.}}, \quad (37)$$

where I'm using  $m_{\chi}$  for the DM mass. Introducing  $x = m_{\chi}/T$  it can be shown that equating Eqs (36) and (37) results in  $x_{f.o.} \approx 30$  for typical values of  $m_{\chi} = 100$  GeV,  $v = 0.3$  and  $\sigma = G_F^2 m_{\chi}^2$ . One can now compute the relic DM abundance

$$\Omega_{\chi} = \frac{m_{\chi} n^{\chi}(T=0)}{\rho_c^{(0)}} = \mathcal{O}(1) \quad (38)$$

This is the so-called WIMP miracle: a thermal relic with weak cross section results in a relic abundance with order of magnitude of the observed one. This mechanism of thermal relic production leads to the *survival of the weakest*: the weaker the cross section the earlier the particle freezes out and consequently the larger its abundance.

There are several experiments looking for DM. There are basically three types of searches: direct production at the LHC (with signatures such as monojets), direct detection of DM particles that surround us in our galaxy in underground laboratories (one can estimate that of the order of a billion particles of DM passes through a typical person per second) and indirect detection through the annihilation of DM into SM particles occurring in dense DM rich regions of the Universe (center of our galaxy, satellite dwarf galaxies, etc). It is beyond the scope of these lectures to discuss these searches but many details can be found in the suggested reading. I just can't resist to mention the recent lower bound of 70 GeV on the mass of thermal DM coming from indirect searches from annihilation (assumed to be exclusively into  $b\bar{b}$ ) in the Milky Way halo [23].

### 3.4 Origin of inhomogeneities

This is another sketchy subsection. Again I refer to the excellent lectures mentioned in the introduction.

#### 3.4.1 The causality problem

CMB is originated at the "last scattering surface" when atoms are formed and the Universe becomes transparent to radiation. That's when radiation decouples from matter. This happened at  $z \approx 1100$  ( $t \approx 380,000$  years after the bang).

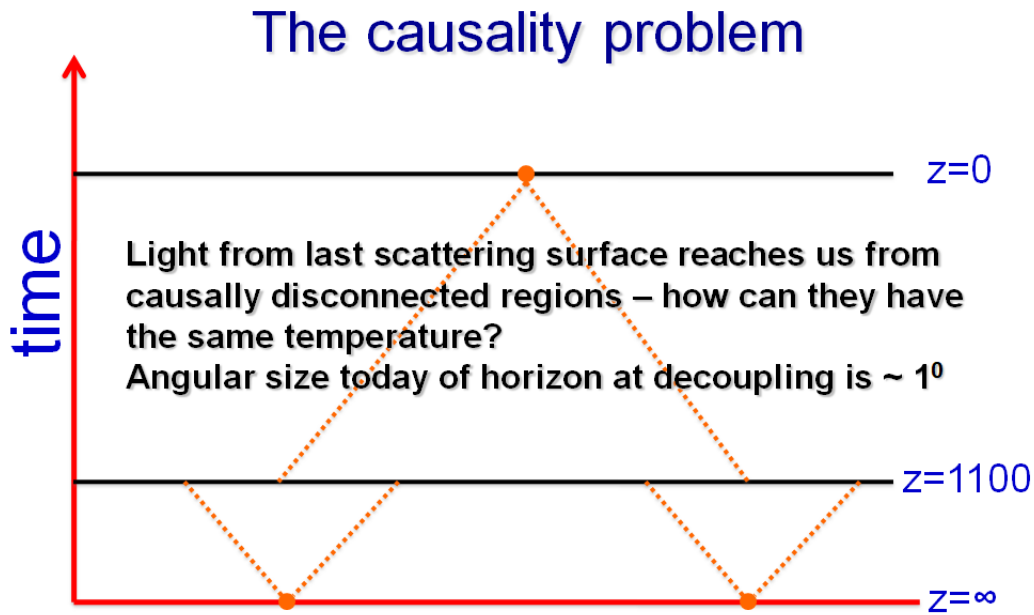
There are regions in the last scattering surface that were never in causal contact - and hence no reason to have the same temperature. Nevertheless the CMB is very uniform over the whole sky, with small variations of 1 part in  $10^5$ . This problem is sketched in Figure 4.

#### 3.4.2 Inflation

Inflation is a period of very fast (exponential) expansion of the Universe. It is similar to the period of dark energy domination we are in now. A single small patch can fill the whole horizon at decoupling. Hence inflation also predicts that the Universe is spatially flat - as observed. Inflation also provides quantum fluctuations that are the seeds for inhomogeneities in the Universe.

The basic idea is simple: the very early Universe is dominated by the energy density of a (surprise!) scalar field - called inflaton - that is slowly rolling down in a potential. There are a plethora of models used to implement inflation and most of them are listed in the "Encyclopedia Inflationaris" [24].





**Fig. 4:** The causality problem.

Inflation must end otherwise we wouldn't be here. And at the end of the inflationary period the Universe is empty and cold. Something has to jump-start the Universe again to the usual hot and dense phase that we know it should have existed in the past. Usually the end of inflation is identified with the end of the so-called slow-roll period when the scalar field is rolling down in a part of the potential that is relatively flat. After the slow-roll phase, the potential becomes steeper and usually has a minimum. The field then starts to oscillate around the minimum of the potential. It is not difficult to show that an oscillating field is equivalent to a nonrelativistic gas of inflaton particles (the mass of the inflaton is set at  $m \approx 10^{12}$  GeV from the amplitude of perturbations in the CMB). The inflaton then decays into radiation (this is model dependent) and results in the so-called reheating of the Universe. A sketch of the potential is shown in Figure 5.

The actual reheating process is more complicated and can be simulated numerically. There is a possibility of preheating, when instabilities in the scalar field perturbations can occur before the coherent oscillation period. The only bound on the reheating temperature  $T_R$  is that it must be larger than temperatures required by BBN (1 GeV).

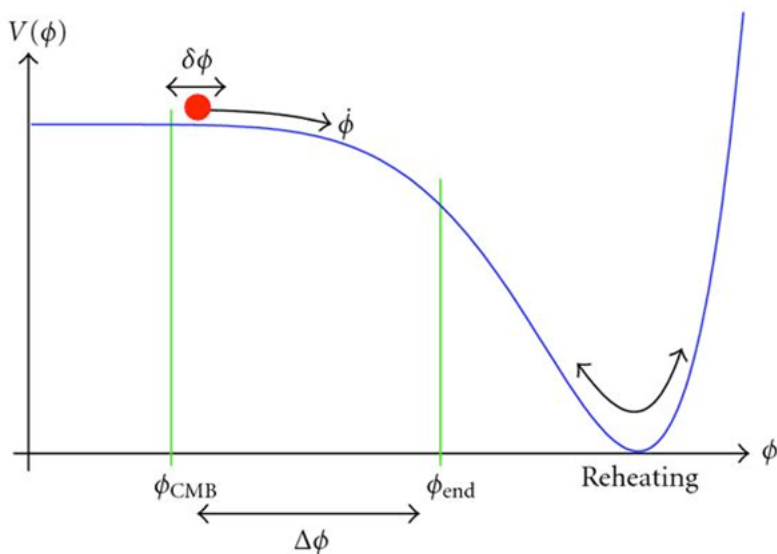
In order to solve the causality problem, one should require that the observable Universe today was to fit in the comoving Hubble radius at the beginning of inflation. A simple calculation that I did in the lectures shows that the amount of inflation needed, characterized by the number of e-foldings denoted by  $N$  is:

$$N = \ln \left( \frac{a_e}{a_i} \right) = 26 - 64, \quad (39)$$

where  $a_i$  and  $a_e$  are the scale factor and the beginning and end of the inflation and the range in  $N$  arises from the uncertainty in the reheating temperature (from 1 to  $10^{15}$  GeV). One usually assumes  $N \approx 50$ .

### 3.4.3 Perturbations generated during inflation

The origin of the inhomogeneities we observe today are quantum fluctuations in the inflaton field  $\delta\phi$  generated during inflation. The size of quantum fluctuations of the inflaton field during inflation is set by



**Fig. 5:** A cartoon of the inflaton potential.

H:

$$\langle (\delta\phi)^2 \rangle = \frac{H^2}{2\pi} \quad (40)$$

In general, perturbations  $\delta(\vec{x}, t)$  can be decomposed in Fourier modes:

$$\delta(\vec{x}, t) = \int d^3k \delta_k(t) e^{i\vec{k}\cdot\vec{x}} \quad (41)$$

and the *power spectrum*  $P(k)$  is defined by

$$\langle \delta_{\vec{k}'} \delta_{\vec{k}} \rangle = (2\pi)^3 \delta^3(\vec{k}' - \vec{k}) P(k). \quad (42)$$

Inflation *predicts* a power-law primordial power spectrum for scalar and tensor (gravitational waves) perturbations

$$\begin{aligned} P_s(k) &= A_s k^{n_s-1} \\ P_t(k) &= A_t k^{n_t} \end{aligned} \quad (43)$$

where  $A_{s,t}$  and  $n_{s,t}$  are respectively the amplitudes and spectral indices of the perturbations. They can be computed in a given model of inflation.

In particular, CMB bounds on the ratio of the tensor-to-scalar amplitudes, denoted by  $r$ , have been used to eliminate several models of inflation. There were some claims that a non-zero value of  $r$  had been detected, which would imply in the detection (albeit indirect) of primordial gravitational waves generated during inflation. However, the claim turned out to be just background and the latest value is  $r < 0.12$  at 95% confidence level [25]. But this upper bound is enough to rule out well-know models for the inflaton field, such as models with a  $\phi^4$  potential.

Even if there is no fully satisfactory model, inflation is great because it

- explain why the Universe is spatially flat;
- solve the causality problem;
- generate almost gaussian, almost scale invariant fluctuations;

- generate both scalar and tensor fluctuations;
- given a inflation potential one can predict the spectrum of scalar and tensor perturbations;
- the scalar (density) perturbations gives rise to the large scale structure of the Universe.

#### 4 Third lecture: The perturbed Universe

In this lecture I briefly describe how the tiny perturbations generated during inflation grow to give rise to the structures in the Universe that we measure today with large surveys of galaxies such as the Dark Energy Survey.

##### 4.1 Growth of perturbations

Tiny fluctuations of the order  $\delta \approx 10^{-5}$  were detected in the CMB around 1991. These fluctuations grew due to gravity in the evolution of the Universe. Their growth is determined by solving the perturbed Einstein's equations. While this can be done analytically in a linearized way, ie keeping only linear terms in the perturbations, when the perturbations become large enough ( $\delta \approx 1$ ) this approximation ceases to be valid. In this case one has to resort to full numerical simulations, such as N-body simulations. These simulations are becoming more and more realistic, with the inclusion of baryons.<sup>5</sup>

In principle one has to use full GR to study the evolution of perturbations. However, at scales smaller than the Hubble radius and for non-relativistic matter one can simplify the problem and use Newtonian physics. Hence one has to study fluid dynamics in an expanding Universe. In this case we have to consider three coupled equations: the continuity equation, the Euler equation and the Poisson equation. As I described in the lecture (you can see the slides), in the linearized approximation we can derive a single equation for the time evolution of the matter density perturbation  $\delta_m$ :

$$\ddot{\delta}_m + 2H(t)\dot{\delta}_m - \frac{3}{3}H(t)^2\Omega_m(t)\delta_m = 0 \quad (44)$$

where

$$\delta_m = \frac{\rho_m - \bar{\rho}_m}{\bar{\rho}_m} \quad (45)$$

Using this simple equation it is easy to show that in a matter-dominated Universe ( $\Omega_m \approx 1$ ,  $a \propto t^{2/3}$ ,  $H \propto 2/(3t)$ ) the matter density perturbation grow as the scale factor  $\delta \propto a$ . On the other hand, in a dark energy dominated Universe ( $\Omega_m \approx 0$ ,  $a \propto e^{Ht}$ ,  $H = \text{const.}$ ) a solution to the equation is  $\delta = \text{const.}$ , ie dark energy prevents matter perturbations from growing. This led to the famous anthropic argument for an upper limit in the cosmological constant put forward by Weinberg: the energy density of the cosmological constant can't be too large otherwise galaxies would not form and we would not exist.

##### 4.1.1 The matter power spectrum in recent times

We already introduced the power spectrum in Eq.(42). Here we discuss some observables related to it after the evolution of perturbations. It can be easily shown that the power spectrum is the Fourier transform of the 2-point correlation function of the density perturbations:

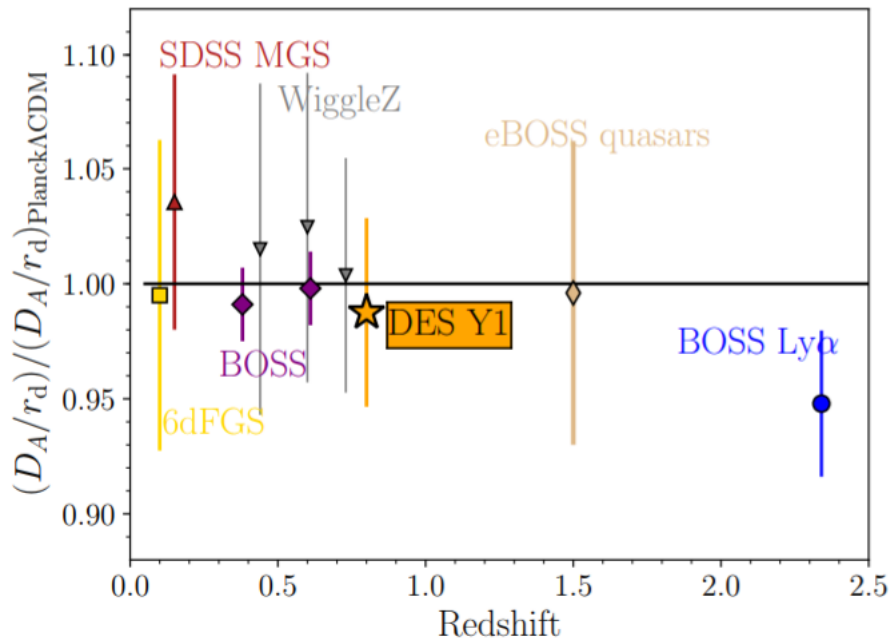
$$P(k) = \int d^3k \xi(r) e^{i\vec{k}\cdot\vec{r}} \quad (46)$$

where the two-point correlation function

$$\xi(r) = \langle \delta(\vec{x}_1) \delta(\vec{x}_2) \rangle \quad (47)$$

---

<sup>5</sup>See eg the Eagle Simulation, <http://icc.dur.ac.uk/Eagle/>



**Fig. 6:** Ratio of angular distance measurements in BAO to values from Planck fiducial cosmology [27].

depends only on  $r = |\vec{x}_1 - \vec{x}_2|$  due to homogeneity and isotropy. Therefore features such as peaks in the 2-point correlation function are transformed into oscillatory features in the power spectrum (think that the Fourier transform of a Dirac  $\delta^3(r - r_*)$  results in  $P(k) = e^{ikr_*}$ ).

Are there peaks in the correlation function of galaxies? Is there a preferred scale in the sky? The answer is affirmative and comes from physics related to the decoupling of matter and radiation at recombination ( $z_{rec} \approx 1100$ ). It is the sound horizon at decoupling.

Before recombination, baryons and photons were strongly coupled, forming a single fluid with pressure and speed. Dark matter, neutrinos and other forms were decoupled. After decoupling, baryons are left behind at a characteristic distance given by the sound horizon at decoupling, which is called the baryon acoustic oscillation (BAO) scale  $r_{BAO}$ :

$$r_{BAO} = \int_{z_{rec}}^{\infty} dz \frac{c_s(z)}{H(z)} \approx 150 \text{ Mpc}, \quad (48)$$

where  $c_s^2 \approx 1/3$  is the speed of propagation of the photon-baryon fluid. This characteristic scale in the distribution of baryons in real space gives rise to oscillations in the matter power spectrum.

The BAO feature in the matter power spectrum and in the real space 2-point correlation function was first detected in 2005. The position of the BAO peak provides a standard ruler in the sky that can be used to determine cosmological parameters. The ratio of an angular distance measure in BAO data to its theoretical value in a  $\Lambda$ CDM cosmology with Planck cosmology (ie, cosmological parameters as derived by the Planck collaboration) is shown in Figure 6 from [27], where one can see the good agreement.

#### 4.2 If the Universe is the answer what is the question?

The question we want to answer is: given our Universe what is the best model that describes it? And at this moment the flat  $\Lambda$ CDM Standard Cosmological model describes all the observations so far. This model is characterized by six parameters: the Hubble constant ( $H_0$ ), the baryon abundance ( $\Omega_b$ ), the abundance of cold dark matter ( $\Omega_{CDM}$ ), the amplitude of the initial scalar perturbations ( $A_s$ ), the spectral

index of the scalar perturbations ( $n_s$ ) and the so-called optical depth ( $\tau$ , related to the ionization history of the Universe). In addition to these baseline parameters, one could add neutrino masses, nonzero spatial curvature and a constant equation of state  $\omega \neq -1$ .

Cosmology is sensitive to the sum of the neutrino masses ( $\Sigma = \sum_i m_i$ ), as we already mentioned in Section 2.3 - it is interesting to notice that the value of  $\Sigma$  depends on the hierarchy of neutrino masses. Using data from neutrino oscillations:

$$\Sigma \geq \begin{cases} (58.5 \pm 0.48) \text{ meV} & \text{for normal hierarchy} \\ (98.6 \pm 0.85) \text{ meV} & \text{for inverted hierarchy} \end{cases} \quad (49)$$

The equality is attained when the lightest mass is zero. Therefore if from cosmology one finds  $\Sigma < 0.098$  eV then one can say that the inverted hierarchy is excluded. There are claims in the literature of strong evidence (in the bayesian sense) for normal hierarchy [26].

Large scale surveys are important for the determination of cosmological parameters. I'm a member of the Dark Energy Survey (DES), which uses photometric techniques to estimate the redshift of galaxies. DES uses a 570 Mpixel digital camera installed in the Blanco 4-meter telescope at the Cerro Tololo International Observatory in Chile. It will finish its 5 years observational period in 2018. The data for the first year has already been mostly analysed and we are now in the process of analysing the data of the three first years. Around 300 million galaxies have already been detected and catalogued. The DES BAO result for the first year of data is shown in Fig. 6.

The distribution of galaxies in the universe provide information about growth of perturbations (and hence is sensitive to Dark Energy or Modified Gravity), information about dark matter (eg hot DM is already ruled out) and a standard ruler (BAO scale). In addition, DES can measure weak gravitational lensing through the small distortions in the shapes of galaxies, can measure the distribution of galaxy clusters and can also measure SNIa. All these probes will be combined to derive the best constraints on cosmological parameters.

I would like to finish this subsection by drawing the following analogy between high energy accelerators and large scale galaxy survey:

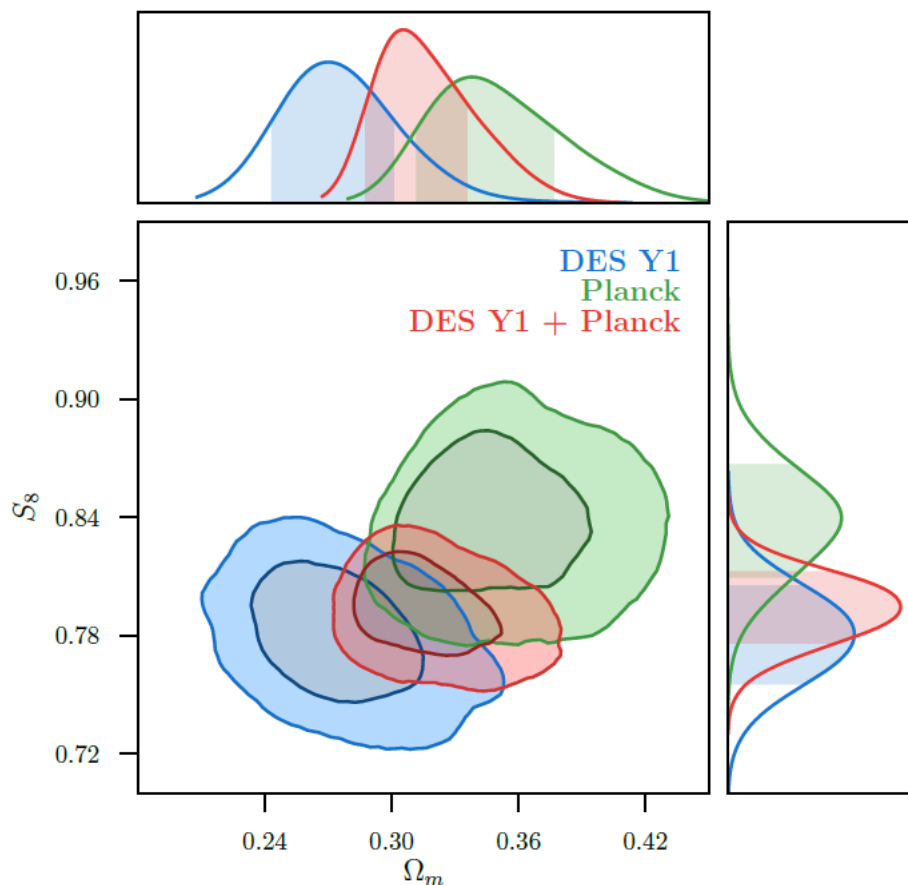
- Energy  $\leftrightarrow$  Redshift
- Luminosity  $\leftrightarrow$  Area & observation time
- Energy resolution  $\leftrightarrow$  Redshift errors
- Energy calibration  $\leftrightarrow$  Redshift calibration using objects with known redshifts
- $p_T$  cuts, etc  $\leftrightarrow$  Magnitude cuts, mask, etc
- Final data set  $\leftrightarrow$  Value added catalogs
- Higgs bump hunting  $\leftrightarrow$  BAO bump hunting
- Perturbation theory in QCD is ok at high energies  $\leftrightarrow$  Perturbation theory in GR is ok at high  $z$

### 4.3 Finding out the best model from data

The way we estimate the best model from data is through the likelihood function, which can generically be written as

$$\mathcal{L}(\vec{p}) \propto e^{\left[ (\vec{x}^{obs} - \vec{x}^{th}(\vec{p}))^t \text{Cov}(\vec{x}^{obs} - \vec{x}^{th}(\vec{p})) \right]} \quad (50)$$

where  $\vec{p}$  is a set of parameters (it turns out that several so-called nuisance parameters must be introduced to model the theoretical prediction),  $\vec{x}^{obs}$  is the data vector,  $\vec{x}^{th}$  is the theoretical modelling of the observable and Cov denotes de covariance matrix. The covariance matrix can be either theoretically modelled, measure from subsets of data (jackknife, bootstrap or subsampling methods) or measured from a set of



**Fig. 7:**  $\Lambda$ CDM constraints from the three combined probes in DES Y1 (blue), Planck with no lensing (green), and their combination (red) [28].

mock data. One can then use Bayes theorem to turn the likelihood function into a probability distribution for the parameters given the data using a prior probability for the parameters. The parameters can then be estimated using the Markov Chain Monte Carlo method to sample the space of parameters with the known probability distribution.

In Figure 7 I show the money plot for the DES analyses of the first year of data using a combination of galaxy distribution and weak lensing data [28]. It shows constraints on only 2 parameters,  $S_8$  (related to the amplitude of the scalar power spectrum  $A_s$ ) and  $\Omega_m$ . All other parameters (including 20 nuisance parameters related to redshift uncertainties, galaxy bias, etc) have been marginalized over. In the same plot it is shown the constraints from CMB obtained by the Planck collaboration. It is amazing that for the first time the two constraints are comparable (usually CMB results are much more constraining than LSS ones) and compatible, given that they come from signals produced with billions of years of difference. We have also analysed a model with constant equation of state denoted by  $\omega$ CDM. Combining DES data with several external datasets resulted in  $\omega = -1.00^{+0.04}_{-0.05}$ . Therefore this analysis confirms that dark energy is compatible with the cosmological constant.

## 5 Coda

I have given a brief tour on some selected topics in cosmology. The current situation in cosmology is somewhat akin to the one in particle physics. As in particle physics, we have a Standard Cosmological Model,  $\Lambda$ CDM, that explains all the cosmological observations so far. It has only 6 parameters in its

minimal version (much less than the 20 or so parameters in the SM). However, in contrast to the SM where all the building blocks have been found and studied, in cosmology we don't know much about the 95% of the Universe that is comprised of dark energy and dark matter. Surveys like DES and the future Large Synoptic Survey Telescope (LSST), of which I'm also a member, will hopefully shed some light on the dark side of the Universe.

## 6 Acknowledgements

It was great fun to give the lectures at CLASHEP in Mexico. I would like to thank Nick Ellis, Martijn Mulders, Kate Ross and Malena Tejada for the organization, the invitation to lecture and for the hospitality in San Juan del Rio.

## References

- [1] E. W. Kolb and M. S. Turner, *Front. Phys.* **69**, 1 (1990).
- [2] S. Dodelson, Amsterdam, Netherlands: Academic Pr. (2003) 440 p
- [3] <http://www.damtp.cam.ac.uk/user/db275/Cosmology/Lectures.pdf>
- [4] P. Binétruy, doi:10.5170/CERN-2014-008.217 arXiv:1504.07050 [gr-qc].
- [5] A. Riotto, CERN Yellow Report CERN-2010-001, pp. 315-362 [arXiv:1010.2642 [hep-ph]].
- [6] G. B. Gelmini, arXiv:1502.01320 [hep-ph].
- [7] M. Lisanti, arXiv:1603.03797 [hep-ph].
- [8] <http://physicschool.web.cern.ch/PhysicsSchool/CLASHEP/CLASHEP2017/programme.html>
- [9] P. Creminelli and F. Vernizzi, *Phys. Rev. Lett.* **119**, no. 25, 251302 (2017) doi:10.1103/PhysRevLett.119.251302 [arXiv:1710.05877 [astro-ph.CO]].
- [10] A. G. Riess *et al.*, *Astrophys. J.* **826**, no. 1, 56 (2016) doi:10.3847/0004-637X/826/1/56 [arXiv:1604.01424 [astro-ph.CO]].
- [11] S. Betts *et al.*, arXiv:1307.4738 [astro-ph.IM].
- [12] M. Lattanzi and M. Gerbino, *Front. in Phys.* **5**, 70 (2018) doi:10.3389/fphy.2017.00070 [arXiv:1712.07109 [astro-ph.CO]].
- [13] S. Alam *et al.* [BOSS Collaboration], *Mon. Not. Roy. Astron. Soc.* **470**, no. 3, 2617 (2017) doi:10.1093/mnras/stx721 [arXiv:1607.03155 [astro-ph.CO]].
- [14] George Gamow, *My World Line : An Informal Autobiography* (1970).
- [15] S. Weinberg, *Rev. Mod. Phys.* **61**, 1 (1989). doi:10.1103/RevModPhys.61.1
- [16] R. H. Cyburt, B. D. Fields, K. A. Olive and T. H. Yeh, *Rev. Mod. Phys.* **88**, 015004 (2016) doi:10.1103/RevModPhys.88.015004 [arXiv:1505.01076 [astro-ph.CO]].
- [17] G. Steigman, *Ann. Rev. Nucl. Part. Sci.* **29**, 313 (1979). doi:10.1146/annurev.ns.29.120179.001525
- [18] A. D. Sakharov, *Pisma Zh. Eksp. Teor. Fiz.* **5**, 32 (1967) [*JETP Lett.* **5**, 24 (1967)] [*Sov. Phys. Usp.* **34**, no. 5, 392 (1991)] [*Usp. Fiz. Nauk* **161**, no. 5, 61 (1991)]. doi:10.1070/PU1991v034n05ABEH002497
- [19] C. Gross, A. Polosa, A. Strumia, A. Urbano and W. Xue, arXiv:1803.10242 [hep-ph].
- [20] S. Clesse and J. García-Bellido, arXiv:1711.10458 [astro-ph.CO].
- [21] V. Irsic *et al.*, *Phys. Rev. D* **96**, no. 2, 023522 (2017) doi:10.1103/PhysRevD.96.023522 [arXiv:1702.01764 [astro-ph.CO]].
- [22] M. Taoso, G. Bertone and A. Masiero, *JCAP* **0803**, 022 (2008) doi:10.1088/1475-7516/2008/03/022 [arXiv:0711.4996 [astro-ph]].
- [23] L. J. Chang, M. Lisanti and S. Mishra-Sharma, arXiv:1804.04132 [astro-ph.CO].
- [24] J. Martin, C. Ringeval and V. Vennin, *Phys. Dark Univ.* **5-6**, 75 (2014) doi:10.1016/j.dark.2014.01.003 [arXiv:1303.3787 [astro-ph.CO]].

- [25] P. A. R. Ade *et al.* [BICEP2 and Planck Collaborations], *Phys. Rev. Lett.* **114**, 101301 (2015) doi:10.1103/PhysRevLett.114.101301 [arXiv:1502.00612 [astro-ph.CO]].
- [26] F. Simpson, R. Jimenez, C. Pena-Garay and L. Verde, *JCAP* **1706**, no. 06, 029 (2017) doi:10.1088/1475-7516/2017/06/029 [arXiv:1703.03425 [astro-ph.CO]].
- [27] T. M. C. Abbott *et al.* [DES Collaboration], [arXiv:1712.06209 [astro-ph.CO]].
- [28] T. M. C. Abbott *et al.* [DES Collaboration], arXiv:1708.01530 [astro-ph.CO].



# Probability and Statistics for Particle Physics

C. Mañá

CIEMAT, Madrid, Spain

## Abstract

The aim of these lectures is to give an overview of the basic concepts and tools on Probability and Statistics that are relevant for analysis in Particle Physics.

## Keywords

Probability; Statistics and Inference; Monte Carlo Methods.

"They say that understanding ought to work by the rules of right reason. These rules are, or ought to be, contained in Logic; but the actual science of logic is conversant at present only with things either certain, impossible, or entirely doubtful, none of which (fortunately) we have to reason on. Therefore the true logic of this world is the calculus of Probabilities, which takes account of the magnitude of the probability which is, or ought to be, in a reasonable man's mind"

*J.C. Maxwell*

## 1 Introduction

These notes are based on a course on Probability and Statistics given to graduate and PhD students at different places and, in particular, at the very nice 2017 CERN Latin-American School of High Energy Physics in San Juan del Rio, Mexico. They contain a humble overview of the basic concepts and ideas one should have in mind before getting involved in data analysis and, although they may have to be tailored for undergraduate students, I hope they will be a useful reading for all of them.

I feel, maybe wrongly, that there is a tendency in a subset of the Particle Physics community to consider statistics as a collection of prescriptions *written in some holy references* that are used blindly with the only arguments that either *"everybody does it that way"* or that *"it has always been done this way"*. In the talks I tried to demystify the *"how to"* recipes not because they are not useful but because, on the one hand, they are applicable under some conditions that tend to be forgotten and, on the other, because if the concepts are clear so will be the way to proceed ("at least formally") for the problems that come across in Particle Physics. At the end, the quote from Laplace given at the beginning of the first section is what it is all about.

There is a countable set of books on probability and statistics and a sizable subset of them are very good out of which I would recommend the following ones (a personal choice function). Section 2 (Lecture 1) deals with probability and this is just a measure, a finite non-negative measure, so it will be very useful to read some sections of *Measure Theory* (2006; Springer) by V.I. Bogachev; in particular the chapters 1 and 2. A large fraction of the material presented in this lecture can be found in more depth, together with other interesting subjects, in the book *Probability: A Graduate Course* (2013; Springer Texts in Statistics) by A. Gut. Section 3 (Lecture 2) is about statistical inference, Bayesian Inference in fact, and a must for this topic is the *Bayesian Theory* (1994; John Wiley & Sons) by J.M. Bernardo and A.F.M. Smith that contains also an enlightening discussion about the Bayesian and frequentist approaches in the appendix B. It is beyond question that in any worthwhile course on statistics the ubiquitous frequentist methodology has to be taught as well and there are excellent references on the subject. Students are encouraged to look for instance at *Statistical Methods in Experimental Physics* (2006; World Scientific) by F. James, *Statistics for Nuclear and Particle Physicists* (1989; Cambridge

University Press) by L.Lyons or *Statistical Data Analysis* (1997; Oxford Science Pub.) by G. Cowan. Last, Section 4 (Lecture 3) is devoted to Monte Carlo simulation, an essential tool in Statistics and Particle Physics.

“*Time is short, my strength is limited,...*”, Kafka dixit, so many interesting subjects that deserve a whole lecture by themselves are left aside. To mention some: an historical development of probability and statistics, Bayesian Networks, Generalized Distributions (a different approach to probability distributions), Decision Theory (Games Theory), ... and Markov Chains for which we shall state only the relevant properties without further explanation. Additional chapters and appendices are given in [4].

“The Theory of Probabilities is basically nothing else but common sense reduced to calculus”

*P.S. Laplace*

## 2 Probability

### 2.1 The Elements of Probability: $(\Omega, \mathcal{B}, \mu)$

The axiomatic definition of probability was introduced by A.N. Kolmogorov in 1933 and starts with the concepts of *sample space*  $(\Omega)$  and *space of events*  $(\mathcal{B}_\Omega)$  with structure of  $\sigma$ -algebra. When the pair  $(\Omega, \mathcal{B}_\Omega)$  is equipped with a *measure*  $\mu$  we have a *measure space*  $(E, \mathcal{B}, \mu)$  and, if the measure is a *probability measure*  $P$  we talk about a *probability space*  $(\Omega, \mathcal{B}_\Omega, P)$ . Let's start with a discussion of these elements.

#### 2.1.1 Events and Sample Space: $(\Omega)$

To learn about the state of nature, we do experiments and observations of the natural world and ask ourselves questions about the outcomes. In a general way, the *object* of questions we may ask about the result of an experiment such that the possible answers are *it occurs* or *it does not occur* are called **events**. There are different kinds of events and among them we have the **elementary events**; that is, those results of the random experiment that **can not** be decomposed in others of lesser entity. The **sample space**  $(\Omega)$  is the set of **all** the possible **elementary outcomes (events)** of a random experiment and they have to be:

- i) **exhaustive**: any possible outcome of the experiment has to be included in  $\Omega$ ;
- ii) **exclusive**: there is no overlap of elementary results.

To study random phenomena we start by specifying the *sample space* and, therefore, we have to have a clear idea of what are the possible results of the experiment. To center the ideas, consider the simple experiment of rolling a die with 6 faces numbered from 1 to 6. We consider as *elementary events*

$$e_i = \{ \text{get the number } i \text{ on the upper face} \}; \quad i = 1, \dots, 6$$

so  $\Omega = \{e_1, \dots, e_6\}$ . Note that any possible outcome of the roll is included in  $\Omega$  and we can not have two or more elementary results simultaneously. But there are other types of events besides the elementary ones. We may be interested for instance in the parity of the number so we would like to consider also the possible results <sup>1</sup>

$$A = \{ \text{get an even number} \} \quad \text{and} \quad A^c = \{ \text{get an odd number} \}$$

<sup>1</sup>Given two sets  $A, B \subset \Omega$ , we shall denote by  $A^c$  the *complement* of  $A$  (that is, the set of all elements of  $\Omega$  that are not in  $A$ ) and by  $A \setminus B \equiv A \cap B^c$  the *set difference* or *relative complement of  $B$  in  $A$*  (that is, the set of elements that are in  $A$  but not in  $B$ ). It is clear that  $A^c = \Omega \setminus A$ .

They are not *elementary* since the result  $A = \{e_2, e_4, e_6\}$  is equivalent to get  $e_2, e_4$  or  $e_6$  and  $A^c = \Omega \setminus A$  to get  $e_1, e_3$  or  $e_5$ . In general, an **event** is any subset<sup>2</sup> of the sample space and we shall distinguish between:

**elementary events:** any element of the *sample space*  $\Omega$ ;

**events:** any subset of the *sample space*;

and two extreme events:

**sure events:**  $S_S = \{ \text{get any result contained in } \Omega \} \equiv \Omega$

**impossible events:**  $S_I = \{ \text{get any result not contained in } \Omega \} \equiv \emptyset$

Any event that is neither *sure* nor *impossible* is called **random event**. Going back to the rolling of the die, sure events are

$$\begin{aligned} S_S &= \{ \text{get a number } n \mid 1 \leq n \leq 6 \} = \Omega && \text{or} \\ S_S &= \{ \text{get a number that is even or odd} \} = \Omega \end{aligned}$$

impossible events are

$$\begin{aligned} S_I &= \{ \text{get an odd number that is not prime} \} = \emptyset && \text{or} \\ S_I &= \{ \text{get the number 7} \} = \emptyset \end{aligned}$$

and random events are any of the  $e_i$  or, for instance,

$$S_r = \{ \text{get an even number} \} = \{e_2, e_4, e_6\}$$

Depending on the number of possible outcomes of the experiment, the the sample space can be:

**finite:** if the number of elementary events is finite;

**Example:** In the rolling of a die,  $\Omega = \{e_i; i = 1, \dots, 6\}$  so  $\dim(\Omega) = 6$ .

**countable:** when there is a one-to-one correspondence between the elements of  $\Omega$  and  $\mathcal{N}$ ;

**Example:** Consider the experiment of flipping a coin and stopping when we get  $H$ . Then  $\Omega = \{H, TH, TTH, TTT, \dots\}$ .

**non-denumerable:** if it is neither of the previous;

**Example:** For the decay time of an unstable particle  $\Omega = \{t \in R \mid t \geq 0\} = [0, \infty)$  and for the production polar angle of a particle  $\Omega = \{\theta \in R \mid 0 \leq \theta \leq \pi\} = [0, \pi]$ .

---

<sup>2</sup>This is not completely true if the sample space is non-denumerable since there are subsets that can not be considered as events. It is however true for the subsets of  $\mathcal{R}^n$  we shall be interested in. We shall talk about that in section 1.2.2.

It is important to note that the *events* are not necessarily numerical entities. We could have for instance the die with colored faces instead of numbers. We shall deal with that when discussing *random quantities*. Last, given a sample space  $\Omega$  we shall talk quite frequently about a *partition* (or a *complete system of events*); that is, a sequence  $\{S_i\}$  of events, finite or countable, such that

$$\Omega = \bigcup_i S_i \quad (\text{complete system}) \quad \text{and} \quad S_i \cap S_j = \emptyset; \quad i \neq j \quad (\text{disjoint events}) \quad (1)$$

### 2.1.2 $\sigma$ -algebras ( $\mathcal{B}_\Omega$ ) and Measurable Spaces ( $\Omega, \mathcal{B}_\Omega$ )

As we have mentioned, in most cases we are interested in events other than the elementary ones. It is therefore interesting to consider a class of events that contains all the possible results of the experiment we are interested in such that when we ask about the union, intersection and complements of events we obtain elements that belong the same class. A non-empty family  $\mathcal{B}_\Omega = \{S_i\}_{i=1}^n$  of subsets of the sample space  $\Omega$  that is *closed* (or *stable*) under the operations of *union* and *complement*; that is

$$S_i \cup S_j \in \mathcal{B}; \quad \forall S_i, S_j \in \mathcal{B} \quad \text{and} \quad S_i^c \in \mathcal{B}; \quad \forall S_i \in \mathcal{B} \quad (2)$$

is an **algebra** (*Boole algebra*) if  $\Omega$  is finite. It is easy to see that if it is closed under unions and complements it is also closed under intersections and the following properties hold for all  $S_i, S_j \in \mathcal{B}_\Omega$ :

$$\begin{array}{lll} \Omega \in \mathcal{B}_\Omega & \emptyset \in \mathcal{B}_\Omega & S_i \cap S_j \in \mathcal{B}_\Omega \\ S_i^c \cup S_j^c \in \mathcal{B}_\Omega & (S_i^c \cup S_j^c)^c \in \mathcal{B}_\Omega & S_i \setminus S_j \in \mathcal{B}_\Omega \\ \bigcup_{i=1}^m S_i \in \mathcal{B}_\Omega & \bigcap_{i=1}^m S_i \in \mathcal{B}_\Omega & \end{array}$$

Given a sample space  $\Omega$  we can construct different Boole algebras depending on the events of interest. The smaller one is  $\mathcal{B}_m = \{\emptyset, \Omega\}$ , the minimum algebra that contains the event  $A \subset \Omega$  has 4 elements:  $\mathcal{B} = \{\emptyset, \Omega, A, A^c\}$  and the largest one,  $\mathcal{B}_M = \{\emptyset, \Omega, \text{all possible subsets of } \Omega\}$  will have  $2^{\dim(\Omega)}$  elements. From  $\mathcal{B}_M$  we can engender any other algebra by a finite number of unions and intersections of its elements.

#### 2.1.2.1 $\sigma$ -algebras

If the sample space is countable, we have to generalize the Boole algebra such that the unions and intersections can be done a countable number of times getting always events that belong to the same class; that is:

$$\bigcup_{i=1}^{\infty} S_i \in \mathcal{B} \quad \text{and} \quad \bigcap_{i=1}^{\infty} S_i \in \mathcal{B} \quad (3)$$

with  $\{S_i\}_{i=1}^{\infty} \in \mathcal{B}$ . These algebras are called  **$\sigma$ -algebras**. Not all the Boole algebras satisfy these properties but the  $\sigma$ -algebras are always Boole algebras (closed under finite union). Consider for instance a finite set  $E$  and the class  $\mathcal{A}$  of subsets of  $E$  that are either finite or have finite complements. The finite union of subsets of  $\mathcal{A}$  belongs to  $\mathcal{A}$  because the finite union of finite sets is a finite set and the finite union of sets that have finite complements has finite complement. However, the countable union of finite sets is countable and its complement will be an infinite set so it does not belong to  $\mathcal{A}$ . Thus,  $\mathcal{A}$  is a Boole algebra but not a  $\sigma$ -algebra. Let now  $E$  be any infinite set and  $\mathcal{B}$  the class of subsets of  $E$  that are either countable or have countable complements. The finite or countable union of countable sets is countable and therefore belongs to  $\mathcal{B}$ . The finite or countable union of sets whose complement is countable has a countable complement and also belongs to  $\mathcal{B}$ . Thus,  $\mathcal{B}$  is a Boole algebra and  $\sigma$ -algebra.

2.1.2.2 Borel  $\sigma$ -algebras

Eventually, we are going to assign a probability to the events of interest that belong to the algebra and, anticipating concepts, probability is just a bounded measure so we need a class of *measurable sets* with structure of a  $\sigma$ -algebra. Now, it turns out that when the sample space  $\Omega$  is a non-denumerable topological space there exist non-measurable subsets that obviously can not be considered as events<sup>3</sup>. We are particularly interested in  $\mathcal{R}$  (or, in general, in  $\mathcal{R}^n$ ) so we have to construct a family  $\mathcal{B}_{\mathcal{R}}$  of measurable subsets of  $\mathcal{R}$  that is

- i) closed under countable number of intersections:  $\{B_i\}_{i=1}^{\infty} \in \mathcal{B}_{\mathcal{R}} \longrightarrow \bigcap_{i=1}^{\infty} B_i \in \mathcal{B}_{\mathcal{R}}$
- ii) closed under complements:  $B \in \mathcal{B}_{\mathcal{R}} \rightarrow B^c = \mathcal{R} \setminus B \in \mathcal{B}_{\mathcal{R}}$

Observe that, for instance, the family of all subsets of  $\mathcal{R}$  satisfies the conditions *i*) and *ii*) and the intersection of any collection of families that satisfy them is a family that also fulfills this conditions but not all are measurable. Measurably is the key condition. Let's start identifying what we shall considered the *basic set* in  $\mathcal{R}$  to engender an algebra. The sample space  $\mathcal{R}$  is a linear set of points and, among it subsets, we have the **intervals**. In particular, if  $a \leq b$  are any two points of  $\mathcal{R}$  we have:

- open intervals:  $(a, b) = \{x \in \mathcal{R} \mid a < x < b\}$
- closed intervals:  $[a, b] = \{x \in \mathcal{R} \mid a \leq x \leq b\}$
- half-open intervals on the right:  $[a, b) = \{x \in \mathcal{R} \mid a \leq x < b\}$
- half-open intervals on the left:  $(a, b] = \{x \in \mathcal{R} \mid a < x \leq b\}$

When  $a = b$  the closed interval reduces to a point  $\{x = a\}$  (*degenerated interval*) and the other three to the null set and, when  $a \rightarrow -\infty$  or  $b \rightarrow \infty$  we have the *infinite intervals*  $(-\infty, b)$ ,  $(-\infty, b]$ ,  $(a, \infty)$  and  $[a, \infty)$ . The whole space  $\mathcal{R}$  can be considered as the interval  $(-\infty, \infty)$  and any interval will be a subset of  $\mathcal{R}$ . Now, consider the class of all intervals of  $\mathcal{R}$  of any of the aforementioned types. It is clear that the intersection of a finite or countable number of intervals is an interval but the union is not necessarily an interval; for instance  $[a_1, b_1] \cup [a_2, b_2]$  with  $a_2 > b_1$  is not an interval. Thus, this class is not additive and therefore not a closed family. However, it is possible to construct an additive class including, along with the intervals, other measurable sets so that any set formed by countably many operations of unions, intersections and complements of intervals is included in the family. Suppose, for instance, that we take the half-open intervals on the right  $[a, b)$ ,  $b > a$  as the initial class of sets<sup>4</sup> to generate the algebra  $\mathcal{B}_{\mathcal{R}}$  so they are in the bag to start with. The open, close and degenerate intervals are

$$(a, b) = \bigcup_{n=1}^{\infty} [a - 1/n, b) ; \quad [a, b] = \bigcap_{n=1}^{\infty} [a, b + 1/n) \quad \text{and} \quad a = \{x \in \mathcal{R} \mid x = a\} = [a, a] \quad (4)$$

so they go also to the bag as well as the half-open intervals  $(a, b] = (a, b) \cup [b, b]$  and the countable union of unitary sets and their complements. Thus, countable sets like  $\mathcal{N}$ ,  $\mathcal{Z}$  or  $\mathcal{Q}$  are in the bag too. Those are the sets we shall deal with.

The smallest family  $\mathcal{B}_{\mathcal{R}}$  (or simply  $\mathcal{B}$ ) of measurable subsets of  $\mathcal{R}$  that contains all intervals and is closed under complements and countable number of intersections has the structure of a  $\sigma$ -algebra, is

---

<sup>3</sup>In particular, there are subsets of  $\mathcal{R}^{n \geq 1}$  for which the Lebesgue measure does not satisfy  $\sigma$ -additivity... and even finite-additivity for  $\mathcal{R}^{n \geq 3}$  (Hausdorff; 1914). Is not difficult to show the existence of Lebesgue non-measurable sets in  $\mathcal{R}$ . One simple example is the *Vitali set* constructed by G. Vitali in 1905 although there are other interesting examples (Hausdorff, Banach-Tarsky) and they all assume the Axiom of Choice. In fact, the work of R.M. Solovay around the 70's shows that one can not prove the existence of Lebesgue non-measurable sets without it. However, one can not specify the choice function so one can prove their existence but can not make an explicit construction in the sense Set Theorists would like. In Probability Theory, we are interested only in Lebesgue measurable sets so those which are not have nothing to do in this business and Borel's algebra contains only measurable sets.

<sup>4</sup>The same algebra is obtained if one starts with  $(a, b)$ ,  $[a, b)$  or  $[a, b]$ .

called **Borel's algebra** and its elements are generically called **Borel's sets** or *borelians* Last, recall that half-open sets are Lebesgue measurable ( $\lambda((a, b]) = b - a$ ) and so is any set built up from a countable number of unions, intersections and complements so all Borel sets are Lebesgue measurable and every Lebesgue measurable set differs from a Borel set by at most a set of measure zero. Whatever has been said about  $\mathcal{R}$  is applicable to the n-dimensional euclidean space  $\mathcal{R}^n$ .

The pair  $(\Omega, \mathcal{B}_\Omega)$  is called **measurable space** and in the next section it will be equipped with a *measure* and "upgraded" to a *measure space* and eventually to a *probability space*.

**2.1.3 Set Functions and Measure Space:  $(\Omega, \mathcal{B}_\Omega, \mu)$**

A function  $f : A \in \mathcal{B}_\Omega \rightarrow \mathcal{R}$  that assigns to each set  $A \in \mathcal{B}_\Omega$  one, and only one real number, finite or not, is called a **set function**. Given a sequence  $\{A_i\}_{i=1}^n$  of subset of  $\mathcal{B}_\Omega$  pair-wise disjoint, ( $A_i \cap A_j = \emptyset; i, j = 1, \dots, n; i \neq j$ ) we say that the *set function* is **additive** (*finitely additive*) if:

$$f \left( \bigcup_{i=1}^n A_i \right) = \sum_{i=1}^n f(A_i) \tag{5}$$

or  **$\sigma$ -additive** if, for a countable the sequence  $\{A_i\}_{i=1}^\infty$  of pair-wise disjoint sets of  $\mathcal{B}$ ,

$$f \left( \bigcup_{i=1}^\infty A_i \right) = \sum_{i=1}^\infty f(A_i) \tag{6}$$

It is clear that any  $\sigma$ -additive set function is additive but the converse is not true. A *countably additive set function* is a **measure on** the algebra  $\mathcal{B}_\Omega$ , a *signed measure* in fact. If the  $\sigma$ -additive set function is  $\mu : A \in \mathcal{B}_\Omega \rightarrow [0, \infty)$  (i.e.,  $\mu(A) \geq 0$ ) for all  $A \in \mathcal{B}_\Omega$  it is a **non-negative measure**. In what follows, whenever we talk about *measures*  $\mu, \nu, ..$  on a  $\sigma$ -algebra we shall assume that they are always non-negative measures without further specification. If  $\mu(A) = 0$  we say that  $A$  is a set of *zero measure*.

The "trio"  $(\Omega, \mathcal{B}_\Omega, \mu)$ , with  $\Omega$  a non-empty set,  $\mathcal{B}_\Omega$  a  $\sigma$ -algebra of the sets of  $\Omega$  and  $\mu$  a measure over  $\mathcal{B}_\Omega$  is called **measure space** and the elements of  $\mathcal{B}_\Omega$  *measurable sets*.

In the particular case of the n-dimensional euclidean space  $\Omega = \mathcal{R}^n$ , the  $\sigma$ -algebra is the Borel algebra and **all** the Borel sets are measurable. Thus, the intervals  $I$  of any kind are *measurable sets* and satisfy that

- i) If  $I \in \mathcal{R}$  is measurable  $\rightarrow I^c = \mathcal{R} - I$  is measurable;
- ii) If  $\{I_i\}_{i=1}^\infty \in \mathcal{R}$  are measurable  $\rightarrow \cup_{i=1}^\infty I_i$  is measurable;

*Countable sets* are Borel sets of zero measure for, if  $\mu$  is the Lebesgue measure, we have that  $\mu([a, b)) = b - a$  and therefore:

$$\mu(\{a\}) = \lim_{n \rightarrow \infty} \mu([a, a + 1/n)) = \lim_{n \rightarrow \infty} \frac{1}{n} = 0 \tag{7}$$

Thus, *any point* is a Borel set with *zero Lebesgue measure* and, being  $\mu$  a  $\sigma$ -additive function, any countable set has zero measure. The converse is not true since there are borelians with zero measure that are not countable (i.e. Cantor's ternary set).

In general, a measure  $\mu$  over  $\mathcal{B}$  satisfies that, for any  $A, B \in \mathcal{B}$  not necessarily disjoint:

- m.1)  $\mu(A \cup B) = \mu(A) + \mu(B \setminus A)$
- m.2)  $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$   $(\mu(A \cup B) \leq \mu(A) + \mu(B))$
- m.3) If  $A \subseteq B$ , then  $\mu(B \setminus A) = \mu(B) - \mu(A)$   $(\geq 0 \text{ since } \mu(B) \geq \mu(A))$
- m.4)  $\mu(\emptyset) = 0$

Observe that, for (m.1),  $A \cup B$  is the union of two disjoint sets,  $A$  and  $B \setminus A$ , and the measure is an additive set function. For (m.2), we have that  $A \cap B^c$  and  $B$  are disjoint and its union is  $A \cup B$  so  $\mu(A \cup B) = \mu(A \cap B^c) + \mu(B)$ . On the other hand  $A \cap B^c$  and  $A \cap B$  are disjoint at its union is  $A$  so  $\mu(A \cap B^c) + \mu(A \cap B) = \mu(A)$ . Property (m.3) follows from (m.1) considering that, if  $A \subseteq B$ , then  $A \cup B = B$ . Last, (m.4) is a consequence of (m.3) with  $B = A$ .

A measure  $\mu$  over a measurable space  $(\Omega, \mathcal{B}_\Omega)$  is **finite** if  $\mu(\Omega) < \infty$  and  $\sigma$ -**finite** if  $\Omega = \cup_{i=1}^\infty A_i$ , with  $A_i \in \mathcal{B}_\Omega$  and  $\mu(A_i) < \infty$ . Clearly, any finite measure is  $\sigma$ -finite but the converse is not necessarily true. For instance, the Lebesgue measure  $\lambda$  in  $(\mathcal{R}^n, \mathcal{B}_{\mathcal{R}^n})$  is not finite because  $\lambda(\mathcal{R}^n) = \infty$  but is  $\sigma$ -finite because

$$\mathcal{R}^n = \bigcup_{k \in \mathcal{N}} [-k, k]^n$$

and  $\lambda([-k, k]^n) = (2k)^n$  is finite. As we shall see in lecture 2, in some circumstances we shall be interested in the limiting behaviour of  $\sigma$ -finite measures over a sequence of compact sets. As a second example, consider the measurable space  $(\mathcal{R}, \mathcal{B})$  and  $\mu$  such that for  $A \subset \mathcal{B}$  is  $\mu(A) = \text{card}(A)$  if  $A$  is finite and  $\infty$  otherwise. Since  $\mathcal{R}$  is an uncountable union of finite sets,  $\mu$  is not  $\sigma$ -finite in  $\mathcal{R}$ . However, it is  $\sigma$ -finite in  $(\mathcal{N}, \mathcal{B}_{\mathcal{N}})$ .

### 2.1.3.1 Probability Measure

Let  $(\Omega, \mathcal{B}_\Omega)$  be a measurable space. A measure  $P$  over  $\mathcal{B}_\Omega$  (that is, with domain in  $\mathcal{B}_\Omega$ ), image in the closed interval  $[0, 1] \in \mathcal{R}$  and such that  $P(\Omega) = 1$  (finite) is called a **probability measure** and its properties are just those of finite (non-negative) measures. Expliciting the axioms, a *probability measure* is a *set function* with *domain* in  $\mathcal{B}_\Omega$  and *image* in the closed interval  $[0, 1] \in \mathcal{R}$  that satisfies three *axioms*:

- i) **additive:** is an additive set function;
- ii) **no negativity:** is a measure;
- iii) **certainty:**  $P(\Omega) = 1$ .

These properties coincide obviously with those of the *frequency and combinatorial probability* (see Note 1). All probability measures are finite ( $P(\Omega) = 1$ ) and any bounded measure can be converted in a *probability measure* by proper normalization. The *measurable space*  $(\Omega, \mathcal{B}_\Omega)$  provided with and probability measure  $P$  is called the **probability space**  $(\Omega, \mathcal{B}_\Omega, P)$ . It is straight forward to see that if  $A, B \in \mathcal{B}$ , then:

- p.1)  $P(A^c) = 1 - P(A)$
- p.2)  $P(\emptyset) = 0$
- p.3)  $P(A \cup B) = P(A) + P(B \setminus A) = P(A) + P(B) - P(A \cap B) \leq P(A) + P(B)$

The property p.3 can be extended by recurrence to an arbitrary number of events  $\{A_i\}_{i=1}^n \in \mathcal{B}$  for if  $S_k = \cup_{j=1}^k A_j$ , then  $S_k = A_k \cup S_{k-1}$  and  $P(S_n) = P(A_n) + P(S_{n-1}) - P(A_n \cap S_{n-1})$ .

Last, note that in the probability space  $(\mathcal{R}, \mathcal{B}, P)$  (or in  $(\mathcal{R}^n, \mathcal{B}_n, P)$ ), the set of points  $W = \{\forall x \in \mathcal{R} \mid P(x) > 0\}$  is countable. Consider the partition

$$W = \bigcup_{k=1}^\infty W_k \quad \text{where} \quad W_k = \{\forall x \in \mathcal{R} \mid 1/(k+1) < P(x) \leq 1/k\} \quad (8)$$

If  $x \in W$  then it belongs to one  $W_k$  and, conversely, if  $x$  belongs to one  $W_k$  then it belongs to  $W$ . Each set  $W_k$  has at most  $k$  points for otherwise the sum of probabilities of its elements is  $P(W_k) > 1$ . Thus,

the sets  $W_k$  are finite and since  $W$  is a countable union of finite sets is a countable set. In consequence, we can assign finite probabilities on at most a countable subset of  $\mathcal{R}$ .

### NOTE 1: What is *probability*?

It is very interesting to see how along the 500 years of history of probability many people (Galileo, Fermat, Pascal, Huygens, Bernoulli, Gauss, De Moivre, Poisson, ...) have approached different problems and developed concepts and theorems (Laws of Large Numbers, Central Limit, Expectation, Conditional Probability,...) and a proper definition of probability has been so elusive. Certainly there is a before and after Kolmogorov's "*General Theory of Measure and Probability Theory*" and "*Grundbegriffe der Wahrscheinlichkeitsrechnung*" so from the mathematical point of view the question is clear after 1930's. But, as Poincaré said in 1912: "*It is very difficult to give a satisfactory definition of Probability*". Intuitively, What is probability?

The first "*definition*" of probability was the *Combinatorial Probability* ( $\sim 1650$ ). This is an objective concept (i.e., independent of the individual) and is based on Bernoulli's *Principle of Symmetry or Insufficient Reason*: all the possible outcomes of the experiment equally likely. For its evaluation we have to know the cardinal ( $\nu(\cdot)$ ) of all possible results of the experiment ( $\nu(\Omega)$ ) and the probability for an event  $A \subset \Omega$  is "*defined*" by the Laplace's rule:  $P(A) = \nu(A)/\nu(\Omega)$  This concept of probability, implicitly admitted by Pascal and Fermat and explicitly stated by Laplace, is an *a priori probability* in the sense that can be evaluated before or even without doing the experiment. It is however meaningless if  $\Omega$  is a countable set ( $\nu(\Omega) = \infty$ ) and one has to justify the validity of the Principle of Symmetry that not always holds originating some interesting debates. For instance, in a problem attributed to D'Alembert, a player  $A$  tosses a coin twice and wins if  $H$  appears in at least one toss. According to Fermat, one can get  $\{(TT), (TH), (HT), (HH)\}$  and  $A$  will lose only in the first case so being the four cases equally likely, the probability for  $A$  to win is  $P = 3/4$ . Pascal gave the same result. However, for Roberval one should consider only  $\{(TT), (TH), (H\cdot)\}$  because if  $A$  has won already if  $H$  appears at the first toss so  $P = 2/3$ . Obviously, Fermat and Pascal were right because, in this last case, the three possibilities are not all equally likely and the *Principle of Symmetry* does not apply.

The second interpretation of probability is the *Frequentist Probability*, and is based on the idea of *frequency of occurrence* of an event. If we repeat the experiment  $n$  times and a particular event  $A_i$  appears  $n_i$  times, the relative frequency of occurrence is  $f(A_i) = n_i/n$ . As  $n$  grows, it is observed (experimental fact) that this number stabilizes around a certain value and in consequence the probability of occurrence of  $A_i$  is defined as  $P(A_i) \equiv \lim_{n \rightarrow \infty}^{exp} f(A_i)$ . This is an objective concept inasmuch it is independent of the observer and is *a posteriori* since it is based on what has been observed after the experiment has been done through an *experimental limit* that obviously is not attainable. In this sense, it is more a practical rule than a definition. It was also implicitly assumed by Pascal and Fermat (letters of de Mere to Pascal: *I have observed in my die games...*), by Bernoulli in his *Ars Conjectandi* of 1705 (*Law of Large Numbers*) and finally was clearly explicated at the beginning of the XX'th century (Fisher and Von Mises).

Both interpretations of probability are restricted to observable quantities. What happens for instance if they are not directly observable?, What if we can not repeat the experiment a large number of times and/or under the same conditions? Suppose that you jump from the third floor down to ground (imaginary experiment). Certainly, we can talk about the probability that you break your leg but, how many times can we repeat the experiment under the same conditions?

During the XX'th century several people tried to pin down the concept of probability. Pierce and, mainly, Popper argued that probability represents the *propensity* of Nature to give a particular result in a **single** trial without any need to appeal at "*large numbers*". This assumes that the *propensity*, and therefore the probability, exists in an *objective* way even though the *causes* may be difficult to



understand. Others, like Knight, proposed that randomness is not a measurable property but just a problem of knowledge. If we toss a coin and know precisely its shape, mass, acting forces, environmental conditions,... we should be able to determine with certainty if the result will be head or tail but since we lack the necessary information we can not predict the outcome with certainty so we are lead to consider that as a random process and use the Theory of Probability. Physics suggests that it is not only a question of knowledge but randomness is deeply in the way Nature behaves.

The idea that probability *is a quantification of the degree of belief that we have in the occurrence of an event* was used, in a more inductive manner, by Bayes and, as we shall see, Bayes's theorem and the idea of information play an essential role in its axiomatization. To quote again Poincare, " ... *the probability of the causes, the most important from the point of view of scientific applications.*". It was still an open the question whether this quantification is *subjective* or not. In the 20's, Keynes argued that it is not because, if we know all the elements and factors of the experiment, what is likely to occur or not is determined in an objective sense regardless what is our opinion. On the contrary, Ramsey and de Finetti argued that the probability that is to be assigned to a particular event depends on the degree of knowledge we have (*personal beliefs*) and those do not have to be shared by everybody so it is *subjective*. Furthermore they started the way towards a mathematical formulation of this concept of probability consistent with Kolmogorov's axiomatic theory. Thus, within the Bayesian spirit, it is logical and natural to consider that **probability is a measure of the degree of belief we have in the occurrence of an event** that characterizes the random phenomena and we shall assign probabilities to events based on the prior knowledge we have. In fact, to some extent, all statistical procedures used for the analysis of natural phenomena are subjective inasmuch they all are based on a mathematical idealizations of Nature and all require a priory judgments and hypothesis that have to be assumed.

#### 2.1.4 Random Quantities

In many circumstances, the possible outcomes of the experiments are not numeric (a die with colored faces, a person may be sick or healthy, a particle may decay in different modes,...) and, even in the case they are, the possible outcomes of the experiment may form a non-denumerable set. Ultimately, we would like to deal with numeric values and benefit from the algebraic structures of the real numbers and the theory behind measurable functions and for this, given a measurable space  $(\Omega, \mathcal{B}_\Omega)$ , we define a function  $X(w) : w \in \Omega \rightarrow \mathcal{R}$  that assigns to each event  $w$  of the sample space  $\Omega$  *one and only one real number*.

In a more formal way, consider two measurable spaces  $(\Omega, \mathcal{B}_\Omega)$  and  $(\Omega', \mathcal{B}'_\Omega)$  and a function

$$X(w) : w \in \Omega \rightarrow X(w) \in \Omega' \quad (9)$$

Obviously, since we are interested in the events that conform the  $\sigma$ -algebra  $\mathcal{B}_\Omega$ , the same structure has to be maintained in  $(\Omega', \mathcal{B}'_\Omega)$  by the application  $X(w)$  for otherwise we wont be able to answer the questions of interest. Therefore, we require the function  $X(w)$  to be *Lebesgue measurable with respect to the  $\sigma$ -algebra  $\mathcal{B}_\Omega$* ; i.e.:

$$X^{-1}(B') = B \subseteq \mathcal{B}_\Omega \quad \forall B' \in \mathcal{B}'_\Omega \quad (10)$$

so we can ultimately identify  $P(B')$  with  $P(B)$ . Usually, we are interested in the case that  $\Omega' = \mathcal{R}$  (or  $\mathcal{R}^n$ ) so  $\mathcal{B}'_\Omega$  is the Borel  $\sigma$ -algebra and, since we have generated the Borel algebra  $\mathcal{B}$  from half-open intervals on the left  $I_x = (-\infty, x]$  with  $x \in \mathcal{R}$ , we have that  $X(w)$  will be a Lebesgue measurable function over the Borel algebra (*Borel measurable*) if, and only if:

$$X^{-1}(I_x) = \{w \in \Omega \mid X(w) \leq x\} \in \mathcal{B}_\Omega \quad \forall x \in \mathcal{R} \quad (11)$$

We could have generated as well the Borel algebra from open, closed or half-open intervals on the right so any of the following relations, all equivalent, serve to define a Borel measurable function  $X(w)$ :

- 1)  $\{w|X(w) > c\} \in \mathcal{B}_\Omega \quad \forall c \in \mathcal{R};$
- 3)  $\{w|X(w) \geq c\} \in \mathcal{B}_\Omega \quad \forall c \in \mathcal{R};$
- 4)  $\{w|X(w) < c\} \in \mathcal{B}_\Omega \quad \forall c \in \mathcal{R};$
- 5)  $\{w|X(w) \leq c\} \in \mathcal{B}_\Omega \quad \forall c \in \mathcal{R}$

To summarize:

- Given a probability space  $(\Omega, \mathcal{B}_\Omega, Q)$ , a **random variable** is a function  $X(w) : \Omega \rightarrow \mathcal{R}$ , Borel measurable over the  $\sigma$ -algebra  $\mathcal{B}_\Omega$ , that allows us to work with the induced probability space  $(\mathcal{R}, \mathcal{B}, \mathcal{P})$ <sup>5</sup>.

Form this definition, it is clear that the name “*random variable*” is quite unfortunate inasmuch it is a univoque function, neither random nor variable. Thus, at least to get rid of *variable*, the term “**random quantity**” it is frequently used to design a numerical entity associated to the outcome of an experiment; outcome that is uncertain before we actually do the experiment and observe the result, and distinguish between the random quantity  $X(w)$ , that we shall write in upper cases and usually as  $X$  assuming understood the  $w$  dependence, and the value  $x$  (lower case) taken in a particular realization of the experiment. If the function  $X$  takes values in  $\Omega_X \subseteq \mathcal{R}$  it will be a *one dimensional random quantity* and, if the image is  $\Omega_X \subseteq \mathcal{R}^n$ , it will be an ordered  $n$ -tuple of real numbers  $(X_1, X_2, \dots, X_n)$ . Furthermore, attending to the cardinality of  $\Omega_X$ , we shall talk about *discrete random quantities* if it is finite or countable and about *continuous random quantities* if it is uncountable. This will be explained in more depth in section 3.1. Last, if for each  $w \in \Omega$  is  $|X(w)| < k$  with  $k$  finite, we shall talk about a *bounded random quantity*.

The properties of random quantities are those of the measurable functions. In particular, if  $X(w) : \Omega \rightarrow \Omega'$  is measurable with respect to  $\mathcal{B}_\Omega$  and  $Y(x) : \Omega' \rightarrow \Omega''$  is measurable with respect to  $\mathcal{B}_{\Omega'}$ , the function  $Y(X(w)) : \Omega \rightarrow \Omega''$  is measurable with respect to  $\mathcal{B}_\Omega$  and therefore is a random quantity. We have then that

$$P(Y \leq y) = P(Y(X) \leq y) = P(X \in Y^{-1}(I_X)) \quad (12)$$

where  $Y^{-1}(I_X)$  is the set  $\{x|x \in \Omega'\}$  such that  $Y(x) \leq y$ .

**Example 2.1:** Consider the measurable space  $(\Omega, \mathcal{B}_\Omega)$  and  $X(w) : \Omega \rightarrow \mathcal{R}$ . Then:

- $X(w) = k$ , constant in  $\mathcal{R}$ . Denoting by  $A = \{w \in \Omega | X(w) > c\}$  we have that if  $c \geq k$  then  $A = \emptyset$  and if  $c < k$  then  $A = \Omega$ . Since  $\{\emptyset, E\} \in \mathcal{B}_\Omega$  we conclude that  $X(w)$  is a measurable function. In fact, it is left as an exercise to show that for the minimal algebra  $\mathcal{B}_\Omega^{\min} = \{\emptyset, \Omega\}$ , the only functions that are measurable are  $X(w) = \text{constant}$ .
- Let  $G \in \mathcal{B}_\Omega$  and  $X(w) = \mathbf{1}_G(w)$  (see Note 2). We have that if  $I_a = (-\infty, a]$  with  $a \in \mathcal{R}$ , then  $a \in (-\infty, 0) \rightarrow X^{-1}(I_a) = \emptyset$ ,  $a \in [0, 1) \rightarrow X^{-1}(I_a) = G^c$ , and  $a \in [1, \infty) \rightarrow X^{-1}(I_a) = \Omega$  so  $X(w)$  is a measurable function with respect to  $\mathcal{B}_\Omega$ . A simple function

$$X(w) = \sum_{k=1}^n a_k \mathbf{1}_{A_k}(w) \quad (13)$$

where  $a_k \in \mathcal{R}$  and  $\{A_k\}_{k=1}^n$  is a partition of  $\Omega$  is Borel measurable and any random quantity that takes a finite number of values can be expressed in this way.

- Let  $\Omega = [0, 1]$ . It is obvious that if  $G$  is a non-measurable Lebesgue subset of  $[0, 1]$ , the function  $X(w) = \mathbf{1}_{G^c}(w)$  is not measurable over  $\mathcal{B}_{[0,1]}$  because  $a \in [0, 1) \rightarrow X^{-1}(I_a) = G \notin \mathcal{B}_{[0,1]}$ .
- Consider a coin tossing, the elementary events

$$e_1 = \{H\}, \quad \text{and} \quad e_2 = \{T\} \quad \longrightarrow \quad \Omega = \{e_1, e_2\} \quad (14)$$

<sup>5</sup> It is important to note that a random variable  $X(w) : \Omega \rightarrow \mathcal{R}$  is measurable **with respect to** the  $\sigma$ -algebra  $\mathcal{B}_\Omega$ .

the algebra  $\mathcal{B}_\Omega = \{\emptyset, \Omega, \{e_1\}, \{e_2\}\}$  and the function  $X : \Omega \rightarrow \mathcal{R}$  that denotes the number of heads

$$X(e_1) = 1 \quad \text{and} \quad X(e_2) = 0 \tag{15}$$

Then, for  $I_a = (-\infty, a]$  with  $a \in \mathcal{R}$  we have that:

$$a \in (-\infty, 0) \rightarrow X^{-1}(I_a) = \emptyset \in \mathcal{B}_\Omega \tag{16}$$

$$a \in [0, 1) \rightarrow X^{-1}(I_a) = e_2 \in \mathcal{B}_\Omega \tag{17}$$

$$a \in [1, \infty) \rightarrow X^{-1}(I_a) = \{e_1, e_2\} = \Omega \in \mathcal{B}_\Omega \tag{18}$$

so  $X(w)$  is measurable in  $(\Omega, \mathcal{B}_\Omega, P)$  and therefore an admissible random quantity with  $P(X = 1) = P(e_1)$  and  $P(X = 0) = P(e_2)$ . It will not be an admissible random quantity for the trivial minimum algebra  $\mathcal{B}_\Omega^{min} = \{\emptyset, \Omega\}$  since  $e_2 \notin \mathcal{B}_\Omega^{min}$ .

**Example 2.2:** Let  $\Omega = [0, 1]$  and consider the sequence of functions  $X_n(w) = 2^n \mathbf{1}_{\Omega_n}(w)$  where  $w \in \Omega$ ,  $\Omega_n = [1/2^n, 1/2^{n-1}]$  and  $n \in \mathcal{N}$ . Is each  $X_n(w)$  measurable iff  $\forall r \in \mathcal{R}$ ,  $A = \{w \in \Omega \mid X_n(w) > r\}$  is a Borel set of  $\mathcal{B}_\Omega$ . Then:

- 1)  $r \in (2^n, \infty) \rightarrow A = \emptyset \in \mathcal{B}_\Omega$  with  $\lambda(A) = 0$ ;
- 2)  $r \in [0, 2^n] \rightarrow A = [1/2^n, 1/2^{n-1}] \in \mathcal{B}_\Omega$  with  $\lambda(A) = 2/2^n - 1/2^n = 1/2^n$ .
- 3)  $r \in (-\infty, 0) \rightarrow A = [0, 1] = \Omega$  with  $\lambda(\Omega) = 1$ .

Thus, each  $X_n(w)$  is a measurable function.

**Problem 2.1:** Consider the experiment of tossing two coins, the elementary events

$$e_1 = \{H, H\}, \quad e_2 = \{H, T\}, \quad e_3 = \{T, H\}, \quad e_4 = \{T, T\}$$

the sample space  $\Omega = \{e_1, e_2, e_3, e_4\}$  and the two algebras

$$\begin{aligned} \mathcal{B}_1 &= \{\emptyset, \Omega, \{e_1\}, \{e_4\}, \{e_1, e_2, e_3\}, \{e_2, e_3, e_4\}, \{e_1, e_4\}, \{e_2, e_3\}\} \\ \mathcal{B}_2 &= \{\emptyset, \Omega, \{e_1, e_2\}, \{e_3, e_4\}\} \end{aligned}$$

The functions  $X(w) : \Omega \rightarrow \mathcal{R}$  such that  $X(e_1) = 2; X(e_2) = X(e_3) = 1; X(e_4) = 0$  (number of heads) and  $Y(w) : \Omega \rightarrow \mathcal{R}$  such that  $Y(e_1) = Y(e_2) = 1; Y(e_3) = Y(e_4) = 0$ , with respect to which algebras are admissible random quantities? (sol.:  $X$  wrt  $\mathcal{B}_1$ ;  $Y$  wrt  $\mathcal{B}_2$ )

**Problem 2.2:** Let  $X_i(w) : \mathcal{R} \rightarrow \mathcal{R}$  with  $i = 1, \dots, n$  be random quantities. Show that

$$Y = \max\{X_1, X_2\}, \quad Y = \min\{X_1, X_2\}, \quad Y = \sup\{X_k\}_{k=1}^n \quad \text{and} \quad Y = \inf\{X_k\}_{k=1}^n$$

are admissible random quantities.

Hint: It is enough to observe that

$$\begin{aligned} \{w \mid \max\{X_1, X_2\} \leq x\} &= \{w \mid X_1(w) \leq x\} \cap \{w \mid X_2(w) \leq x\} \in \mathcal{B} \\ \{w \mid \min\{X_1, X_2\} \leq x\} &= \{w \mid X_1(w) \leq x\} \cup \{w \mid X_2(w) \leq x\} \in \mathcal{B} \\ \{w \mid \sup_n X_n(w) \leq x\} &= \bigcap_n \{w \mid X_n(w) \leq x\} \in \mathcal{B} \\ \{w \mid \inf_n X_n(w) < x\} &= \bigcup_n \{w \mid X_n(w) < x\} \in \mathcal{B} \end{aligned}$$

**NOTE 2: Indicator Function.** This is one of the most useful functions in maths. Given subset  $A \subset \Omega$  we define the *Indicator Function*  $\mathbf{1}_A(x)$  for all elements  $x \in \Omega$  as:

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} \tag{19}$$

Given two sets  $A, B \subset \Omega$ , the following relations are obvious:

$$\begin{aligned} \mathbf{1}_{A \cap B}(x) &= \min\{\mathbf{1}_A(x), \mathbf{1}_B(x)\} = \mathbf{1}_A(x) \mathbf{1}_B(x) \\ \mathbf{1}_{A \cup B}(x) &= \max\{\mathbf{1}_A(x), \mathbf{1}_B(x)\} = \mathbf{1}_A(x) + \mathbf{1}_B(x) - \mathbf{1}_A(x) \mathbf{1}_B(x) \\ \mathbf{1}_{A^c}(x) &= 1 - \mathbf{1}_A(x) \end{aligned} \quad (20)$$

It is also called "*Characteristic Function*" but in Probability Theory we reserve this name for the Fourier Transform.

## 2.2 Conditional Probability and Bayes Theorem

Suppose an experiment that consists on rolling a die with faces numbered from one to six and the event  $e_2 = \{ \text{get the number two on the upper face} \}$ . If the die is fair, based on the Principle of Insufficient Reason you and your friend would consider reasonable to assign equal chances to any of the possible outcomes and therefore a probability of  $P_1(e_2) = 1/6$ . Now, if I look at the die and tell you, and only you, that the outcome of the roll is an even number, you will change your beliefs on the occurrence of event  $e_2$  and assign the new value  $P_2(e_2) = 1/3$ . Both of you assign different probabilities because you do not share the same knowledge so it may be a truism but it is clear that *the probability we assign to an event is subjective and is conditioned by the information we have about the random process*. In one way or another, probabilities are always conditional degrees of belief since there is always some state of information (even before we do the experiment we know that whatever number we shall get is not less than one and not greater than six) and we always assume some hypothesis (the die is fair so we can rely on the Principle of Symmetry).

Consider a probability space  $(\Omega, B_\Omega, P)$  and two events  $A, B \subset B_\Omega$  that are not disjoint so  $A \cap B \neq \emptyset$ . The probability for both  $A$  and  $B$  to happen is  $P(A \cap B) \equiv P(A, B)$ . Since  $\Omega = B \cup B^c$  and  $B \cap B^c = \emptyset$  we have that:

$$P(A) \equiv P(A \cap \Omega) = \underbrace{P(A \cap B)}_{\substack{\text{probability for A} \\ \text{and B to occur}}} + \underbrace{P(A \cap B^c) = P(A \setminus B)}_{\substack{P(A \setminus B) : \text{probability for} \\ \text{A to happen and not B}}} \quad (21)$$

What is the probability for  $A$  to happen if we know that  $B$  has occurred? The probability of  $A$  *conditioned* to the occurrence of  $B$  is called **conditional probability of  $A$  given  $B$**  and is expressed as  $P(A|B)$ . This is equivalent to calculate the probability for  $A$  to happen in the probability space  $(\Omega', B'_\Omega, P')$  with  $\Omega'$  the reduced sample space where  $B$  has already occurred and  $B'_\Omega$  the corresponding sub-algebra that does not contain  $B^c$ . We can set  $P(A|B) \propto P(A \cap B)$  and define (Kolmogorov) the conditional probability for  $A$  to happen once  $B$  has occurred as:

$$P(A|B) \stackrel{\text{def.}}{=} \frac{P(A \cap B)}{P(B)} = \frac{P(A, B)}{P(B)} \quad (22)$$

provided that  $P(B) \neq 0$  for otherwise the conditional probability is not defined. This normalization factor ensures that  $P(B|B) = P(B \cap B)/P(B) = 1$ . Conditional probabilities satisfy the basic axioms of probability:

- i) **non-negative** since  $(A \cap B) \subset B \rightarrow 0 \leq P(A|B) \leq 1$
- ii) **unit measure (certainty)** since  $P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$

iii)  $\sigma$ -**additive**: For a countable sequence of disjoint set  $\{A_i\}_{i=1}^{\infty}$

$$P\left(\bigcup_{i=1}^{\infty} A_i|B\right) = \frac{P\left(\left(\bigcup_{i=1}^{\infty} A_i\right) \cap B\right)}{P(B)} = \frac{\sum_{i=1}^{\infty} P(A_i \cap B)}{P(B)} = \sum_{i=1}^{\infty} P(A_i|B) \quad (23)$$

Generalizing, for  $n$  events  $\{A_i\}_{i=1}^n$  we have, with  $j = 0, \dots, n - 1$  that

$$P(A_1, \dots, A_n) = P(A_n, \dots, A_{n-j}|A_j, \dots, A_1)P(A_j, \dots, A_1) = \quad (24)$$

$$= P(A_n|A_1, \dots, A_{n-1})P(A_{n-1}|A_2, \dots, A_1)P(A_2|A_1)P(A_1) \quad (25)$$

### 2.2.1 Statistically Independent Events

Two events  $A, B \in B_{\Omega}$  are **statistically independent** when the occurrence of one does not give any information about the occurrence of the other<sup>6</sup>; that is, when

$$P(A, B) = P(A)P(B) \quad (26)$$

A necessary and sufficient condition for  $A$  and  $B$  to be independent is that  $P(A|B) = P(A)$  (which implies  $P(B|A) = P(B)$ ). Necessary because

$$P(A, B) = P(A)P(B) \quad \longrightarrow \quad P(A|B) = \frac{P(A, B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A) \quad (27)$$

Sufficient because

$$P(A|B) = P(A) \quad \longrightarrow \quad P(A, B) = P(A|B)P(B) = P(A)P(B) \quad (28)$$

If this is not the case, we say that they are *statistically dependent* or **correlated**. In general, we have that:

$P(A|B) > P(A) \rightarrow$  the events  $A$  and  $B$  are **positively correlated**; that is, that  $B$  has already occurred increases the chances for  $A$  to happen;

$P(A|B) < P(A) \rightarrow$  the events  $A$  and  $B$  are **negatively correlated**; that is, that  $B$  has already occurred reduces the chances for  $A$  to happen;

$P(A|B) = P(A) \rightarrow$  the events  $A$  and  $B$  are **not correlated** so the occurrence of  $B$  does not modify the chances for  $A$  to happen.

Given a finite collection of events  $\mathcal{A} = \{A_i\}_{i=1}^n$  with  $A_{\forall i} \subset B_{\Omega}$ , they are statistically independent if

$$P(A_1, \dots, A_m) = P(A_1) \cdots P(A_m) \quad (29)$$

<sup>6</sup> In fact for the events  $A, B \in B_{\Omega}$  we should talk about *conditional independence* for it is true that if  $C \in B_{\Omega}$ , it may happen that  $P(A, B) = P(A)P(B)$  but conditioned on  $C$ ,  $P(A, B|C) \neq P(A|C)P(B|C)$  so  $A$  and  $B$  are related through the event  $C$ . On the other hand, that  $P(A|B) \neq P(A)$  does not imply that  $B$  has a "direct" effect on  $A$ . Whether this is the case or not has to be determined by reasoning on the process and/or additional evidences. Bernard Shaw said that we all should buy an umbrella because there is statistical evidence that doing so you have a higher life expectancy. And this is certainly true. However, it is more reasonable to suppose that instead of the umbrellas having any mysterious influence on our health, in London, at the beginning of the XX<sup>th</sup> century, if you can afford to buy an umbrella you have most likely a well-off status, healthy living conditions, access to medical care,...

for any finite subsequence  $\{A_k\}_{k=j}^m$ ;  $1 \leq j < m \leq n$  of events. Thus, for instance, for a sequence of 3 events  $\{A_1, A_2, A_3\}$  the condition of independence requires that:

$$P(A_1, A_2) = P(A_1)P(A_2) ; \quad P(A_1, A_3) = P(A_1)P(A_3) ; \quad P(A_2, A_3) = P(A_2)P(A_3) \\ \text{and} \quad P(A_1, A_2, A_3) = P(A_1)P(A_2)P(A_3) \quad (30)$$

so the events  $\{A_1, A_2, A_3\}$  may be statistically dependent and be pairwise independent.

**Example 2.3:** In four cards ( $C_1, C_2, C_3$  and  $C_4$ ) we write the numbers 1 ( $C_1$ ), 2 ( $C_2$ ), 3 ( $C_3$ ) and 123 ( $C_4$ ) and make a fair random extraction. Let be the events

$$A_i = \{ \text{the chosen card has the number } i \}$$

with  $i = 1, 2, 3$ . Since the extraction is fair we have that:

$$P(A_i) = P(C_i) + P(C_4) = 1/2 \quad (31)$$

Now, I look at the card and tell you that it has number  $j$ . Since you know that  $A_j$  has happened, you know that the extracted card was either  $C_j$  or  $C_4$  and the only possibility to have  $A_i \neq A_j$  is that the extracted card was  $C_4$  so the conditional probabilities are

$$P(A_i|A_j) = 1/2 ; \quad i, j = 1, 2, 3 ; \quad i \neq j \quad (32)$$

The, since

$$P(A_i|A_j) = P(A_i) ; \quad i, j = 1, 2, 3 ; \quad i \neq j \quad (33)$$

any two events ( $A_i, A_j$ ) are (pairwise) independent. However:

$$P(A_1, A_2, A_3) = P(A_1|A_2, A_3)P(A_2|A_3)P(A_3) \quad (34)$$

and if I tell you that events  $A_2$  and  $A_3$  have occurred then you are certain that chosen card is  $C_4$  and therefore  $A_1$  has happened too so  $P(A_1|A_2, A_3) = 1$ . But

$$P(A_1, A_2, A_3) = 1 \frac{1}{2} \frac{1}{2} \neq P(A_1)P(A_2)P(A_3) = \frac{1}{8} \quad (35)$$

so the events  $\{A_1, A_2, A_3\}$  are **not** independent even though they are pairwise independent.

**Example 2.4: Bonferroni's Inequality.** Given a finite collection  $A = \{A_1, \dots, A_n\} \subset \mathcal{B}$  of events, Bonferroni's inequality states that:

$$P(A_1 \cap \dots \cap A_n) \equiv P(A_1, \dots, A_n) \geq P(A_1) + \dots + P(A_n) - (n - 1) \quad (36)$$

and gives a lower bound for the joint probability  $P(A_1, \dots, A_n)$ . For  $n = 1$  it is trivially true since  $P(A_1) \geq P(A_1)$ . For  $n = 2$  we have that

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \leq 1 \quad \longrightarrow \quad P(A_1 \cap A_2) \geq P(A_1) + P(A_2) - 1 \quad (37)$$

Proceed then by induction. Assume the statement is true for  $n-1$  and see if it is so for  $n$ . If  $B_{n-1} = A_1 \cap \dots \cap A_{n-1}$  and apply the result we got for  $n = 2$  we have that

$$P(A_1 \cap \dots \cap A_n) = P(B_{n-1} \cap A_n) \geq P(B_{n-1}) + P(A_n) - 1 \quad (38)$$

but

$$P(B_{n-1}) = P(B_{n-2} \cap A_{n-1}) \geq P(B_{n-2}) + P(A_{n-1}) - 1 \quad (39)$$

so

$$P(A_1 \cap \dots \cap A_n) \geq P(B_{n-2}) + P(A_{n-1}) + P(A_n) - 2 \quad (40)$$

and therefore the inequality is demonstrated.

### 2.2.2 Theorem of Total Probability

Consider a probability space  $(\Omega, B_\Omega, P)$  and a partition  $\mathcal{S} = \{S_i\}_{i=1}^n$  of the sample space. Then, for any event  $A \in B_\Omega$  we have that  $A = A \cap \Omega = A \cap (\bigcup_{i=1}^n S_i)$  and therefore:

$$P(A) = P\left(A \cap \left[\bigcup_{i=1}^n S_i\right]\right) = P\left(\bigcup_{i=1}^n [A \cap S_i]\right) = \sum_{i=1}^n P(A \cap S_i) = \sum_{i=1}^n P(A|S_i) \cdot P(S_i) \quad (41)$$

Consider now a second different partition of the sample space  $\{B_k\}_{k=1}^m$ . Then, for each set  $B_k$  we have

$$P(B_k) = \sum_{i=1}^n P(B_k|S_i)P(S_i); \quad k = 1, \dots, m \quad (42)$$

and

$$\sum_{k=1}^m P(B_k) = \sum_{i=1}^n P(S_i) \left[ \sum_{k=1}^m P(B_k|S_i) \right] = \sum_{i=1}^n P(S_i) = 1 \quad (43)$$

Last, a similar expression can be written for conditional probabilities. Since

$$P(A, B, S) = P(A|B, S)P(B, S) = P(A|B, S)P(S|B)P(B) \quad (44)$$

and

$$P(A, B) = \sum_{i=1}^n P(A, B, S_i) \quad (45)$$

we have that

$$P(A|B) = \frac{P(A, B)}{P(B)} = \frac{1}{P(B)} \sum_{i=1}^n P(A, B, S_i) = \sum_{i=1}^n P(A|B, S_i)P(S_i|B) \quad (46)$$

**Example 2.5:** We have two indistinguishable urns:  $U_1$  with three white and two black balls and  $U_2$  with two white balls and three black ones. What is the probability that in a random extraction we get a white ball?

Consider the events:

$$A_1 = \{\text{choose urn } U_1\}; \quad A_2 = \{\text{choose urn } U_2\} \quad \text{and} \quad B = \{\text{get a white ball}\} \quad (47)$$

It is clear that  $A_1 \cap A_2 = \emptyset$  and that  $A_1 \cup A_2 = \Omega$ . Now:

$$P(B|A_1) = \frac{3}{5}; \quad P(B|A_2) = \frac{2}{5} \quad \text{and} \quad P(A_1) = P(A_2) = \frac{1}{2} \quad (48)$$

so we have that

$$P(B) = \sum_{i=1}^2 P(B|A_i) \cdot P(A_i) = \frac{3}{5} \frac{1}{2} + \frac{2}{5} \frac{1}{2} = \frac{1}{2} \quad (49)$$

as expected since out of 10 balls, 5 are white.

### 2.2.3 Bayes Theorem

Given a probability space  $(\Omega, B_\Omega, P)$  we have seen that the joint probability for two events  $A, B \in B_\Omega$  can be expressed in terms of conditional probabilities as:

$$P(A, B) = P(A|B)P(B) = P(B|A)P(A) \quad (50)$$

The Bayes Theorem (Bayes ~1770's and independently Laplace ~1770's) states that if  $P(B) \neq 0$ , then

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (51)$$

apparently a trivial statement but with profound consequences. Let's see other expressions of the theorem. If  $\mathcal{H} = \{H_i\}_{i=1}^n$  is a partition of the sample space then

$$P(A, H_i) = P(A|H_i)P(H_i) = P(H_i|A)P(A) \quad (52)$$

and from the Total Probability Theorem

$$P(A) = \sum_{k=1}^n P(A|H_k)P(H_k) \quad (53)$$

so we have a different expression for Bayes' Theorem:

$$P(H_i|A) = \frac{P(A|H_i)P(H_i)}{P(A)} = \frac{P(A|H_i)P(H_i)}{\sum_{k=1}^n P(A|H_k)P(H_k)} \quad (54)$$

Let's summarize the meaning of these terms <sup>7</sup>:

$P(H_i)$  : is the probability of occurrence of the event  $H_i$  before we know if event  $A$  has happened or not; that is, the degree of confidence we have in the occurrence of the event  $H_i$  before we do the experiment so it is called **prior probability**;

$P(A|H_i)$  : is the probability for event  $A$  to happen given that event  $H_i$  has occurred. This may be different depending on  $i = 1, 2, \dots, n$  and when considered as function of  $H_i$  is usually called **likelihood**;

$P(H_i|A)$  : is the degree of confidence we have in the occurrence of event  $H_i$  given that the event  $A$  has happened. The knowledge that the event  $A$  has occurred provides information about the random process and modifies the beliefs we had in  $H_i$  before the experiment was done (expressed by  $P(H_i)$ ) so it is called **posterior probability**;

$P(A)$  : is simply the normalizing factor.

Clearly, if the events  $A$  and  $H_i$  are independent, the occurrence of  $A$  does not provide any information on the chances for  $H_i$  to happen. Whether it has occurred or not does not modify our beliefs about  $H_i$  and therefore  $P(H_i|A) = P(H_i)$ .

In first place, it is interesting to note that the occurrence of  $A$  restricts the sample space for  $\mathcal{H}$  and modifies the prior chances  $P(H_i)$  for  $H_i$  in the same proportion as the occurrence of  $H_i$  modifies the probability for  $A$  because

$$P(A|H_i)P(H_i) = P(H_i|A)P(A) \quad \longrightarrow \quad \frac{P(H_i|A)}{P(H_i)} = \frac{P(A|H_i)}{P(A)} \quad (55)$$

<sup>7</sup>Although is usually the case, the terms *prior* and *posterior* do not necessarily imply a temporal ordering.



Second, from Bayes Theorem we can obtain *relative posterior probabilities* (in the case, for instance, that  $P(A)$  is unknown) because

$$\frac{P(H_i|A)}{P(H_j|A)} = \frac{P(A|H_i)}{P(A|H_j)} \frac{P(H_i)}{P(H_j)} \quad (56)$$

Last, conditioning all the probabilities to  $H_0$  (maybe some conditions that are assumed) we get a third expression of Bayes Theorem

$$P(H_i|A, H_0) = \frac{P(A|H_i, H_0)P(H_i|H_0)}{P(A|H_0)} = \frac{P(A|H_i, H_0)P(H_i|H_0)}{\sum_{k=1}^n P(A|H_k, H_0)P(H_k|H_0)} \quad (57)$$

where  $H_0$  represents to some initial state of information or some conditions that are assumed. The *posterior degree of credibility* we have on  $H_i$  is certainly meaningful when we have an initial degree of information and therefore is relative to our *prior* beliefs. And those are subjective inasmuch different people may assign a different prior degree of credibility based on their previous knowledge and experiences. Think for instance in soccer pools. Different people will assign different prior probabilities to one or other team depending on what they know before the match and this information may not be shared by all of them. However, to the extent that they share common prior knowledge they will arrive to the same conclusions.

Bayes' rule provides a natural way to include new information and update our beliefs in a sequential way. After the event (data)  $D_1$  has been observed, we have

$$P(H_i) \longrightarrow P(H_i|D_1) = \frac{P(D_1|H_i)}{P(D_1)} P(H_i) \propto P(D_1|H_i)P(H_i) \quad (58)$$

Now, if we get additional information provided by the observation of  $D_2$  (new data) we "update" or beliefs on  $H_i$  as:

$$P(H_i|D_1) \longrightarrow P(H_i|D_2, D_1) = \frac{P(D_2|H_i, D_1)}{P(D_2|D_1)} P(H_i|D_1) = \frac{P(D_2|H_i, D_1)P(D_1|H_i)P(H_i)}{P(D_2, D_1)} \quad (59)$$

and so on with further evidences.

**Example 2.6:** An important interpretation of Bayes Theorem is that based on the relation **cause-effect**. Suppose that the event  $A$  (*effect*) has been produced by a certain *cause*  $H_i$ . We consider all possible causes (so  $\mathcal{H}$  is a complete set) and among them we have interest in those that seem more plausible to explain the observation of the event  $A$ . Under this scope, we interpret the terms appearing in Bayes's formula as:

$P(A|H_i, H_0)$  : is the probability that the **effect**  $A$  is produced by the **cause** (or hypothesis)  $H_i$ ;

$P(H_i, H_0)$  : is the prior degree of credibility we assign to the **cause**  $H_i$  before we know that  $A$  has occurred;

$P(H_i|A, H_0)$  : is the posterior probability we have for  $H_i$  being the cause of the event (effect)  $A$  that has already been observed.

Let's see an example of a clinical diagnosis just because the problem is general enough and conclusions may be more disturbing. If you want, replace *individuals* by *events* and for instance (*sick, healthy*) by (*signal, background*). Now, the incidence of certain rare disease is of 1 every 10,000 people and there is an efficient diagnostic test such that:

- 1) If a person is sick, the tests gives positive in 99% of the cases;
- 2) If a person is healthy, the tests may fail and give positive (false positive) in 0.5% of the cases;

In this case, the **effect** is to give positive ( $T$ ) in the test and the exclusive and exhaustive hypothesis for the **cause** are:

$$H_1 : \text{ be sick} \quad \text{and} \quad H_2 : \text{ be healthy} \quad (60)$$

with  $H_2 = H_1^c$ . A person, say you, is chosen **randomly** ( $H_0$ ) among the population to go under the test and give positive. Then you are scared when they tell you: "*The probability of giving positive being healthy is 0.5%, very small*" (p-value). There is nothing wrong with the statement but it has to be correctly interpreted and usually it is not. It means no more and no less than what the expression  $P(T|H_2)$  says: "*under the assumption that you are healthy ( $H_2$ ) the chances of giving positive are 0.5%*" and this is nothing else but a feature of the test. It doesn't say anything about  $P(H_1|T)$ , the chances you have to be sick giving positive in the test that, in the end, is what you are really interested in. The two probabilities are related by an additional piece of information that appears in Bayes's formula:  $P(H_1|H_0)$ ; that is, *under the hypothesis that you have been chosen at random ( $H_0$ ), What are the prior chances to be sick?*. From the prior knowledge we have, the degree of credibility we assign to both hypothesis is

$$P(H_1|H_0) = \frac{1}{10000} \quad \text{and} \quad P(H_2|H_0) = 1 - P(H_1) = \frac{9999}{10000} \quad (61)$$

On the other hand, if  $T$  denotes the event *give positive in the test* we know that:

$$P(T|H_1) = \frac{99}{100} \quad \text{and} \quad P(T|H_2) = \frac{5}{1000} \quad (62)$$

Therefore, Bayes's Theorem tells that the probability to be sick positive in the test is

$$P(H_1|T) = \frac{P(T|H_1) \cdot P(H_1|H_0)}{\sum_{i=1}^2 P(T|H_i) \cdot P(H_i|H_0)} = \frac{\frac{99}{100} \cdot \frac{1}{10000}}{\frac{99}{100} \cdot \frac{1}{10000} + \frac{5}{1000} \cdot \frac{9999}{10000}} \simeq 0.02 \quad (63)$$

Thus, even if the test looks very efficient and you gave positive, the fact that you were chosen at random and that the incidence of the disease in the population is very small, reduces dramatically the degree of belief you assign to be sick. Clearly, if you were not chosen randomly but because there is a suspicion from to other symptoms that you are sic, prior probabilities change.

### 2.3 Distribution Function

A one-dimensional Distribution Function is a real function  $F : \mathcal{R} \rightarrow \mathcal{R}$  that:

- p.1) is monotonous non-decreasing:  $F(x_1) \leq F(x_2) \quad \forall x_1 < x_2 \in \mathcal{R}$
- p.2) is everywhere continuous on the right:  $\lim_{\epsilon \rightarrow 0^+} F(x + \epsilon) = F(x) \quad \forall x \in \mathcal{R}$
- p.3)  $F(-\infty) \equiv \lim_{x \rightarrow -\infty} F(x) = 0$  and  $F(\infty) \equiv \lim_{x \rightarrow +\infty} F(x) = 1$ .

and there is [?] a unique Borel measure  $\mu$  on  $\mathcal{R}$  that satisfies  $\mu((-\infty, x]) = F(x)$  for all  $x \in \mathcal{R}$ . In the Theory of Probability, we define the Probability Distribution Function <sup>8</sup> of the random quantity  $X(w) : \Omega \rightarrow \mathcal{R}$  as:

$$F(x) \stackrel{def.}{=} P(X \leq x) = P(X \in (-\infty, x]) ; \quad \forall x \in \mathcal{R} \quad (64)$$

Note that the Distribution Function  $F(x)$  is defined for all  $x \in \mathcal{R}$  so if  $\text{supp}\{P(X)\} = [a, b]$ , then  $F(x) = 0 \quad \forall x < a$  and  $F(x) = 1 \quad \forall x \geq b$ . From the definition, it is easy to show the following important properties:

<sup>8</sup>The condition  $P(X \leq x)$  is due to the requirement that  $F(x)$  be continuous on the right. This is not essential in the sense that any non-decreasing function  $G(x)$ , defined on  $\mathcal{R}$ , bounded between 0 and 1 and continuous on the left ( $G(x) = \lim_{\epsilon \rightarrow 0^+} G(x - \epsilon)$ ) determines a distribution function defined as  $F(x)$  for all  $x$  where  $G(x)$  is continuous and as  $F(x + \epsilon)$  where  $G(x)$  is discontinuous. In fact, in the general theory of measure it is more common to consider continuity on the left.

a)  $\forall x \in \mathcal{R}$  we have that:

- a.1)  $P(X \leq x) \stackrel{\text{def.}}{=} F(x)$
- a.2)  $P(X < x) = F(x - \epsilon)$  ;
- a.3)  $P(X > x) = 1 - P(X \leq x) = 1 - F(x)$  ;
- a.4)  $P(X \geq x) = 1 - P(X < x) = 1 - F(x - \epsilon)$  ;

b)  $\forall x_1 < x_2 \in \mathcal{R}$  we have that:

- b.1)  $P(x_1 < X \leq x_2) = P(X \in (x_1, x_2]) = F(x_2) - F(x_1)$  ;
- b.2)  $P(x_1 \leq X \leq x_2) = P(X \in [x_1, x_2]) = F(x_2) - F(x_1 - \epsilon)$   
(thus, if  $x_1 = x_2$  then  $P(X = x_1) = F(x_1) - F(x_1 - \epsilon)$ );
- b.3)  $P(x_1 < X < x_2) = P(X \in (x_1, x_2)) = F(x_2 - \epsilon) - F(x_1) =$   
 $= F(x_2) - F(x_1) - P(X = x_2)$  ;
- b.4)  $P(x_1 \leq X < x_2) = P(X \in [x_1, x_2)) = F(x_2 - \epsilon) - F(x_1 - \epsilon) =$   
 $= F(x_2) - F(x_1) - P(X = x_2) + P(X = x_1)$  .

The Distribution Function is discontinuous at all  $x \in \mathcal{R}$  where  $F(x - \epsilon) \neq F(x + \epsilon)$ . Let  $D$  be the set of all points of discontinuity. If  $x \in D$ , then  $F(x - \epsilon) < F(x + \epsilon)$  since it is monotonous non-decreasing. Thus, we can associate to each  $x \in D$  a rational number  $r(x) \in \mathcal{Q}$  such that  $F(x - \epsilon) < r(x) < F(x + \epsilon)$  and all will be different because if  $x_1 < x_2 \in D$  then  $F(x_1 + \epsilon) \leq F(x_2 - \epsilon)$ . Then, since  $\mathcal{Q}$  is a countable set, we have that the set of points of discontinuity of  $F(x)$  is either **finite** or **countable**. At each of them the distribution function has a “jump” of amplitude (property b.2):

$$F(x) - F(x - \epsilon) = P(X = x) \quad (65)$$

and will be continuous on the right (condition p.2).

Last, for each Distribution Function there is a **unique** probability measure  $P$  defined over the Borel sets of  $\mathcal{R}$  that assigns the probability  $F(x_2) - F(x_1)$  to each half-open interval  $(x_1, x_2]$  and, conversely, to any probability measure defined on a measurable space  $(\mathcal{R}, \mathcal{B})$  corresponds one Distribution Function. Thus, the Distribution Function of a random quantity contains all the information needed to describe the properties of the random process.

### 2.3.1 Discrete and Continuous Distribution Functions

Consider the probability space  $(\Omega, \mathcal{F}, Q)$ , the random quantity  $X(w) : w \in \Omega \rightarrow X(w) \in \mathcal{R}$  and the induced probability space  $(\mathcal{R}, \mathcal{B}, P)$ . The function  $X(w)$  is a **discrete random quantity** if its range (image)  $D = \{x_1, \dots, x_i, \dots\}$ , with  $x_i \in \mathcal{R}$ ,  $i = 1, 2, \dots$  is a finite or countable set; that is, if  $\{A_k; k = 1, 2, \dots\}$  is a finite or countable partition of  $\Omega$ , the function  $X(w)$  is either:

$$\text{simple: } X(w) = \sum_{k=1}^n x_k \mathbf{1}_{A_k}(w) \quad \text{or} \quad \text{elementary: } X(w) = \sum_{k=1}^{\infty} x_k \mathbf{1}_{A_k}(w) \quad (66)$$

Then,  $P(X = x_k) = Q(A_k)$  and the corresponding Distribution Function, defined for all  $x \in \mathcal{R}$ , will be

$$F(x) = P(X \leq x) = \sum_{\forall x_k \in D} P(X = x_k) \mathbf{1}_A(x_k) = \sum_{\forall x_k \leq x} P(X = x_k) \quad (67)$$

with  $A = (-\infty, x] \cap D$  and satisfies:

- i)  $F(-\infty) = 0$  and  $F(+\infty) = 1$ ;
- ii) is a monotonous non decreasing step-function;

- iii) continuous on the right ( $F(x + \epsilon) = F(x)$ ) and therefore constant but on the finite or countable set of points of discontinuity  $D = \{x_1, \dots\}$  where

$$F(x_k) - F(x_k - \epsilon) = P(X = x_k) \quad (68)$$

Familiar examples of discrete Distribution Functions are Poisson, Binomial, Multinomial,...

The random quantity  $X(w) : \Omega \rightarrow \mathcal{R}$  is **continuous** if its range is a non-denumerable set; that is, if for all  $x \in \mathcal{R}$  we have that  $P(X = x) = 0$ . In this case, the Distribution Function  $F(x) = P(X \leq x)$  is continuous for all  $x \in \mathcal{R}$  because

- i) from condition (p.2):  $F(x + \epsilon) = F(x)$ ;
- ii) from property (b.2):  $F(x - \epsilon) = F(x) - P(X = x) = F(x)$

Now, consider the measure space  $(\Omega, \mathcal{B}_\Omega, \mu)$  with  $\mu$  countably additive. If  $f : \Omega \rightarrow [0, \infty)$  is integrable with respect to  $\mu$ , it is clear that  $\nu(A) = \int_A f d\mu$  for  $A \in \mathcal{B}_\Omega$  is also a non-negative countably additive set function. More generally, we have:

• **Radon-Nikodym Theorem** (Radon(1913), Nikodym(1930)) : If  $\nu$  and  $\mu$  are two  $\sigma$ -additive measures on the measurable space  $(\Omega, \mathcal{B}_\Omega)$  such that  $\nu$  is **absolutely continuous** with respect to  $\mu$  ( $\nu \ll \mu$ ; that is, for every set  $A \in \mathcal{B}_\Omega$  for which  $\mu(A) = 0$  it is  $\nu(A) = 0$ ), then there exists a  $\mu$ -integrable function  $p(x)$  such that

$$\nu(A) = \int_A d\nu(w) = \int_A \frac{d\nu(w)}{d\mu(w)} d\mu(w) = \int_A p(w) d\mu(w) \quad (69)$$

and, conversely, if such a function exists then  $\nu \ll \mu$  (see appendix 1.3 for the main properties).

The function  $p(w) = d\nu(w)/d\mu(w)$  is called *Radon density* and is unique up to at most a set of measure zero; that is, if

$$\nu(A) = \int_A p(w) d\mu(w) = \int_A f(w) d\mu(w) \quad (70)$$

then then  $\mu\{x | p(x) \neq f(x)\} = 0$ . Furthermore, if  $\nu$  and  $\mu$  are *equivalent* ( $\nu \sim \mu$ ;  $\mu \ll \nu$  and  $\nu \ll \mu$ ) then  $d\nu/d\mu > 0$  almost everywhere. In consequence, if we have a probability space  $(\mathcal{R}, \mathcal{B}, P)$  with  $P$  equivalent to the Lebesgue measure, there exists a non-negative Lebesgue integrable function (see Appendix 2)  $p : \mathcal{R} \rightarrow [0, \infty)$ , unique a.e., such that

$$P(A) \equiv P(X \in A) = \int_A p(x) dx \quad ; \quad \forall A \in \mathcal{B} \quad (71)$$

The function  $p(x)$  is called **probability density function** and satisfies:

- i)  $p(x) \geq 0 \forall x \in \mathcal{R}$ ;
- ii) at any bounded interval of  $\mathcal{R}$ ,  $p(x)$  is bounded and is Riemann-integrable;
- iii)  $\int_{-\infty}^{+\infty} p(x) dx = 1$ .

Thus, for an **absolutely continuous random quantity**  $X$ , the Distribution Function  $F(x)$  can be expressed as

$$F(x) = P(X \leq x) = \int_{-\infty}^x p(w) dw \quad (72)$$

Usually we shall be interested in random quantities that take values in a subset  $D \subset \mathcal{R}$ . It will then be understood that  $p(x)$  is  $p(x)\mathbf{1}_D(x)$  so it is defined for all  $x \in \mathcal{R}$ . Thus, for instance, if  $\text{supp}\{p(x)\} = [a, b]$  then

$$\int_{-\infty}^{+\infty} p(x) dx \equiv \int_{-\infty}^{+\infty} p(x) \mathbf{1}_{[a,b]}(x) dx = \int_a^b p(x) dx = 1 \quad (73)$$

and therefore

$$F(x) = P(X \leq x) = \mathbf{1}_{(a,\infty)}(x)\mathbf{1}_{[b,\infty)}(x) + \mathbf{1}_{[a,b)}(x) \int_a^x p(u) du \quad (74)$$

Note that from the previous considerations, the value of the integral will not be affected if we modify the integrand on a countable set of points. In fact, what we actually integrate is an equivalence class of functions that differ only in a set of measure zero. Therefore, a probability density function  $p(x)$  has to be continuous for all  $x \in \mathcal{R}$  but, at most, on a countable set of points. If  $F(x)$  is not differentiable at a particular point,  $p(x)$  is not defined on it but the set of those points is of zero measure. However, if  $p(x)$  is continuous in  $\mathcal{R}$  then  $F'(x) = p(x)$  and the value of  $p(x)$  is univocally determined by  $F(x)$ . We also have that

$$P(X \leq x) = F(x) = \int_{-\infty}^x p(w) dw \quad \longrightarrow \quad P(X > x) = 1 - F(x) = \int_x^{+\infty} p(w) dw \quad (75)$$

and therefore:

$$P(x_1 < X \leq x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} p(w) dw \quad (76)$$

Thus, since  $F(x)$  is continuous at all  $x \in \mathcal{R}$ :

$$P(x_1 < X \leq x_2) = P(x_1 < X < x_2) = P(x_1 \leq X < x_2) = P(x_1 \leq X \leq x_2) \quad (77)$$

and therefore  $P(X = x) = 0 \forall x \in \mathcal{R}$  ( $\lambda([x]) = 0$ ) even though  $X = x$  is a possible outcome of the experiment so, in this sense, unlike discrete random quantities “probability” 0 does not correspond necessarily to impossible events. Well known examples absolutely continuous Distribution Functions are the Normal, Gamma, Beta, Student, Dirichlet, Pareto, ...

Last, if the continuous probability measure  $P$  is not absolutely continuous with respect to the Lebesgue measure  $\lambda$  in  $\mathcal{R}$ , then the probability density function does not exist. Those are called **singular** random quantities for which  $F(x)$  is continuous but  $F'(x) = 0$  almost everywhere. A well known example is the Dirac’s singular measure  $\delta_{x_0}(A) = \mathbf{1}_A(x_0)$  that assigns a measure 1 to a set  $A \in \mathcal{B}$  if  $x_0 \in A$  and 0 otherwise. As we shall see in the examples 1.9 and 1.20, dealing with these cases is no problem because the Distribution Function always exists. The Lebesgue’s *General Decomposition Theorem* establishes that any Distribution Function can be expressed as a convex combination:

$$F(x) = \sum_{i=1}^{N_d} a_i F_d(x) + \sum_{j=1}^{N_{ac}} b_j F_{ac}(x) + \sum_{k=1}^{N_s} c_k F_s(x) \quad (78)$$

of a discrete Distribution Functions ( $F_d(x)$ ), absolutely continuous ones ( $F_{ac}(x)$  with derivative at every point so  $F'(x) = p(x)$ ) and singular ones ( $F_s(x)$ ). For the cases we shall deal with,  $c_k = 0$ .

**Example 2.7:** Consider a real parameter  $\mu > 0$  and a discrete random quantity  $X$  that can take values  $\{0, 1, 2, \dots\}$  with a Poisson probability law:

$$P(X = k|\mu) = e^{-\mu} \frac{\mu^k}{\Gamma(k+1)} ; \quad k = 0, 1, 2, \dots \quad (79)$$

The Distribution Function will be

$$F(x|\mu) = P(X \leq x|\mu) = e^{-\mu} \sum_{k=0}^{m=[x]} \frac{\mu^k}{\Gamma(k+1)} \quad (80)$$

where  $m = [x]$  is the largest integer less or equal to  $x$ . Clearly, for  $\epsilon \rightarrow 0^+$ :

$$F(x + \epsilon|\mu) = F([x + \epsilon]|\mu) = F([x]|\mu) = F(x|\mu) \quad (81)$$

so it is continuous on the right and for  $k = 0, 1, 2, \dots$

$$F(k|\mu) - F(k - 1|\mu) = P(X = k|\mu) = e^{-\mu} \frac{\mu^k}{\Gamma(k+1)} \quad (82)$$

Therefore, for reals  $x_2 > x_1 > 0$  such that  $x_2 - x_1 < 1$ ,  $P(x_1 < X \leq x_2) = F(x_2) - F(x_1) = 0$ .

**Example 2.8:** Consider the function  $g(x) = e^{-ax}$  with  $a > 0$  real and support in  $(0, \infty)$ . It is non-negative and Riemann integrable in  $\mathcal{R}^+$  so we can define a probability density

$$p(x|a) = \frac{e^{-ax}}{\int_0^\infty e^{-ax} dx} \mathbf{1}_{(0,\infty)}(x) = a e^{-ax} \mathbf{1}_{(0,\infty)}(x) \quad (83)$$

and the Distribution Function

$$F(x) = P(X \leq x) = \int_{-\infty}^x p(u|a) du = \begin{cases} 0 & x < 0 \\ 1 - e^{-ax} & x \geq 0 \end{cases} \quad (84)$$

Clearly,  $F(-\infty) = 0$  and  $F(+\infty) = 1$ . Thus, for an absolutely continuous random quantity  $X \sim p(x|a)$  we have that for reals  $x_2 > x_1 > 0$ :

$$P(X \leq x_1) = F(x_1) = 1 - e^{-ax_1} \quad (85)$$

$$P(X > x_1) = 1 - F(x_1) = e^{-ax_1} \quad (86)$$

$$P(x_1 < X \leq x_2) = F(x_2) - F(x_1) = e^{-ax_1} - e^{-ax_2} \quad (87)$$

**Example 2.9:** The ternary Cantor Set  $C_s(0, 1)$  is constructed iteratively. Starting with the interval  $C_{s_0} = [0, 1]$ , at each step one removes the open middle third of each of the remaining segments. That is; at step one the interval  $(1/3, 2/3)$  is removed so  $C_{s_1} = [0, 1/3] \cup [2/3, 1]$  and so on. If we denote by  $D_n$  the union of the  $2^{n-1}$  disjoint open intervals removed at step  $n$ , each of length  $1/3^n$ , the Cantor set is defined as  $C_s(0, 1) = [0, 1] \setminus \bigcup_{n=1}^\infty D_n$ . It is easy to check that any element  $X$  of the Cantor Set can be expressed as

$$X = \sum_{n=1}^\infty \frac{X_n}{3^n} \quad (88)$$

with  $\text{supp}\{X_n\} = \{0, 2\}$ <sup>9</sup> and that  $C_s(0, 1)$  is a closed set, uncountable, nowhere dense in  $[0, 1]$  and with zero measure. The Cantor Distribution, whose support is the Cantor Set, is defined assigning a probability  $P(X_n = 0) = P(X_n = 2) = 1/2$ . Thus,  $X$  is a continuous random quantity with support on a non-denumerable set of measure zero and can not be described by a probability density function. The Distribution Function  $F(x) = P(X \leq x)$  (Cantor Function; Figure ??) is an example of singular Distribution.

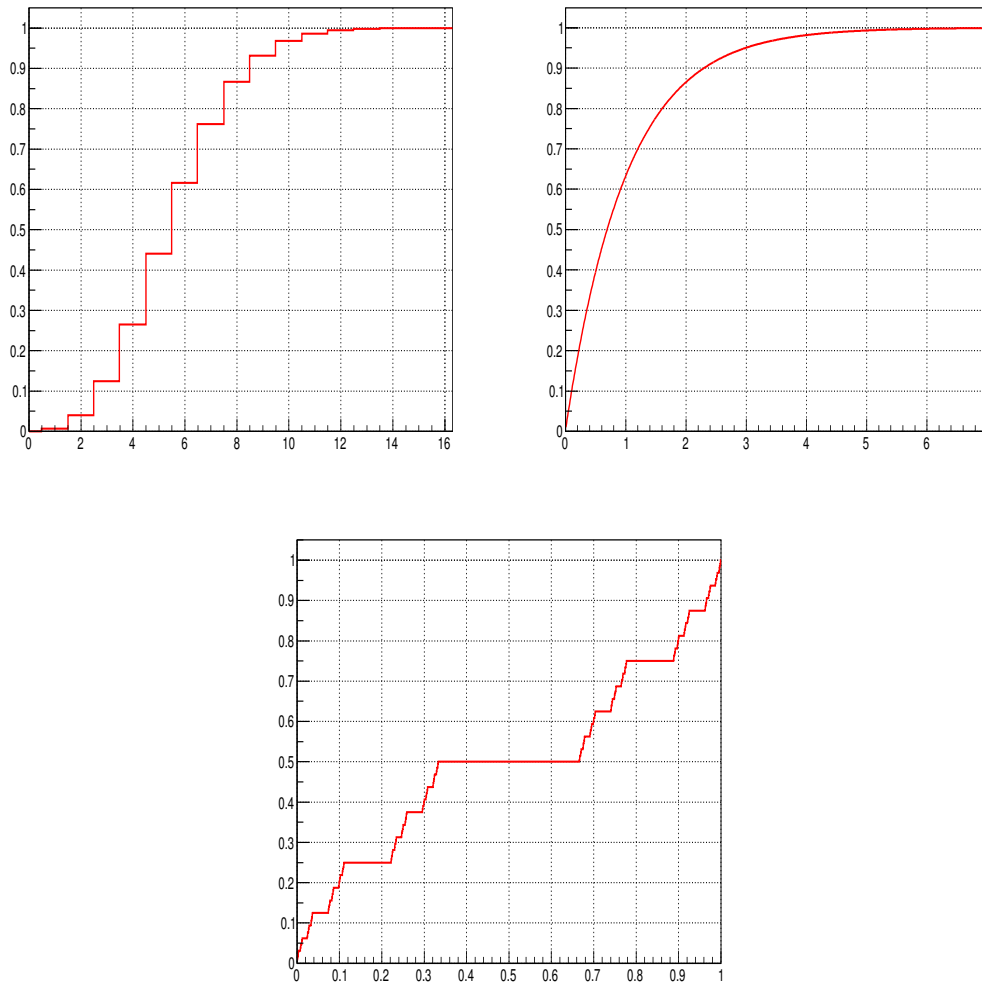
### 2.3.2 Distributions in more dimensions

The previous considerations can be extended to random quantities in more dimensions but with some care. Let's consider the two-dimensional case:  $\mathbf{X} = (X_1, X_2)$ . The Distribution Function will be defined as:

$$F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2) ; \quad \forall (x_1, x_2) \in \mathcal{R}^2 \quad (89)$$

and satisfies:

<sup>9</sup>Note that the representation of a real number  $r \in [0, 1]$  as  $(a_1, a_2, \dots) : \sum_{n=1}^\infty a_n 3^{-n}$  with  $a_i = \{0, 1, 2\}$  is not unique. In fact  $x = 1/3 \in C_s(0, 1)$  and can be represented by  $(1, 0, 0, 0, \dots)$  or  $(0, 2, 2, 2, \dots)$ .



**Fig. 1:** Empiric Distribution Functions (ordinate) form a Monte Carlo sampling ( $10^6$  events) of the Poisson  $Po(x|5)$  (discrete; upper left), Exponential  $Ex(x|1)$  (absolute continuous; upper right) and Cantor (singular; bottom) Distributions.

i) monotonous no-decreasing in both variables; that is, if  $(x_1, x_2), (x'_1, x'_2) \in \mathcal{R}^2$ :

$$x_1 \leq x'_1 \longrightarrow F(x_1, x_2) \leq F(x'_1, x_2) \quad \text{and} \quad x_2 \leq x'_2 \longrightarrow F(x_1, x_2) \leq F(x_1, x'_2) \quad (90)$$

ii) continuous on the right at  $(x_1, x_2) \in \mathcal{R}^2$ :

$$F(x_1 + \epsilon, x_2) = F(x_1, x_2 + \epsilon) = F(x_1, x_2) \quad (91)$$

iii)  $F(-\infty, x_2) = F(x_1, -\infty) = 0$  and  $F(+\infty, +\infty) = 1$ .

Now, if  $(x_1, x_2), (x'_1, x'_2) \in \mathcal{R}^2$  with  $x_1 < x'_1$  and  $x_2 < x'_2$  we have that:

$$P(x_1 < X_1 \leq x'_1, x_2 < X_2 \leq x'_2) = F(x'_1, x'_2) - F(x_1, x'_2) - F(x'_1, x_2) + F(x_1, x_2) \geq 0 \quad (92)$$

and

$$P(x_1 \leq X_1 \leq x'_1, x_2 \leq X_2 \leq x'_2) = F(x'_1, x'_2) - F(x_1 - \epsilon, x'_2) - F(x'_1, x_2 - \epsilon) + F(x_1 - \epsilon, x_2 - \epsilon) \geq 0 \quad (93)$$

so, for discrete random quantities, if  $x_1 = x'_1$  and  $x_2 = x'_2$ :

$$P(X_1 = x_1, X_2 = x_2) = F(x_1, x_2) - F(x_1 - \epsilon_1, x_2) - F(x_1, x_2 - \epsilon) + F(x_1 - \epsilon, x_2 - \epsilon) \geq 0 \quad (94)$$

will give the amplitude of the jump of the Distribution Function at the points of discontinuity.

As for the one-dimensional case, for absolutely continuous random quantities we can introduce a two-dimensional probability density function  $p(\mathbf{x}) : \mathcal{R}^2 \rightarrow \mathcal{R}$ :

- i)  $p(\mathbf{x}) \geq 0 ; \quad \forall \mathbf{x} \in \mathcal{R}^2$ ;
- ii) At every bounded interval of  $\mathcal{R}^2$ ,  $p(\mathbf{x})$  is bounded and Riemann integrable;
- iii)  $\int_{\mathcal{R}^2} p(\mathbf{x}) d\mathbf{x} = 1$

such that:

$$F(x_1, x_2) = \int_{-\infty}^{x_1} du_1 \int_{-\infty}^{x_2} du_2 p(u_1, u_2) \quad \longleftrightarrow \quad p(x_1, x_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F(x_1, x_2) \quad (95)$$

### 2.3.2.1 Marginal and Conditional Distributions

It may happen that we are interested only in one of the two random quantities say, for instance,  $X_1$ . Then we ignore all aspects concerning  $X_2$  and obtain the one-dimensional Distribution Function

$$F_1(x_1) = P(X_1 \leq x_1) = P(X_1 \leq x_1, X_2 \leq +\infty) = F(x_1, +\infty) \quad (96)$$

that is called the **Marginal Distribution Function** of the random quantity  $X_1$ . In the same manner, we have  $F_2(x_2) = F(+\infty, x_2)$  for the random quantity  $X_2$ . For absolutely continuous random quantities,

$$F_1(x_1) = F(x_1, +\infty) = \int_{-\infty}^{x_1} du_1 \int_{-\infty}^{+\infty} p(u_1, u_2) du_2 = \int_{-\infty}^{x_1} p(u_1) du_1 \quad (97)$$

with  $p(x_1)$  the **marginal probability density function**<sup>10</sup> of the random quantity  $X_1$ :

$$p_1(x_1) = \frac{\partial}{\partial x_1} F_1(x_1) = \int_{-\infty}^{+\infty} p(x_1, u_2) du_2 \quad (98)$$

In the same manner, we have for  $X_2$

$$p_2(x_2) = \frac{\partial}{\partial x_2} F_2(x_2) = \int_{-\infty}^{+\infty} p(u_1, x_2) du_1 \quad (99)$$

As we have seen, given a probability space  $(\Omega, B_\Omega, P)$ , for any two sets  $A, B \in B_\Omega$  the conditional probability for  $A$  given  $B$  was *defined* as

$$P(A|B) \stackrel{def.}{=} \frac{P(A \cap B)}{P(B)} \equiv \frac{P(A, B)}{P(B)} \quad (100)$$

provided  $P(B) \neq 0$ . Intimately related to this definition is the Bayes' rule:

$$P(A, B) = P(A|B) P(B) = P(B|A) P(A) \quad (101)$$

Consider now the discrete random quantity  $\mathbf{X} = (X_1, X_2)$  with values on  $\Omega \subset \mathcal{R}^2$ . It is then natural to define

$$P(X_1 = x_1 | X_2 = x_2) \stackrel{def.}{=} \frac{P(X_1 = x_1, X_2 = x_2)}{P(X_2 = x_2)} \quad (102)$$

---

<sup>10</sup>It is habitual to avoid the indices and write  $p(x)$  meaning "the probability density function of the variable  $x$ " since the distinctive features are clear within the context.



and therefore

$$F(x_1|x_2) = \frac{P(X_1 \leq x_1, X_2 = x_2)}{P(X_2 = x_2)}$$

whenever  $P(X_2 = x_2) \neq 0$ . For absolutely continuous random quantities we can express the probability density as

$$p(x_1, x_2) = p(x_1|x_2) p(x_2) = p(x_2|x_1) p(x_1) \quad (103)$$

and define the **conditional probability density function** as

$$p(x_1|x_2) \stackrel{\text{def.}}{=} \frac{p(x_1, x_2)}{p(x_2)} = \frac{\partial}{\partial x_1} F(x_1|x_2) \quad (104)$$

provided again that  $p_2(x_2) \neq 0$ . This is certainly is an admissible density <sup>11</sup> since  $p(x_1|x_2) \geq 0$   $\forall (x_1, x_2) \in \mathcal{R}^2$  and  $\int_{\mathcal{R}} p(x_1|x_2) dx_1 = 1$ .

As stated already, two events  $A, B \in \mathcal{B}_\Omega$  are statistically independent iff:

$$P(A, B) \equiv P(A \cap B) = P(A) P(B) \quad (105)$$

Then, we shall say that two discrete random quantities  $X_1$  and  $X_2$  are **statistically independent** if  $F(x_1, x_2) = F_1(x_1)F_2(x_2)$ ; that is, if

$$P(X_1 = x_1, X_2 = x_2) = P(X_1 = x_1) P(X_2 = x_2) \quad (106)$$

for discrete random quantities and

$$p(x_1, x_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F(x_1)F(x_2) = p(x_1) p(x_2) \iff p(x_1|x_2) = p(x_1) \quad (107)$$

and for absolutely continuous random quantities.

**Example 2.10:** Consider the probability space  $(\Omega, \mathcal{B}_\Omega, \lambda)$  with  $\Omega = [0, 1]$  and  $\lambda$  the Lebesgue measure. If  $F$  is an arbitrary Distribution Function,  $X : w \in [0, 1] \rightarrow F^{-1}(w) \in \mathcal{R}$  is a random quantity and is distributed as  $F(w)$ . Take the Borel set  $I = (-\infty, r]$  with  $r \in \mathcal{R}$ . Since  $F$  is a Distribution Function is monotonous and non-decreasing we have that:

$$X^{-1}(I) = \{w \in \Omega \mid X(w) \leq r\} = \{w \in [0, 1] \mid F^{-1}(w) \leq r\} = \{w \in \Omega \mid w \leq F(r)\} = [0, F(r)] \in \mathcal{B}_\Omega \quad (108)$$

and therefore  $X(w) = F^{-1}(w)$  is measurable over  $\mathcal{B}_\mathcal{R}$  and is distributed as

$$P(X(w) \leq x) = P(F^{-1}(w) \leq x) = P(w \leq F(x)) = \int_0^{F(x)} d\lambda = F(x) \quad (109)$$

**Example 2.11:** Consider the probability space  $(\mathcal{R}, \mathcal{B}, \mu)$  with  $\mu$  the probability measure

$$\mu(A) = \int_{A \in \mathcal{B}} dF \quad (110)$$

<sup>11</sup>Recall that for continuous random quantities  $P(X_2 = x_2) = P(X_1 = x_1) = 0$ . One can justify this expression with kind of heuristic arguments; essentially considering  $X_1 \in \Lambda_1 = (-\infty, x_1]$ ,  $X_2 \in \Delta_\epsilon(x_2) = [x_2, x_2 + \epsilon]$  and taking the limit  $\epsilon \rightarrow 0^+$  of

$$P(X_1 \leq x_1 \mid X_2 \in \Delta_\epsilon(x_2)) = \frac{P(X_1 \leq x_1, X_2 \in \Delta_\epsilon(x_2))}{P(X_2 \in \Delta_\epsilon(x_2))} = \frac{F(x_1, x_2 + \epsilon) - F(x_1, x_2)}{F_2(x_2 + \epsilon) - F_2(x_2)}$$

See however [?]; Vol 2; Chap. 10. for the Radon-Nikodym density with conditional measures.

The function  $X : w \in \mathcal{R} \rightarrow F(w) \in [0, 1]$  is measurable on  $\mathcal{B}$ . Take  $I = [a, b] \in \mathcal{B}_{[0,1]}$ . Then

$$X^{-1}(I) = \{w \in \mathcal{R} \mid a \leq F(w) < b\} = \{w \in \mathcal{R} \mid F^{-1}(a) \leq w < F^{-1}(b)\} = [w_a, w_b) \in \mathcal{B}_{\mathcal{R}} \quad (111)$$

It is distributed as  $X \sim Un(x|0, 1)$ :

$$P(X(w) \leq x) = P(F(w) \leq x) = P(w \leq F^{-1}(x)) = \int_{-\infty}^{F^{-1}(x)} dF = x \quad (112)$$

This is the basis of the Inverse Transform sampling method that we shall see in Lecture 3 on Monte Carlo techniques.

**Example 2.12:** Suppose that the number of eggs a particular insect may lay ( $X_1$ ) follows a Poisson distribution  $X_1 \sim Po(x_1|\mu)$ :

$$P(X_1 = x_1|\mu) = e^{-\mu} \frac{\mu^{x_1}}{\Gamma(x_1 + 1)} ; \quad x_1 = 0, 1, 2, \dots \quad (113)$$

Now, if the probability for an egg to hatch is  $\theta$  and  $X_2$  represent the number of off springs, given  $x_1$  eggs the probability to have  $x_2$  descendants follows a Binomial law  $X_2 \sim Bi(x_2|x_1, \theta)$ :

$$P(X_2 = x_2|x_1, \theta) = \binom{x_1}{x_2} \theta^{x_2} (1 - \theta)^{x_1 - x_2} ; \quad 0 \leq x_2 \leq x_1 \quad (114)$$

In consequence

$$\begin{aligned} P(X_1 = x_1, X_2 = x_2|\mu, \theta) &= P(X_2 = x_2|X_1 = x_1, \theta) P(X_1 = x_1|\mu) = \\ &= \binom{x_1}{x_2} \theta^{x_2} (1 - \theta)^{x_1 - x_2} e^{-\mu} \frac{\mu^{x_1}}{\Gamma(x_1 + 1)} ; \quad 0 \leq x_2 \leq x_1 \end{aligned} \quad (115)$$

Suppose that we have not observed the number of eggs that were laid. What is the distribution of the number of off springs? This is given by the marginal probability

$$P(X_2 = x_2|\theta, \mu) = \sum_{x_1=x_2}^{\infty} P(X_1 = x_1, X_2 = x_2) = e^{-\mu\theta} \frac{(\mu\theta)^{x_2}}{\Gamma(x_2 + 1)} = Po(x_2|\mu\theta) \quad (116)$$

Now, suppose that we have found  $x_2$  new insects. What is the distribution of the number of eggs laid? This will be the conditional probability  $P(X_1 = x_1|X_2 = x_2, \theta, \mu)$  and, since  $P(X_1 = x_1, X_2 = x_2) = P(X_1 = x_1|X_2 = x_2)P(X_2 = x_2)$  we have that:

$$P(X_1 = x_1|X_2 = x_2, \mu, \theta) = \frac{P(X_1 = x_1, X_2 = x_2)}{P(X_2 = x_2)} = \frac{1}{(x_1 - x_2)!} (\mu(1 - \theta))^{x_1 - x_2} e^{-\mu(1 - \theta)} \quad (117)$$

with  $0 \leq x_2 \leq x_1$ ; that is, again a Poisson with parameter  $\mu(1 - \theta)$ .

**Example 2.13:** Let  $X_1$  and  $X_2$  two independent Poisson distributed random quantities with parameters  $\mu_1$  and  $\mu_2$ . How is  $Y = X_1 + X_2$  distributed? Since they are independent:

$$P(X_1 = x_1, X_2 = x_2|\mu_1, \mu_2) = e^{-(\mu_1 + \mu_2)} \frac{\mu_1^{x_1}}{\Gamma(x_1 + 1)} \frac{\mu_2^{x_2}}{\Gamma(x_2 + 1)} \quad (118)$$

Then, since  $X_2 = Y - X_1$ :

$$P(X_1 = x, Y = y) = P(X_1 = x, X_2 = y - x) = e^{-(\mu_1 + \mu_2)} \frac{\mu_1^x}{\Gamma(x + 1)} \frac{\mu_2^{(y-x)}}{\Gamma(y - x + 1)} \quad (119)$$

Being  $X_2 = y - x \geq 0$  we have the condition  $y \geq x$  so the marginal probability for  $Y$  will be

$$P(Y = y) = e^{-(\mu_1 + \mu_2)} \sum_{x=0}^y \frac{\mu_1^x}{\Gamma(x + 1)} \frac{\mu_2^{(y-x)}}{\Gamma(y - x + 1)} = e^{-(\mu_1 + \mu_2)} \frac{(\mu_1 + \mu_2)^y}{\Gamma(y + 1)} \quad (120)$$

that is,  $Po(y|\mu_1 + \mu_2)$ .

**Example 2.14:** Consider a two-dimensional random quantity  $X = (X_1, X_2)$  that takes values in  $\mathcal{R}^2$  with the probability density function  $N(x_1, x_2|\mu = \mathbf{0}, \sigma = \mathbf{1}, \rho)$ :

$$p(x_1, x_2|\rho) = \frac{1}{2\pi} \frac{1}{\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(x_1^2 - 2\rho x_1 x_2 + x_2^2)} \quad (121)$$

with  $\rho \in (-1, 1)$ . The marginal densities are:

$$X_1 \sim p(x_1) = \int_{-\infty}^{+\infty} p(x_1, u_2) du_2 = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_1^2} \quad (122)$$

$$X_2 \sim p(x_2) = \int_{-\infty}^{+\infty} p(u_1, x_2) du_1 = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_2^2} \quad (123)$$

and since

$$p(x_1, x_2|\rho) = p(x_1)p(x_2) \frac{1}{\sqrt{1-\rho^2}} e^{-\frac{\rho}{2(1-\rho^2)}(x_1^2\rho - 2x_1x_2 + x_2^2\rho)}$$

both quantities will be independent iff  $\rho = 0$ . The conditional densities are

$$p(x_1|x_2, \rho) = \frac{p(x_1, x_2)}{p(x_2)} = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(x_1 - x_2\rho)^2} \quad (124)$$

$$p(x_2|x_1, \rho) = \frac{f(x_1, x_2)}{f_1(x_1)} = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(x_2 - x_1\rho)^2} \quad (125)$$

and when  $\rho = 0$  (thus independent)  $p(x_1|x_2) = p(x_1)$  and  $p(x_2|x_1) = p(x_2)$ . Last, it is clear that

$$p(x_1, x_2|\rho) = p(x_2|x_1, \rho)p(x_1) = p(x_1|x_2, \rho)p(x_2) \quad (126)$$

## 2.4 Stochastic Characteristics

### 2.4.1 Mathematical Expectation

Consider a random quantity  $X(w) : \Omega \rightarrow \mathcal{R}$  that can be either discrete

$$X(w) = \begin{cases} X(w) = \sum_{k=1}^n x_k \mathbf{1}_{A_k}(w) \\ X(w) = \sum_{k=1}^{\infty} x_k \mathbf{1}_{A_k}(w) \end{cases} \longrightarrow P(X = x_k) = P(A_k) = \int_{\mathcal{R}} \mathbf{1}_{A_k}(w) dP(w) \quad (127)$$

or absolutely continuous for which

$$P(X(w) \in A) = \int_{\mathcal{R}} \mathbf{1}_A(w) dP(w) = \int_A dP(w) = \int_A p(w) dw \quad (128)$$

The **mathematical expectation** of a n-dimensional random quantity  $\mathbf{Y} = g(\mathbf{X})$  is defined as <sup>12</sup>:

$$E[\mathbf{Y}] = E[g(\mathbf{X})] \stackrel{def.}{=} \int_{\mathcal{R}^n} g(\mathbf{x}) dP(\mathbf{x}) = \int_{\mathcal{R}^n} g(\mathbf{x}) p(\mathbf{x}) d\mathbf{x} \quad (129)$$

<sup>12</sup> In what follows we consider the Stieltjes-Lebesgue integral so  $\int \rightarrow \sum$  for discrete random quantities and in consequence:

$$\int_{-\infty}^{\infty} g(x) dP(x) = \int_{-\infty}^{\infty} g(x) p(x) dx \longrightarrow \sum_{\forall x_k} g(x_k) P(X = x_k)$$

In general, the function  $g(x)$  will be unbounded on  $\text{supp}\{X\}$  so both the sum and the integral have to be **absolutely convergent** for the *mathematical expectation* to exist.

In a similar way, we define the *conditional expectation*. If  $\mathbf{X} = (X_1, \dots, X_m, \dots, X_n)$ ,  $\mathbf{W} = (X_1, \dots, X_m)$  and  $\mathbf{Z} = (X_{m+1}, \dots, X_n)$  we have for  $\mathbf{Y} = g(\mathbf{W})$  that

$$E[\mathbf{Y}|\mathbf{Z}_0] = \int_{\mathcal{R}^m} g(\mathbf{w}) p(\mathbf{w}|\mathbf{z}_0) d\mathbf{w} = \int_{\mathcal{R}^m} g(\mathbf{w}) \frac{p(\mathbf{w}, \mathbf{z}_0)}{p(\mathbf{z}_0)} d\mathbf{w} \quad (130)$$

#### 2.4.2 Moments of a Distribution

Given a random quantity  $X \sim p(x)$ , we define the *moment or order  $n$*  ( $\alpha_n$ ) as:

$$\alpha_n = E[X^n] \stackrel{\text{def.}}{=} \int_{-\infty}^{\infty} x^n p(x) dx \quad (131)$$

Obviously, they exist if  $x^n p(x) \in L_1(\mathcal{R})$  so it may happen that a particular probability distribution has only a finite number of moments. It is also clear that if the moment of order  $n$  exists, so do the moments of lower order and, if it does not, neither those of higher order. In particular, the moment of order 0 always exists (that, due to the normalization condition, is  $\alpha_0 = 1$ ) and those of even order, if exist, are non-negative. A specially important moment is that order 1: the **mean** (*mean value*)  $\mu = E[X]$  that has two important properties:

- It is a **linear operator** since  $X = c_0 + \sum_{i=1}^n c_i X_i \rightarrow E[X] = c_0 + \sum_{i=1}^n c_i E[X_i]$
- If  $X = \prod_{i=1}^n c_i X_i$  with  $\{X_i\}_{i=1}^n$  independent random quantities, then  $E[X] = \prod_{i=1}^n c_i E[X_i]$ .

We can define as well the moments ( $\beta_n$ ) with respect to any point  $c \in \mathcal{R}$  as:

$$\beta_n = E[(X - c)^n] \stackrel{\text{def.}}{=} \int_{-\infty}^{\infty} (x - c)^n p(x) dx \quad (132)$$

so  $\alpha_n$  are also called *central moments or moments with respect to the origin*. It is easy to see that the non-central moment of second order,  $\beta_2 = E[(X - c)^2]$ , is minimal for  $c = \mu = E[X]$ . Thus, of special relevance are the *moments or order  $n$  with respect to the mean*

$$\mu_n \equiv E[(X - \mu)^n] = \int_{-\infty}^{\infty} (x - \mu)^n p(x) dx \quad (133)$$

and, among them, the moment of order 2: the **variance**  $\mu_2 = V[X] = \sigma^2$ . It is clear that  $\mu_0 = 1$  and, if exists,  $\mu_1 = 0$ . Note that:

- $V[X] = \sigma^2 = E[(X - \mu)^2] > 0$
- It is **not a linear operator** since  $X = c_0 + c_1 X_1 \rightarrow V[X] = \sigma_X^2 = c_1^2 V[X_1] = c_1^2 \sigma_{X_1}^2$
- If  $X = \sum_{i=1}^n c_i X_i$  and  $\{X_i\}_{i=1}^n$  are independent random quantities,  $V[X] = \sum_{i=1}^n c_i^2 V[X_i]$ .

Usually, is less tedious to calculate the moments with respect to the origin and evidently they are related so, from the binomial expansion

$$(X - \mu)^n = \sum_{k=0}^n \binom{n}{k} X^k (-\mu)^{n-k} \rightarrow \mu_n = \sum_{k=0}^n \binom{n}{k} \alpha_k (-\mu)^{n-k} \quad (134)$$

The previous definitions are trivially extended to  $n$ -dimensional random quantities. In particular, for 2 dimensions,  $\mathbf{X} = (X_1, X_2)$ , we have the moments of order  $(n, m)$  with respect to the origin:

$$\alpha_{nm} = E[X_1^n X_2^m] = \int_{\mathcal{R}^2} x_1^n x_2^m p(x_1, x_2) dx_1 dx_2 \quad (135)$$

so that  $\alpha_{01} = \mu_1$  and  $\alpha_{02} = \mu_2$ , and the moments order  $(n, m)$  with respect to the mean:

$$\mu_{nm} = E[(X_1 - \mu_1)^n (X_2 - \mu_2)^m] = \int_{\mathcal{R}^2} (x_1 - \mu_1)^n (x_2 - \mu_2)^m p(x_1, x_2) dx_1 dx_2 \quad (136)$$

for which

$$\mu_{20} = E[(X_1 - \mu_1)^2] = V[X_1] = \sigma_1^2 \quad \text{and} \quad \mu_{02} = E[(X_2 - \mu_2)^2] = V[X_2] = \sigma_2^2 \quad (137)$$

The moment

$$\mu_{11} = E[(X_1 - \mu_1)(X_2 - \mu_2)] = \alpha_{11} - \alpha_{10}\alpha_{01} = V[X_1, X_2] = V[X_2, X_1] \quad (138)$$

is called **covariance** between the random quantities  $X_1$  and  $X_2$  and, if they are independent,  $\mu_{11} = 0$ . The second order moments with respect to the mean can be condensed in matrix form, the **covariance matrix** defined as:

$$\mathbf{V}[\mathbf{X}] = \begin{pmatrix} \mu_{20} & \mu_{11} \\ \mu_{11} & \mu_{02} \end{pmatrix} = \begin{pmatrix} V[X_1, X_1] & V[X_1, X_2] \\ V[X_1, X_2] & V[X_2, X_2] \end{pmatrix} \quad (139)$$

Similarly, for  $\mathbf{X} = (X_1, X_2, \dots, X_n)$  we have the moments with respect to the origin

$$\alpha_{k_1, k_2, \dots, k_n} = E[X_1^{k_1} X_2^{k_2} \dots X_n^{k_n}]; \quad (140)$$

the moments with respect to the mean

$$\mu_{k_1, k_2, \dots, k_n} = E[(X_1 - \mu_1)^{k_1} (X_2 - \mu_2)^{k_2} \dots (X_n - \mu_n)^{k_n}] \quad (141)$$

and the covariance matrix:

$$\mathbf{V}[\mathbf{X}] = \begin{pmatrix} \mu_{20\dots 0} & \mu_{11\dots 0} & \cdots & \mu_{10\dots 1} \\ \mu_{11\dots 0} & \mu_{02\dots 0} & \cdots & \mu_{01\dots 1} \\ \vdots & \vdots & \cdots & \vdots \\ \mu_{10\dots 1} & \mu_{01\dots 1} & \cdots & \mu_{00\dots 2} \end{pmatrix} = \begin{pmatrix} V[X_1, X_1] & V[X_1, X_2] & \cdots & V[X_1, X_n] \\ V[X_1, X_2] & V[X_2, X_2] & \cdots & V[X_2, X_n] \\ \vdots & \vdots & \cdots & \vdots \\ V[X_1, X_n] & V[X_2, X_n] & \cdots & V[X_n, X_n] \end{pmatrix} \quad (142)$$

The covariance matrix  $\mathbf{V}[\mathbf{X}] = E[(\mathbf{X} - \mu)(\mathbf{X} - \mu)^T]$  has the following properties that are easy to prove from basic matrix algebra relations:

- 1) It is a **symmetric** matrix ( $\mathbf{V} = \mathbf{V}^T$ ) with **positive diagonal elements** ( $V_{ii} \geq 0$ );
- 2) It is **positive defined** ( $\mathbf{x}^T \mathbf{V} \mathbf{x} \geq 0$ ;  $\forall \mathbf{x} \in \mathcal{R}^n$  with the equality when  $\forall i x_i = 0$ );
- 3) Being  $\mathbf{V}$  symmetric, all the eigenvalues are real and the corresponding eigenvectors orthogonal. Furthermore, since it is positive defined all eigenvalues are positive;
- 4) If  $\mathbf{J}$  is a diagonal matrix whose elements are the eigenvalues of  $\mathbf{V}$  and  $\mathbf{H}$  a matrix whose columns are the corresponding eigenvectors, then  $\mathbf{V} = \mathbf{H} \mathbf{J} \mathbf{H}^{-1}$  (Jordan dicit);
- 5) Since  $\mathbf{V}$  is symmetric, there is an orthogonal matrix  $\mathbf{C}$  ( $\mathbf{C}^T = \mathbf{C}^{-1}$ ) such that  $\mathbf{C} \mathbf{V} \mathbf{C}^T = \mathbf{D}$  with  $\mathbf{D}$  a diagonal matrix whose elements are the eigenvalues of  $\mathbf{V}$ ;
- 6) Since  $\mathbf{V}$  is symmetric and positive defined, there is a non-singular matrix  $\mathbf{C}$  such that  $\mathbf{V} = \mathbf{C} \mathbf{C}^T$ ;
- 7) Since  $\mathbf{V}$  is symmetric and positive defined, the inverse  $\mathbf{V}^{-1}$  is also symmetric and positive defined;
- 8) (Cholesky Factorization) Since  $\mathbf{V}$  is symmetric and positive defined, there exists a unique lower triangular matrix  $\mathbf{C}$  ( $C_{ij} = 0$ ;  $\forall i < j$ ) with positive diagonal elements such that  $\mathbf{V} = \mathbf{C} \mathbf{C}^T$  (more about this in section 3).

Among other things to be discussed later, the moments of the distribution are interesting because they give an idea of the shape and location of the probability distribution and, in many cases, the distribution parameters are expressed in terms of the moments.

### 2.4.2.1 Position parameters

Let  $X \sim p(x)$  with support in  $\Omega \subset \mathcal{R}$ . The position parameters choose a *characteristic* value of  $X$  and indicate more or less where the distribution is located. Among them we have the **mean value**

$$\mu = \alpha_1 = E[X] = \int_{-\infty}^{\infty} x p(x) dx \quad (143)$$

The mean is bounded by the minimum and maximum values the random quantity can take but, clearly, if  $\Omega \subset \mathcal{R}$  it may happen that  $\mu \notin \Omega$ . If, for instance,  $\Omega = \Omega_1 \cup \Omega_2$  is the union of two disconnected regions,  $\mu$  may lay in between and therefore  $\mu \notin \Omega$ . On the other hand, as has been mentioned the integral has to be absolute convergent and there are some probability distributions for which there is no mean value. There are however other interesting location quantities. The **mode** is the value  $x_0$  of  $X$  for which the distribution is maximum; that is,

$$x_0 = \sup_{x \in \Omega} p(x) \quad (144)$$

Nevertheless, it may happen that there are several relative maximums so we talk about uni-modal, bi-modal,... distributions. The **median** is the value  $x_m$  such that

$$F(x_m) = P(X \leq x_m) = 1/2 \longrightarrow \int_{-\infty}^{x_m} p(x) dx = \int_{x_m}^{\infty} p(x) dx = P(X > x_m) = 1/2 \quad (145)$$

For discrete random quantities, the distribution function is either a finite or countable combination of indicator functions  $\mathbf{1}_{A_k}(x)$  with  $\{A_k\}_{k=1}^{n,\infty}$  a partition of  $\Omega$  so it may happen that  $F(x) = 1/2 \forall x \in A_k$ . Then, any value of the interval  $A_k$  can be considered the median. Last, we may consider the **quantiles**  $\alpha$  defined as the value  $q_\alpha$  of the random quantity such that  $F(q_\alpha) = P(X \leq q_\alpha) = \alpha$  so the *median* is the *quantile*  $q_{1/2}$ .

### 2.4.2.2 Dispersion parameters

There are many ways to give an idea of how *dispersed* are the values the random quantity may take. Usually they are based on the mathematical expectation of a function that depends on the difference between  $X$  and some characteristic value it may take; for instance  $E[|X - \mu|]$ . By far, the most usual and important one is the already defined **variance**

$$V[X] = \sigma^2 = E[(X - E[X])^2] = \int_{\mathcal{R}} (x - \mu)^2 p(x) dx \quad (146)$$

provided it exists. Note that if the random quantity  $X$  has dimension  $D[X] = d_X$ , the variance has dimension  $D[\sigma^2] = d_X^2$  so to have a quantity that gives an idea of the *dispersion* and has the same dimension one defines the **standard deviation**  $\sigma = +\sqrt{V[X]} = +\sqrt{\sigma^2}$  and, if both the mean value ( $\mu$ ) and the variance exist, the **standardized** random quantity

$$Y = \frac{X - \mu}{\sigma} \quad (147)$$

for which  $E[Y] = 0$  and  $V[Y] = \sigma_Y^2 = 1$ .

### 2.4.2.3 Asymmetry and Peakiness parameters

Related to higher order non-central moments, there are two dimensionless quantities of interest: the skewness and the kurtosis. The first non-trivial odd moment with respect to the mean is that of order 3:  $\mu_3$ . Since it has dimension  $D[\mu_3] = d_X^3$  we define the **skewness** ( $\gamma_1$ ) as the dimensionless quantity

$$\gamma_1 \stackrel{def}{=} \frac{\mu_3}{\mu_2^{3/2}} = \frac{\mu_3}{\sigma^3} = \frac{E[(X - \mu)^3]}{\sigma^3} \quad (148)$$

The skewness is  $\gamma_1 = 0$  for distributions that are symmetric with respect to the mean,  $\gamma_1 > 0$  if the probability content is more concentrated on the right of the mean and  $\gamma_1 < 0$  if it is to the left of the mean. Note however that there are many asymmetric distributions for which  $\mu_3 = 0$  and therefore  $\gamma_1 = 0$ . For unimodal distributions, it is easy to see that

$$\gamma_1 = 0 \quad \text{mode} = \text{median} = \text{mean}$$

$$\gamma_1 > 0 \quad \text{mode} < \text{median} < \text{mean}$$

$$\gamma_1 < 0 \quad \text{mode} > \text{median} > \text{mean}$$

The **kurtosis** is defined, again for dimensional considerations, as

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{\sigma^4} = \frac{E[(X - \mu)^4]}{\sigma^4} \quad (149)$$

and gives an idea of how *peaked* is the distribution. For the Normal distribution  $\gamma_2 = 3$  so in order to have a reference one defines the *extended kurtosis* as  $\gamma_2^{ext} = \gamma_2 - 3$ . Thus,  $\gamma_2^{ext} > 0$  ( $< 0$ ) indicates that the distribution is *more (less) peaked* than the Normal. Again,  $\gamma_2^{ext} = 0$  for the Normal density and for any other distribution for which  $\mu_4 = 3\sigma^4$ . Last you can check that  $\forall a, b \in \mathcal{R} E[(X - \mu - a)^2(X - \mu - b)^2] > 0$  so, for instance, defining  $u = a + b$ ,  $w = ab$  and taking derivatives,  $\gamma_2 \geq 1 + \gamma_1^2$ .

**Example 2.15:** Consider the discrete random quantity  $X \sim Pn(k|\lambda)$  with

$$P(X = k) \equiv Pn(k|\lambda) = e^{-\lambda} \frac{\lambda^k}{\Gamma(k+1)} \quad ; \quad \lambda \in \mathcal{R}^+ ; k = 0, 1, 2, \dots \quad (150)$$

The moments with respect to the origin are

$$\alpha_n(\lambda) \equiv E[X^n] = e^{-\lambda} \sum_{k=0}^{\infty} k^n \frac{\lambda^k}{k!} \quad (151)$$

If  $a_k$  denotes the  $k^{th}$  term of the sum, then

$$a_{k+1} = \frac{\lambda}{k+1} \left(1 + \frac{1}{k}\right)^n a_k \quad \longrightarrow \quad \lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| \rightarrow 0$$

so being the series absolute convergent all order moments exist. Taking the derivative of  $\alpha_n(\lambda)$  with respect to  $\lambda$  one gets the recurrence relation

$$\alpha_{n+1}(\lambda) = \lambda \left( \alpha_n(\lambda) + \frac{d\alpha_n(\lambda)}{d\lambda} \right) \quad ; \quad \alpha_0(\lambda) = 1 \quad (152)$$

so we can easily get

$$\alpha_0 = 1 ; \quad \alpha_1 = \lambda ; \quad \alpha_2 = \lambda(\lambda + 1) ; \quad \alpha_3 = \lambda(\lambda^2 + 3\lambda + 1) ; \quad \alpha_4 = \lambda(\lambda^3 + 6\lambda^2 + 7\lambda + 1) \quad (153)$$

and from them

$$\mu_0 = 1 ; \quad \mu_1 = 0 ; \quad \mu_2 = \lambda ; \quad \mu_3 = \lambda ; \quad \mu_4 = \lambda(3\lambda + 1) \quad (154)$$

Thus, for the Poisson distribution  $Po(n|\lambda)$  we have that:

$$E[X] = \lambda \quad ; \quad V[X] = \lambda \quad ; \quad \gamma_1 = \lambda^{-1/2} \quad ; \quad \gamma_2 = 3 + \lambda^{-1} \quad (155)$$

**Example 2.16:** Consider  $X \sim Ga(x|a, b)$  with:

$$p(x) = \frac{a^b}{\Gamma(b)} e^{-ax} x^{b-1} \mathbf{1}_{(0, \infty)}(x) \lambda \quad ; \quad a, b \in \mathcal{R}^+ \quad (156)$$

The moments with respect to the origin are

$$\alpha_n = E[X^n] = \frac{a^b}{\Gamma(b)} \int_0^\infty e^{-ax} x^{b+n-1} dx = \frac{\Gamma(b+n)}{\Gamma(b)} a^{-n} \quad (157)$$

being the integral absolute convergent. Thus we have:

$$\mu_n = \frac{1}{a^n \Gamma(b)} \sum_{k=0}^n \binom{n}{k} (-b)^{n-k} \Gamma(b+k) \quad (158)$$

and in consequence

$$E[X] = \frac{b}{a} \quad ; \quad V[X] = \frac{b}{a^2} \quad ; \quad \gamma_1 = \frac{2}{\sqrt{b}} \quad ; \quad \gamma_2^{ext.} = \frac{6}{b} \quad (159)$$

**Example 2.17:** For the Cauchy distribution  $X \sim Ca(x|1, 1)$ ,

$$p(x) = \frac{1}{\pi} \frac{1}{1+x^2} \mathbf{1}_{(-\infty, \infty)}(x) \quad (160)$$

we have that

$$\alpha_n = E[X^n] = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x^n}{1+x^2} dx \quad (161)$$

and clearly the integral diverges for  $n > 1$  so there are no moments but the trivial one  $\alpha_0$ . Even for  $n = 1$ , the integral

$$\int_{-\infty}^{\infty} \frac{|x|}{(1+x^2)} dx = 2 \int_0^{\infty} \frac{x}{(1+x^2)} dx = \lim_{a \rightarrow \infty} \ln(1+a^2) \quad (162)$$

is not absolute convergent so, in strict sense, there is no mean value. However, the mode and the median are  $x_0 = x_m = 0$ , the distribution is symmetric about  $x = 0$  and for  $n = 1$  there exists the Cauchy's Principal Value and is equal to 0. Had we introduced the Probability Distributions as a subset of Generalized Distributions, the Principal Value is an admissible distribution. It is left as an exercise to show that for:

- **Pareto:**  $X \sim Pa(x|\nu, x_m)$  with  $p(x|x_m, \nu) \propto x^{-(\nu+1)} \mathbf{1}_{[x_m, \infty)}(x)$  ;  $x_m, \nu \in \mathcal{R}^+$
- **Student:**  $X \sim St(x|\nu)$  with  $p(x|\nu) \propto (1+x^2/\nu)^{-(\nu+1)/2} \mathbf{1}_{(-\infty, \infty)}(x)$  ;  $\nu \in \mathcal{R}^+$

the moments  $\alpha_n = E[X^n]$  exist iff  $n < \nu$ .

Another distribution of interest in physics is the Landau Distribution that describes the energy lost by a particle when traversing a material under certain conditions. The probability density, given as the inverse Laplace Transform, is:

$$p(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{s \log s + xs} ds \quad (163)$$

with  $c \in \mathcal{R}^+$  and closing the contour on the left along a counterclockwise semicircle with a branch-cut along the negative real axis it has a real representation

$$p(x) = \frac{1}{\pi} \int_0^\infty e^{-(r \log r + xr)} \sin(\pi r) dr \quad (164)$$

The actual expression of the distribution of the energy loss is quite involved and some simplifying assumptions have been made; among other things, that the energy transfer in the collisions is unbounded (no kinematic constraint). But nothing is for free and the price to pay is that the Landau Distribution has no moments other than the trivial of order zero. This is why instead of mean and variance one talks about the *most probable energy loss* and the *full-width-half-maximum*.



## 2.4.2.4 Correlation Coefficient

The **covariance** between the random quantities  $X_i$  and  $X_j$  was defined as:

$$V[X_i, X_j] = V[X_j, X_i] = E[(X_i - \mu_i)(X_j - \mu_j)] = E[X_i X_j] - E[X_i]E[X_j] \quad (165)$$

If  $X_i$  and  $X_j$  are independent, then  $E[X_i X_j] = E[X_i]E[X_j]$  and  $V[X_i, X_j] = 0$ . Conversely, if  $V[X_i, X_j] \neq 0$  then  $E[X_i X_j] \neq E[X_i]E[X_j]$  and in consequence  $X_i$  and  $X_j$  are not statistically independent. Thus, the *covariance*  $V[X_i, X_j]$  serves to quantify, to some extent, the degree of *statistical dependence* between the random quantities  $X_i$  and  $X_j$ . Again, for dimensional considerations one defines the *correlation coefficient*

$$\rho_{ij} = \frac{V[X_i, X_j]}{\sqrt{V[X_i]V[X_j]}} = \frac{E[X_i X_j] - E[X_i]E[X_j]}{\sigma_i \sigma_j} \quad (166)$$

Since  $p(x_i, x_j)$  is a non-negative function we can write

$$V[X_i, X_j] = \int_{\mathcal{R}^2} \left\{ (x_i - \mu_i) \sqrt{p(x_i, x_j)} \right\} \left\{ (x_j - \mu_j) \sqrt{p(x_i, x_j)} \right\} dx_i dx_j \quad (167)$$

and from the Cauchy-Schwarz inequality:

$$-1 \leq \rho_{ij} \leq 1 \quad (168)$$

The extreme values  $(+1, -1)$  will be taken when  $E[X_i X_j] = E[X_i]E[X_j] \pm \sigma_i \sigma_j$  and  $\rho_{ij} = 0$  when  $E[X_i X_j] = E[X_i]E[X_j]$ . In particular, it is immediate to see that if there is a linear relation between both random quantities; that is,  $X_i = aX_j + b$ , then  $\rho_{ij} = \pm 1$ . Therefore, it is a **linear correlation coefficient**. Note however that:

- If  $X_i$  and  $X_j$  are linearly related,  $\rho_{ij} = \pm 1$ , but  $\rho_{ij} = \pm 1$  **does not** imply necessarily a linear relation;
- If  $X_i$  and  $X_j$  are statistically independent, then  $\rho_{ij} = 0$  but  $\rho_{ij} = 0$  **does not** imply necessarily statistical independence as the following example shows.

**Example 2.18:** Let  $X_1 \sim p(x_1)$  and define a random quantity  $X_2$  as

$$X_2 = g(X_1) = a + bX_1 + cX_1^2 \quad (169)$$

Obviously,  $X_1$  and  $X_2$  are not statistically independent for there is a clear parabolic relation. However

$$V[X_1, X_2] = E[X_1 X_2] - E[X_1]E[X_2] = b\sigma^2 + c(\alpha_3 - \mu^3 - \mu\sigma^2) \quad (170)$$

with  $\mu$ ,  $\sigma^2$  and  $\alpha_3$  respectively the mean, variance and moment of order 3 with respect to the origin of  $X_1$  and, if we take  $b = c\sigma^{-2}(\mu^3 + \mu\sigma^2 - \alpha_3)$  then  $V[X_1, X_2] = 0$  and so is the (linear) correlation coefficient.

**NOTE 3: Information as a measure of independence.** The *Mutual Information* (see [?]) serves also to quantify the degree of statistical dependence between random quantities. Consider for instance the two-dimensional random quantity  $\mathbf{X} = (X_1, X_2) \sim p(x_1, x_2)$ . Then:

$$I(X_1 : X_2) = \int_{\mathbf{X}} dx_1 dx_2 p(x_1, x_2) \ln \left( \frac{p(x_1, x_2)}{p(x_1)p(x_2)} \right) \quad (171)$$

and  $I(X_1 : X_2) \geq 0$  with equality iff  $p(x_1, x_2) = p(x_1)p(x_2)$ . Let's look as an example to the bi-variate normal distribution:  $N(\mathbf{x}|\mu, \Sigma)$ :

$$p(\mathbf{x}|\phi) = (2\pi)^{-1} |\det[\Sigma]|^{-1/2} \exp \left\{ -\frac{1}{2} ((\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu)) \right\} \quad (172)$$

with covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \quad \text{and} \quad \det[\Sigma] = \sigma_1^2\sigma_2^2(1 - \rho^2) \quad (173)$$

Since  $X_i \sim N(x_i | \mu_i, \sigma_i)$ ;  $i = 1, 2$  we have that:

$$I(X_1 : X_2) = \int_{\mathbf{x}} dx_1 dx_2 p(\mathbf{x} | \mu, \Sigma) \ln \left( \frac{p(\mathbf{x} | \mu, \Sigma)}{p(x_1 | \mu_1, \sigma_1) p(x_2 | \mu_2, \sigma_2)} \right) = -\frac{1}{2} \ln(1 - \rho^2) \quad (174)$$

Thus, if  $X_1$  and  $X_2$  are independent ( $\rho = 0$ ),  $I(X_1 : X_2) = 0$  and when  $\rho \rightarrow \pm 1$ ,  $I(X_1 : X_2) \rightarrow \infty$ .

### 2.4.3 The "Error Propagation Expression"

Consider a n-dimensional random quantity  $\mathbf{X} = (X_1, \dots, X_n)$  with  $E[X_i] = \mu_i$  and the random quantity  $Y = g(\mathbf{X})$  with  $g(x)$  an infinitely differentiable function. If we make a Taylor expansion of  $g(\mathbf{X})$  around  $E[\mathbf{X}] = \mu$  we have

$$Y = g(\mathbf{X}) = g(\mu) + \sum_{i=1}^n \left( \frac{\partial g(\mathbf{x})}{\partial x_i} \right)_{\mu} Z_i + \frac{1}{2!} \sum_{i=1}^n \sum_{j=1}^n \left( \frac{\partial^2 g(\mathbf{x})}{\partial x_i \partial x_j} \right)_{\mu} Z_i Z_j + R \quad (175)$$

where  $Z_i = X_i - \mu_i$ . Now,  $E[Z_i] = 0$  and  $E[Z_i Z_j] = V[X_i, X_j] = V_{ij}$  so

$$E[Y] = E[g(\mathbf{X})] = g(\mu) + \frac{1}{2!} \sum_{i=1}^n \sum_{j=1}^n \left( \frac{\partial^2 g(\mathbf{x})}{\partial x_i \partial x_j} \right)_{\mu} V_{ij} + \dots \quad (176)$$

and therefore

$$Y - E[Y] = \sum_{i=1}^n \left( \frac{\partial g(\mathbf{x})}{\partial x_i} \right)_{\mu} Z_i + \frac{1}{2!} \sum_{i=1}^n \sum_{j=1}^n \left( \frac{\partial^2 g(\mathbf{x})}{\partial x_i \partial x_j} \right)_{\mu} (Z_i Z_j - V_{ij}) + \dots \quad (177)$$

Neglecting all but the first term

$$V[Y] = E[(Y - E[Y])^2] = \sum_{i=1}^n \sum_{j=1}^n \left( \frac{\partial g(x)}{\partial x_i} \right)_{\mu} \left( \frac{\partial g(x)}{\partial x_j} \right)_{\mu} V[X_i, X_j] + \dots \quad (178)$$

This is the first order approximation to  $V[Y]$  and usually is reasonable but has to be used with care. On the one hand, we have assumed that higher order terms are negligible and this is not always the case so further terms in the expansion may have to be considered. Take for instance the simple case  $Y = X_1 X_2$  with  $X_1$  and  $X_2$  independent random quantities. The first order expansion gives  $V[Y] \simeq \mu_1^2 \sigma_2^2 + \mu_2^2 \sigma_1^2$  and including second order terms (there are no more)  $V[Y] = \mu_1^2 \sigma_2^2 + \mu_2^2 \sigma_1^2 + \sigma_1^2 \sigma_2^2$ ; the correct result. On the other hand, all this is obviously meaningless if the random quantity  $Y$  has no variance. This is for instance the case for  $Y = X_1 X_2^{-1}$  when  $X_{1,2}$  are Normal distributed.

## 2.5 Integral Transforms

The Integral Transforms of Fourier, Laplace and Mellin are a very useful tool to study the properties of the random quantities and their distribution functions. In particular, they will allow us to obtain the distribution of the sum, product and ratio of random quantities, the moments of the distributions and to study the convergence of a sequence  $\{F_k(x)\}_{k=1}^{\infty}$  of distribution functions to  $F(x)$ .

**2.5.1 The Fourier Transform**

Let  $f : \mathcal{R} \rightarrow \mathcal{C}$  be a complex and integrable function ( $f \in L_1(\mathcal{R})$ ). The *Fourier Transform*  $\mathcal{F}(t)$  with  $t \in \mathcal{R}$  of  $f(x)$  is defined as:

$$\mathcal{F}(t) = \int_{-\infty}^{\infty} f(x) e^{ixt} dx \tag{179}$$

The class of functions for which the Fourier Transform exists is certainly much wider than the probability density functions  $p(x) \in L_1(\mathcal{R})$  (normalized real functions of real argument) we are interested in for which the transform always exists. If  $X \sim p(x)$ , the Fourier Transform is nothing else but the mathematical expectation

$$\mathcal{F}(t) = E[e^{itX}] ; \quad t \in \mathcal{R} \tag{180}$$

and it is called *Characteristic Function*  $\Phi(t)$ . Thus, depending on the character of the random quantity  $X$ , we shall have:

- if  $X$  is **discrete**:  $\Phi(t) = \sum_{x_k} e^{itx_k} P(X = x_k)$
- if  $X$  is **continuous**:  $\Phi(t) = \int_{-\infty}^{+\infty} e^{itx} dP(x) = \int_{-\infty}^{+\infty} e^{itx} p(x) dx$

Attending to its definition, the Characteristic Function  $\Phi(t)$ , with  $t \in \mathcal{R}$ , is a complex function and has the following properties:

- 1)  $\Phi(0) = 1$ ;
- 2)  $\Phi(t)$  is bounded:  $|\Phi(t)| \leq 1$ ;
- 3)  $\Phi(t)$  has Schwarzian symmetry:  $\Phi(-t) = \overline{\Phi(t)}$ ;
- 4)  $\Phi(t)$  is uniformly continuous in  $\mathcal{R}$ .

The first three properties are obvious. For the fourth one, observe that for any  $\epsilon > 0$  there exists a  $\delta > 0$  such that  $|\Phi(t_1) - \Phi(t_2)| < \epsilon$  when  $|t_1 - t_2| < \delta$  with  $t_1$  and  $t_2$  arbitrary in  $\mathcal{R}$  because

$$|\Phi(t + \delta) - \Phi(t)| \leq \int_{-\infty}^{+\infty} |1 - e^{-i\delta x}| dP(x) = 2 \int_{-\infty}^{+\infty} |\sin \delta x/2| dP(x) \tag{181}$$

and this integral can be made arbitrarily small taking a sufficiently small  $\delta$ . These properties, that clearly hold also for a discrete random quantity, are **necessary but not sufficient** for a function  $\Phi(t)$  to be the Characteristic Function of a distribution  $P(x)$  (see example 2.19). Generalizing for a n-dimensional random quantity  $\mathbf{X} = (X_1, \dots, X_n)$ :

$$\Phi(t_1, \dots, t_n) = E[e^{i \mathbf{t} \mathbf{X}}] = E[e^{i(t_1 X_1 + \dots + t_n X_n)}] \tag{182}$$

so, for the discrete case:

$$\Phi(t_1, \dots, t_n) = \sum_{x_1} \dots \sum_{x_n} e^{i(t_1 x_1 + \dots + t_n x_n)} P(X_1 = x_1, \dots, X_n = x_n) \tag{183}$$

and for the continuous case:

$$\Phi(t_1, \dots, t_n) = \int_{-\infty}^{+\infty} dx_1 \dots \int_{-\infty}^{+\infty} dx_n e^{i(t_1 x_1 + \dots + t_n x_n)} p(x_1, \dots, x_n) \tag{184}$$

The n-dimensional Characteristic Function is such that:

- 1)  $\Phi(0, \dots, 0) = 1$
- 2)  $|\Phi(t_1, \dots, t_n)| \leq 1$
- 3)  $\Phi(-t_1, \dots, -t_n) = \overline{\Phi}(t_1, \dots, t_n)$

**NOTE 4: Laplace Transform.** For a function  $f(x) : \mathcal{R}^+ \rightarrow \mathcal{C}$  defined as  $f(x) = 0$  for  $x < 0$ , we may consider also the *Laplace Transform* defined as

$$L(s) = \int_0^{\infty} e^{-sx} f(x) dx \quad (185)$$

with  $s \in \mathcal{C}$  provided it exists. For a non-negative random quantity  $X \sim p(x)$  this is just the mathematical expectation  $E[e^{-sx}]$  and is named *Moment Generating Function* since the derivatives give the moments of the distribution (see 5.1.4). While the Fourier Transform exists for  $f(x) \in L_1(\mathcal{R})$ , the Laplace Transform exists if  $e^{-sx} f(x) \in L_1(\mathcal{R}^+)$  and thus, for a wider class of functions and although it is formally defined for functions with non-negative support, it may be possible to extend the limits of integration to the whole real line (*Bilateral Laplace Transform*). However, for the functions we shall be interested in (probability density functions), both Fourier and Laplace Transforms exist and usually there is no major advantage in using one or the other.

**Example 2.19:** There are several criteria (Bochner, Kintchine, Cramèr,...) specifying sufficient and necessary conditions for a function  $\Phi(t)$ , that satisfies the four aforementioned conditions, to be the Characteristic Function of a random quantity  $X \sim F(x)$ . However, it is easy to find simple functions like

$$g_1(t) = e^{-t^4} \quad \text{and} \quad g_2(t) = \frac{1}{1+t^4} \quad (186)$$

that satisfy the four stated conditions and that can not be Characteristic Functions associated to any distribution. Let's calculate the moments of order one with respect to the origin and the central one of order two. In both cases (see section 5.1.4) we have that:

$$\alpha_1 = \mu = E[X] = 0 \quad \text{and} \quad \mu_2 = \sigma^2 = E[(X - \mu)^2] = 0 \quad (187)$$

that is, the mean value and the variance are zero so the distribution function is zero almost everywhere but for  $X = 0$  where  $P(X = 0) = 1$ ... but this is the *Singular Distribution*  $\text{Sn}(x|0)$  that takes the value 1 if  $X = 0$  and 0 otherwise whose Characteristic Function is  $\Phi(t) = 1$ . In general, any function  $\Phi(t)$  that in a boundary of  $t = 0$  behaves as  $\Phi(t) = 1 + O(t^{2+\epsilon})$  with  $\epsilon > 0$  can not be the Characteristic Function associated to a distribution  $F(x)$  unless  $\Phi(t) = 1$  for all  $t \in \mathcal{R}$ .

**Example 2.20:** The elements of the Cantor Set  $C_S(0, 1)$  can be represented in base 3 as:

$$X = \sum_{n=1}^{\infty} \frac{X_n}{3^n} \quad (188)$$

with  $X_n \in \{0, 2\}$ . This set is non-denumerable and has zero Lebesgue measure so any distribution with support on it is *singular* and, in consequence, has no pdf. The Uniform Distribution on  $C_S(0, 1)$  is defined assigning a probability  $P(X_n = 0) = P(X_n = 2) = 1/2$  (Geometric Distribution). Then, for the random quantity  $X_n$  we have that

$$\Phi_{X_n}(t) = E[e^{itX}] = \frac{1}{2} (1 + e^{2it}) \quad (189)$$

and for  $Y_n = X_n/3^n$ :

$$\Phi_{Y_n}(t) = \Phi_{X_n}(t/3^n) = \frac{1}{2} (1 + e^{2it/3^n}) \quad (190)$$

Being all  $X_n$  statistically independent, we have that

$$\Phi_X(t) = \prod_{n=1}^{\infty} \frac{1}{2} \left( 1 + e^{2it/3^n} \right) = \prod_{n=1}^{\infty} \frac{1}{2} e^{it/3^n} \cos(t/3^n) = e^{it/2} \prod_{n=1}^{\infty} \cos(t/3^n) \quad (191)$$

and, from the derivatives (section 2.5.1.4) it is straight forward to calculate the moments of the distribution. In particular:

$$\Phi_X^1(0) = \frac{i}{2} \longrightarrow E[X] = \frac{1}{2} \quad \text{and} \quad \Phi_X^2(0) = -\frac{3}{8} \longrightarrow E[X^2] = \frac{3}{8} \quad (192)$$

so  $V[X] = 1/8$ .

### 2.5.1.1 Inversion Theorem (Lévy, 1925)

The Inverse Fourier Transform allows us to obtain the distribution function of a random quantity from the Characteristic Function. If  $X$  is a continuous random quantity and  $\Phi(t)$  its Characteristic Function, then the pdf  $p(x)$  will be given by

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \Phi(t) dt \quad (193)$$

provided that  $p(x)$  is continuous at  $x$  and, if  $X$  is discrete:

$$P(X = x_k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx_k} \Phi(t) dt \quad (194)$$

In particular, if the discrete distribution is *reticular* (that is, all the possible values that the random quantity  $X$  may take can be expressed as  $a + bn$  with  $a, b \in \mathcal{R}$ ;  $b \neq 0$  and  $n$  integer) we have that:

$$P(X = x_k) = \frac{b}{2\pi} \int_{-\pi/b}^{\pi/b} e^{-itx_k} \Phi(t) dt \quad (195)$$

From this expressions, we can obtain also the relation between the Characteristic Function and the Distribution Function. For discrete random quantities we shall have:

$$F(x_k) = \sum_{x \leq x_k} P(X = x_k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \sum_{x \leq x_k} e^{-itx} \Phi(t) dt \quad (196)$$

and, in the continuous case, for  $x_1 < x_2 \in \mathcal{R}$  we have that:

$$F(x_2) - F(x_1) = \int_{x_1}^{x_2} p(x) dx = \frac{1}{2\pi i} \int_{-\infty}^{+\infty} \Phi(t) \frac{1}{t} (e^{-itx_1} - e^{-itx_2}) dt \quad (197)$$

so taking  $x_1 = 0$  we have that (Lévy, 1925):

$$F(x) = F(0) + \frac{1}{2\pi i} \int_{-\infty}^{+\infty} \Phi(t) \frac{1}{t} (1 - e^{-itx}) dt \quad (198)$$

The **Inversion Theorem** states that there is a one-to-one correspondence between a distribution function and its Characteristic Function so to each Characteristic Function corresponds **one and only one** distribution function that can be either discrete or continuous but not a combination of both. Therefore, two distribution functions with the same Characteristic Function may differ, at most, on their points of discontinuity that, as we have seen, are a set of zero measure. In consequence, if we have two random quantities  $X_1$  and  $X_2$  with distribution functions  $P_1(x)$  and  $P_2(x)$ , a **necessary** and **sufficient** condition for  $P_1(x) = P_2(x)$  a.e. is that  $\Phi_1(t) = \Phi_2(t)$  for all  $t \in \mathcal{R}$ .

### 2.5.1.2 Changes of variable

Let  $X \sim P(x)$  be a random quantity with Characteristic Function  $\Phi_X(t)$  and  $g(X)$  a one-to-one finite real function defined for all real values of  $X$ . The Characteristic Function of the random quantity  $Y = g(X)$  will be given by:

$$\Phi_Y(t) = E_Y[e^{itY}] = E_X[e^{itg(X)}] \quad (199)$$

that is:

$$\Phi_Y(t) = \int_{-\infty}^{+\infty} e^{itg(x)} dP(x) \quad \text{or} \quad \Phi_Y(t) = \sum_{x_k} e^{itg(x_k)} P(X = x_k) \quad (200)$$

depending on whether  $X$  is continuous or discrete. In the particular case of a linear transformation  $Y = aX + b$  with  $a$  and  $b$  real constants, we have that:

$$\Phi_Y(t) = E_X[e^{it(aX + b)}] = e^{itb} \Phi_X(at) \quad (201)$$

### 2.5.1.3 Sum of random quantities

The Characteristic Function is particularly useful to obtain the Distribution Function of a random quantity defined as the sum of independent random quantities. If  $X_1, \dots, X_n$  are  $n$  independent random quantities with Characteristic Functions  $\Phi_1(t_1), \dots, \Phi_n(t_n)$ , then the Characteristic Function of  $X = X_1 + \dots + X_n$  will be:

$$\Phi_X(t) = E[e^{itX}] = E[e^{it(X_1 + \dots + X_n)}] = \Phi_1(t) \dots \Phi_n(t) \quad (202)$$

that is, the product of the Characteristic Functions of each one, a **necessary** but **not sufficient** condition for the random quantities  $X_1, \dots, X_n$  to be independent. In a similar way, we have that if  $X = X_1 - X_2$  with  $X_1$  and  $X_2$  independent random quantities, then

$$\Phi_X(t) = E[e^{it(X_1 - X_2)}] = \Phi_1(t) \Phi_2(-t) = \Phi_1(t) \bar{\Phi}_2(t) \quad (203)$$

Form these considerations, it is left as an exercise to show that:

- **Poisson:** The sum of  $n$  independent random quantities, each distributed as  $Po(x_k | \mu_k)$  with  $k = 1, \dots, n$  is Poisson distributed with parameter  $\mu_s = \mu_1 + \dots + \mu_n$ .
- **Normal:** The sum of  $n$  independent random quantities, each distributed as  $N(x_k | \mu_k, \sigma_k)$  with  $k = 1, \dots, n$  is Normal distributed with mean  $\mu_s = \mu_1 + \dots + \mu_n$  and variance  $\sigma_s^2 = \sigma_1^2 + \dots + \sigma_n^2$ .
- **Cauchy:** The sum of  $n$  independent random quantities, each Cauchy distributed  $Ca(x_k | \alpha_k, \beta_k)$  with  $k = 1, \dots, n$  is Cauchy distributed with parameters  $\alpha_s = \alpha_1 + \dots + \alpha_n$  and  $\beta_s = \beta_1 + \dots + \beta_n$ .
- **Gamma:** The sum of  $n$  independent random quantities, each distributed as  $Ga(x_k | \alpha, \beta_k)$  with  $k = 1, \dots, n$  is Gamma distributed with parameters  $(\alpha, \beta_1 + \dots + \beta_n)$ .

#### Example 2.21: Difference of Poisson distributed random quantities.

Consider two independent random quantities  $X_1 \sim Po(X_1 | \mu_1)$  and  $X_2 \sim Po(X_2 | \mu_2)$  and let us find the distribution of  $X = X_1 - X_2$ . Since for the Poisson distribution:

$$X_i \sim Po(\mu_i) \quad \longrightarrow \quad \Phi_i(t) = e^{-\mu_i} (1 - e^{it}) \quad (204)$$

we have that

$$\Phi_X(t) = \Phi_1(t) \bar{\Phi}_2(t) = e^{-(\mu_1 + \mu_2)} e^{(\mu_1 e^{it} + \mu_2 e^{-it})} \quad (205)$$

Obviously,  $X$  is a discrete random quantity with integer support  $\Omega_X = \{\dots, -2, -1, 0, 1, 2, \dots\}$ ; that is, a *reticular* random quantity with  $a = 0$  and  $b = 1$ . Then

$$P(X = n) = \frac{1}{2\pi} e^{-\mu_S} \int_{-\pi}^{\pi} e^{-itn} e^{(\mu_1 e^{it} + \mu_2 e^{-it})} dt \quad (206)$$

being  $\mu_S = \mu_1 + \mu_2$ . If we take:

$$z = \sqrt{\frac{\mu_1}{\mu_2}} e^{it} \quad (207)$$

we have

$$P(X = n) = \left(\frac{\mu_1}{\mu_2}\right)^{n/2} e^{-\mu_S} \frac{1}{2\pi i} \oint_C z^{-n-1} e^{\frac{w}{2}(z + 1/z)} dz \quad (208)$$

with  $w = 2\sqrt{\mu_1\mu_2}$  and  $C$  the circle  $|z| = \sqrt{\mu_1/\mu_2}$  around the origin. From the definition of the Modified Bessel Function of first kind

$$I_n(z) = \frac{1}{2\pi i} \oint_C t^{-n-1} e^{\frac{z}{2}(t + 1/t)} dt \quad (209)$$

with  $C$  a circle enclosing the origin anticlockwise and considering that  $I_{-n}(z) = I_n(z)$  we have finally:

$$P(X = n) = \left(\frac{\mu_1}{\mu_2}\right)^{n/2} e^{-(\mu_1 + \mu_2)} I_{|n|}(2\sqrt{\mu_1\mu_2}) \quad (210)$$

#### 2.5.1.4 Moments of a distribution

Consider a continuous random quantity  $X \sim P(x)$  and Characteristic Function

$$\Phi(t) = E[e^{itX}] = \int_{-\infty}^{+\infty} e^{itx} dP(x) \quad (211)$$

and let us assume that there exists the moment of order  $k$ . Then, upon derivation of the Characteristic Function  $k$  times with respect to  $t$  we have:

$$\frac{\partial^k}{\partial t^k} \Phi(t) = E[i^k X^k e^{itX}] = \int_{-\infty}^{+\infty} (ix)^k e^{itx} dP(x) \quad (212)$$

and taking  $t = 0$  we get the *moments with respect to the origin*:

$$E[X^k] = \frac{1}{i^k} \left( \frac{\partial^k}{\partial t^k} \Phi(t) \right)_{t=0} \quad (213)$$

Consider now the Characteristic Function referred to an arbitrary point  $a \in \mathcal{R}$ ; that is:

$$\Phi(t, a) = E[e^{it(X - a)}] = \int_{-\infty}^{+\infty} e^{it(x - a)} dP(x) = e^{-ita} \Phi(t) \quad (214)$$

In a similar way, upon  $k$  times derivation with respect to  $t$  we get the moments with respect to an arbitrary point  $a$ :

$$E[(X - a)^k] = \frac{1}{i^k} \left( \frac{\partial^k}{\partial t^k} \Phi(t, a) \right)_{t=0} \quad (215)$$

and the *central moments* if  $a = E(X) = \mu$ . The extension to  $n$  dimensions immediate: for a  $n$  dimensional random quantity  $\mathbf{X}$  we shall have the for the moment  $\alpha_{k_1 \dots k_n}$  with respect to the origin that

$$\alpha_{k_1 \dots k_n} = E[X_1^{k_1} \dots X_n^{k_n}] = \frac{1}{i^{k_1 + \dots + k_n}} \left( \frac{\partial^{k_1 + \dots + k_n}}{\partial^{k_1} t_1 \dots \partial^{k_n} t_n} \Phi(t_1, \dots, t_n) \right)_{t_1 = \dots = t_n = 0} \quad (216)$$

**Example 2.22:** For the difference of Poisson distributed random quantities analyzed in the previous example, one can easily derive the moments from the derivatives of the Characteristic Function. Since

$$\log \Phi_X(t) = -(\mu_1 + \mu_2) + (\mu_1 e^{it} + \mu_2 e^{-it}) \quad (217)$$

we have that

$$\Phi'_X(0) = i(\mu_1 - \mu_2) \quad \longrightarrow \quad E[X] = \mu_1 - \mu_2 \quad (218)$$

$$\Phi''_X(0) = (\Phi'_X(0))^2 - (\mu_1 + \mu_2) \quad \longrightarrow \quad V[X] = \mu_1 + \mu_2 \quad (219)$$

**Problem 2.3:** The Moyal Distribution, with density

$$p(x) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} (x + e^{-x}) \right\} \mathbf{1}_{(-\infty, \infty)}(x) \quad (220)$$

is sometimes used as an approximation to the Landau Distribution. Obtain the Characteristic Function  $\Phi(t) = \pi^{-1/2} 2^{-it} \Gamma(1/2 - it)$  and show that  $E[x] = \gamma_E + \ln 2$  and  $V[X] = \pi^2/2$ .

### 2.5.2 The Mellin Transform

Let  $f : \mathcal{R}^+ \rightarrow \mathcal{C}$  be a complex and integrable function with support on the real positive axis. The *Mellin Transform* is defined as:

$$M(f; s) = M_f(s) = \int_0^\infty f(x) x^{s-1} dx \quad (221)$$

with  $s \in \mathcal{C}$ , provided that the integral exists. In general, we shall be interested in continuous probability density functions  $f(x)$  such that

$$\lim_{x \rightarrow 0^+} f(x) = O(x^\alpha) \quad \text{and} \quad \lim_{x \rightarrow \infty} f(x) = O(x^\beta) \quad (222)$$

and therefore

$$\begin{aligned} |M(f; s)| &\leq \int_0^\infty |f(x)| x^{Re(s)-1} dx = \int_0^1 |f(x)| x^{Re(s)-1} dx + \int_1^\infty |f(x)| x^{Re(s)-1} dx \leq \\ &\leq C_1 \int_0^1 x^{Re(s)-1+\alpha} dx + C_2 \int_1^\infty x^{Re(s)-1+\beta} dx \end{aligned} \quad (223)$$

The first integral converges for  $-\alpha < Re(s)$  and the second for  $Re(s) < -\beta$  so the Mellin Transform exists and is holomorphic on the band  $-\alpha < Re(s) < -\beta$ , parallel to the imaginary axis  $\Im(s)$  and determined by the conditions of convergence of the integral. We shall denote the holomorphy band (that can be a half of the complex plane or the whole complex plane) by  $S_f = \langle -\alpha, -\beta \rangle$ . Last, to simplify the notation when dealing with several random quantities, we shall write for  $X_n \sim p_n(x)$   $M_n(s)$  of  $M_X(s)$  instead of  $M(p_n; s)$ .



2.5.2.1 *Inversion*

For a given function  $f(t)$  we have that

$$M(f; s) = \int_0^\infty f(t) t^{s-1} dt = \int_0^\infty f(t) e^{(s-1)\ln t} dt = \int_{-\infty}^\infty f(e^u) e^{su} du \quad (224)$$

assuming that the integral exists. Since  $s \in \mathcal{C}$ , we can write  $s = x + iy$  so: the Mellin Transform of  $f(t)$  is the Fourier Transform of  $g(u) = f(e^u)e^{xu}$ . Setting now  $t = e^u$  we have that

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^\infty M(f; s = x + iy) t^{-(x+iy)} dy = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} M(f; s) t^{-s} ds \quad (225)$$

where, due to Cauchy's Theorem,  $\sigma$  lies anywhere within the holomorphy band. The uniqueness of the result holds with respect to this strip so, in fact, the Mellin Transform consists on the pair  $M(s)$  together with the band  $\langle a, b \rangle$ .

**Example 2.23:** It is clear that to determine the function  $f(x)$  from the transform  $F(s)$  we have to specify the strip of analyticity for, otherwise, we do not know which poles should be included. Let's see as an example  $f_1(x) = e^{-x}$ . We have that

$$M_1(z) = \int_0^\infty e^{-x} x^{z-1} dx = \Gamma(z) \quad (226)$$

holomorphic in the band  $\langle 0, \infty \rangle$  so, for the inverse transform, we shall include the poles  $z = 0, -1, -2, \dots$ . For  $f_2(x) = e^{-x} - 1$  we get  $M_2(s) = \Gamma(s)$ , the same function, but

$$\lim_{x \rightarrow 0^+} f(x) \simeq O(x^1) \rightarrow \alpha = 1 \quad \text{and} \quad \lim_{x \rightarrow \infty} f(x) \simeq O(x^0) \rightarrow \beta = 0 \quad (227)$$

Thus, the holomorphy strip is  $\langle -1, 0 \rangle$  and for the inverse transform we shall include the poles  $z = -1, -2, \dots$ . For  $f_3(x) = e^{-x} - 1 + x$  we get  $M_3(s) = \Gamma(s)$ , again the same function, but

$$\lim_{x \rightarrow 0^+} f(x) \simeq O(x^2) \rightarrow \alpha = 2 \quad \text{and} \quad \lim_{x \rightarrow \infty} f(x) \simeq O(x^1) \rightarrow \beta = 1 \quad (228)$$

Thus, the holomorphy strip is  $\langle -2, -1 \rangle$  and for the inverse transform we include the poles  $z = -2, -3, \dots$

2.5.2.2 *Useful properties*

Consider a positive random quantity  $X$  with continuous density  $p(x)$  and  $x \in [0, \infty)$ , the Mellin Transform  $M_X(s)$  (defined only for  $x \geq 0$ )

$$M(p; s) = \int_0^\infty x^{s-1} p(x) dx = E[X^{s-1}] \quad (229)$$

and the Inverse Transform

$$p(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} x^{-s} M(p; s) ds \quad (230)$$

defined for all  $x$  where  $p(x)$  is continuous with the line of integration contained in the strip of analyticity of  $M(p; s)$ . Then:

- **Moments:**  $E[X^n] = M_X(n + 1)$ ;

- For the positive random quantity  $Z = aX^b$  ( $a, b \in \mathcal{R}$  and  $a > 0$ ) we have that

$$M_Z(s) = \int_0^\infty z^{s-1} f(z) dz = \int_0^\infty a^{s-1} x^{b(s-1)} p(x) dx = a^{s-1} M_X(bs - b + 1) \quad (231)$$

$$2\pi i p(z) = \int_{c-i\infty}^{c+i\infty} z^{-s} M_X(bs - b + 1) ds \quad (232)$$

In particular, for  $Z = 1/X$  ( $a = 1$  and  $b = -1$ ) we have that

$$M_{Z=1/X}(s) = M_X(2 - s) \quad (233)$$

• If  $Z = X_1 X_2 \cdots X_n$  with  $\{X_i\}_{i=1}^n$   $n$  independent positive defined random quantities, each distributed as  $p_i(x_i)$ , we have that

$$M_Z(s) = \int_0^\infty z^{s-1} p(z) dz = \prod_{i=1}^n \int_0^\infty x_i^{s-1} p_i(x_i) dx_i = \prod_{i=1}^n E[X_i^{s-1}] = \prod_{i=1}^n M_i(s) \quad (234)$$

$$2\pi i p(z) = \int_{c-i\infty}^{c+i\infty} z^{-s} M_1(s) \cdots M_n(s) ds \quad (235)$$

In particular, for  $n = 2$ ,  $X = X_1 X_2$  it is easy to check that

$$p(x) = \int_0^\infty p_1(w) p_2(x/w) dw/w = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} x^{-s} M_1(s) M_2(s) ds \quad (236)$$

Obviously, the strip of holomorphy is  $S_1 \cap S_2$ .

• For  $X = X_1/X_2$ , with both  $X_1$  and  $X_2$  positive defined and independent, we have that

$$M_X(s) = M_1(s) M_2(2 - s) \quad (237)$$

and therefore

$$p(x) = \int_0^\infty p_1(wx) p_2(w) w dw = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} x^{-s} M_1(s) M_2(2 - s) ds \quad (238)$$

• Consider the distribution function  $F(x) = \int_0^x p(u) du$  of the random quantity  $X$ . Since  $dF(x) = p(x) dx$  we have that

$$M(p(x); s) = \int_0^\infty x^{s-1} dF(x) = [x^{s-1} F(x)]_0^\infty - (s-1) \int_0^\infty x^{s-2} F(x) dx \quad (239)$$

and therefore, if  $\lim_{x \rightarrow 0^+} [x^{s-1} F(x)] = 0$  and  $\lim_{x \rightarrow \infty} [x^{s-1} F(x)] = 0$  we have, shifting  $s \rightarrow s - 1$ , that

$$M(F(x); s) = M\left(\int_0^x p(u) du; s\right) = -\frac{1}{s} M(p(x); s + 1) \quad (240)$$

### 2.5.2.3 Some useful examples

• **Ratio and product of two independent Exponential distributed random quantities**

Consider  $X_1 \sim Ex(x_1|a_1)$  and  $X_2 \sim Ex(x_2|a_2)$ . The Mellin transform of  $X \sim Ex(x|a)$  is

$$M_X(s) = \int_0^\infty x^{s-1} p(x|a) dx = a \int_0^\infty x^{s-1} e^{-ax} dx = \frac{\Gamma(s)}{a^{s-1}} \quad (241)$$

and therefore, for  $Z = 1/X$ :

$$M_Z(s) = M_X(2-s) = \frac{\Gamma(2-s)}{a^{1-s}} \quad (242)$$

In consequence, we have that

$$\begin{aligned} \bullet X = X_1 X_2 \quad \longrightarrow \quad M_X(z) &= M_1(z) M_2(z) = \frac{\Gamma(z)^2}{(a_1 a_2)^{z-1}} \\ p(x) &= \frac{a_1 a_2}{2\pi i} \int_{c-i\infty}^{c+i\infty} (a_1 a_2 x)^{-z} \Gamma(z)^2 dz \end{aligned} \quad (243)$$

The poles of the integrand are at  $z_n = -n$  and the residuals<sup>13</sup> are

$$Res(f(z), z_n) = \frac{(a_1 a_2 x)^n}{(n!)^2} (2\psi(n+1) - \ln(a_1 a_2 x)) \quad (244)$$

and therefore

$$p(x) = a_1 a_2 \sum_{n=0}^{\infty} \frac{(a_1 a_2 x)^n}{(n!)^2} (2\psi(n+1) - \ln(a_1 a_2 x)) \quad (245)$$

If we define  $w = 2\sqrt{a_1 a_2 x}$

$$p(x) = 2 a_1 a_2 K_0(2\sqrt{a_1 a_2 x}) \mathbf{1}_{(0,\infty)}(x) \quad (246)$$

from the Neumann Series expansion the Modified Bessel Function  $K_0(w)$ .

$$\begin{aligned} \bullet Y = X_1 X_2^{-1} \quad \longrightarrow \quad M_Y(z) &= M_1(z) M_2(2-z) = \left(\frac{a_2}{a_1}\right)^{z-1} \frac{\pi(1-z)}{\sin(z\pi)} \\ p(x) &= \frac{a_1 a_2^{-1}}{2i} \int_{c-i\infty}^{c+i\infty} (a_1 a_2^{-1} x)^{-z} \frac{1-z}{\sin(z\pi)} dz \end{aligned} \quad (247)$$

Considering again the poles of  $M_Y(z)$  at  $z_n = -n$  we get the residuals

$$Res(f(z), z_n) = (1+n) (-1)^n \left(\frac{a_1}{a_2}\right)^{n+1} x^n \quad (248)$$

and therefore:

$$p(x) = \frac{a_1}{a_2} \sum_{n=0}^{\infty} (1+n) (-1)^n \left(\frac{a_1 x}{a_2}\right)^n = \frac{a_1 a_2}{(a_2 + a_1 x)^2} \mathbf{1}_{(1,\infty)}(x) \quad (249)$$

To summarize, if  $X_1 \sim Ex(x_1|a_1)$  and  $X_2 \sim Ex(x_2|a_2)$  are independent random quantities:

$$\begin{aligned} X = X_1 X_2 &\sim 2 a_1 a_2 K_0(2\sqrt{a_1 a_2 x}) \mathbf{1}_{(1,\infty)}(x) \\ Y = X_1/X_2 &\sim \frac{a_1 a_2}{(a_2 + a_1 x)^2} \mathbf{1}_{(0,\infty)}(x) \end{aligned} \quad (250)$$

### • Ratio and product of two independent Gamma distributed random quantities

<sup>13</sup>In the following examples,  $-\pi \leq \arg(z) < \pi$ .

Consider  $Y \sim Ga(x|a, b)$ . Then  $X = aY \sim Ga(x|1, b)$  with Mellin Transform

$$M_X(s) = \frac{\Gamma(b + s - 1)}{\Gamma(b)} \quad (251)$$

The, if  $X_1 \sim Ga(x_1|1, b_1)$  and  $X_2 \sim Ga(x_2|1, b_2)$ ;  $b_1 \neq b_2$ :

- $X = X_1 X_2^{-1} \rightarrow M_X(z) = M_1(z) M_2(z) = \frac{\Gamma(b_1 - 1 + z)}{\Gamma(b_1)} \frac{\Gamma(b_2 + 1 - z)}{\Gamma(b_2)}$

$$2\pi i \Gamma(b_1) \Gamma(b_2) p(x) = \int_{c-i\infty}^{c+i\infty} x^{-z} \Gamma(b_1 - 1 + z) \Gamma(b_2 + 1 - z) dz \quad (252)$$

Closing the contour on the left of the line  $Re(z) = c$  contained in the strip of holomorphy  $< 0, \infty >$  we have poles of order one at  $b_i - 1 + z_n = -n$  with  $n = 0, 1, 2, \dots$ , that is, at  $z_n = 1 - b_i - n$ . Expansion around  $z = z_n + \epsilon$  gives the residuals

$$Res(f(z), z_n) = \frac{(-1)^n}{n!} \Gamma(b_1 + b_2 + n) x^{n+b_1-1} \quad (253)$$

and therefore the quantity  $X = X_1/X_2$  is distributed as

$$p(x) = \frac{x^{b_1-1}}{\Gamma(b_1) \Gamma(b_2)} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \Gamma(b_1 + b_2 + n) x^n = \frac{\Gamma(b_1 + b_2)}{\Gamma(b_1) \Gamma(b_2)} \frac{x^{b_1-1}}{(1+x)^{b_1+b_2}} \mathbf{1}_{(0,\infty)}(x) \quad (254)$$

- $X = X_1 X_2 \rightarrow M_X(z) = M_1(z) M_2(z) = \frac{\Gamma(b_1 - 1 + z)}{\Gamma(b_1)} \frac{\Gamma(b_2 - 1 + z)}{\Gamma(b_2)}$

Without loss of generality, we may assume that  $b_2 > b_1$  so the strip of holomorphy is  $< 1 - b_1, \infty >$ . Then, with  $c > 1 - b_1$  real

$$\Gamma(b_1) \Gamma(b_2) p(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} x^{-z} \Gamma(b_1 - 1 + z) \Gamma(b_2 - 1 + z) dz \quad (255)$$

Considering the definition of the Modified Bessel Functions

$$I_\nu(x) = \sum_{n=0}^{\infty} \frac{1}{n! \Gamma(1 + n + \nu)} \left(\frac{x}{2}\right)^{2n+\nu} \quad \text{and} \quad K_\nu(x) = \frac{\pi}{2} \frac{I_{-\nu}(x) - I_\nu(x)}{\sin(\nu\pi)} \quad (256)$$

we get that

$$p(x) = \frac{2}{\Gamma(b_1) \Gamma(b_2)} x^{(b_1+b_2)/2-1} K_\nu(2\sqrt{x}) \mathbf{1}_{(0,\infty)}(x) \quad (257)$$

with  $\nu = b_2 - b_1 > 0$ .

To summarize, if  $X_1 \sim Ga(x_1|a_1, b_1)$  and  $X_2 \sim Ga(x_2|a_2, b_2)$  are two independent random quantities and  $\nu = b_2 - b_1 > 0$  we have that

$$\begin{aligned} X = X_1 X_2 &\sim \frac{2a_1^{b_1} a_2^{b_2}}{\Gamma(b_1) \Gamma(b_2)} \left(\frac{a_2}{a_1}\right)^{\nu/2} x^{(b_1+b_2)/2-1} K_\nu(2\sqrt{a_1 a_2 x}) \mathbf{1}_{(0,\infty)}(x) \\ X = X_1/X_2 &\sim \frac{\Gamma(b_1 + b_2)}{\Gamma(b_1) \Gamma(b_2)} \frac{a_1^{b_1} a_2^{b_2} x^{b_1-1}}{(a_2 + a_1 x)^{b_1+b_2}} \mathbf{1}_{(0,\infty)}(x) \end{aligned} \quad (258)$$

• **Ratio and product of two independent Uniform distributed random quantities**

Consider  $X \sim Un(x|0, 1)$ . Then  $M_X(z) = 1/z$  with  $S = \langle 0, \infty \rangle$ . For  $X = X_1 \cdots X_n$  we have

$$p(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{-z \ln x} z^{-n} dz = \frac{(-\ln x)^{n-1}}{\Gamma(n)} \mathbf{1}_{(0,1]}(x) \quad (259)$$

being  $z = 0$  the only pole or order  $n$ .

For  $X = X_1/X_2$  one has to be careful when defining the contours. In principle,

$$M_X(s) = M_1(s)M_2(2-s) = \frac{1}{s} \frac{1}{2-s} \quad (260)$$

so the strip of holomorphy is  $S = \langle 0, 2 \rangle$  and there are two poles, at  $s = 0$  and  $s = 2$ . If  $\ln x < 0 \rightarrow x < 1$  we shall close the Bromwich the contour on the left enclosing the pole at  $s = 0$  and if  $\ln x > 0 \rightarrow x > 1$  we shall close the contour on the right enclosing the pole at  $s = 2$  so the integrals converge. Then it is easy to get that

$$p(x) = \frac{1}{2} [\mathbf{1}_{(0,1]}(x) + x^{-2} \mathbf{1}_{(1,\infty)}(x)] = Un(x|0, 1) + Pa(x|1, 1) \quad (261)$$

Note that

$$E[X^n] = M_X(n+1) = \frac{1}{n+1} \frac{1}{1-n} \quad (262)$$

and therefore there are no moments for  $n \geq 1$ .

**Example 2.24:** Show that if  $X_i \sim Be(x_i|a_i, b_i)$  with  $a_i, b_i > 0$ , then

$$M_i(s) = \frac{\Gamma(a_i + b_i)}{\Gamma(a_i)} \frac{\Gamma(s + a_i - 1)}{\Gamma(s + a_i + b_i - 1)} \quad (263)$$

with  $S = \langle 1 - a_i, \infty \rangle$  and therefore:

•  $X = X_1 X_2$

$$p(x) = N_p x^{a_1-1} (1-x)^{b_1+b_2-1} F(a_1 - a_2 + b_1, b_2, b_1 + b_2, 1-x) \mathbf{1}_{(0,1)}(x) \quad (264)$$

with

$$N_p = \frac{\Gamma(a_1 + b_1)\Gamma(a_2 + b_2)}{\Gamma(a_1)\Gamma(a_2)\Gamma(b_1 + b_2)} \quad (265)$$

•  $X = X_1/X_2$

$$p(x) = N_1 x^{-(a_2+1)} F(1 - b_2, a_1 + a_2, b_1 + a_1 + a_2, x^{-1}) \mathbf{1}_{(1,\infty)}(x) + \quad (266)$$

$$+ N_2 x^{a_1-1} F(1 - b_1, a_1 + a_2, b_2 + a_1 + a_2, x) \mathbf{1}_{(0,1)}(x) \quad (267)$$

with

$$N_k = \frac{B(a_1 + a_2, b_k)}{B(a_1 + b_1)B(a_2 + b_2)} \quad (268)$$

**Example 2.25:** Consider a random quantity

$$X \sim p(x|a, b) = \frac{2 a^{(b+1)/2}}{\Gamma(b/2 + 1/2)} e^{-ax^2} x^b \quad (269)$$

with  $a, b > 0$  and  $x \in [0, \infty)$ . Show that

$$M(s) \sim p(x|a, b) = \frac{\Gamma(b/2 + s/2)}{\Gamma(b/2 + 1/2)} a^{-(s-1)/2} \quad (270)$$

with  $S = \langle -b, \infty \rangle$  and, from this, derive that the probability density function of  $X = X_1 X_2$ , with  $X_1 \sim p(x_1|a_1, b_1)$  and  $X_2 \sim p(x_2|a_2, b_2)$  independent, is given by:

$$p(x) = \frac{4\sqrt{a_1 a_2}}{\Gamma(b_1/2 + 1/2)\Gamma(b_2/2 + 1/2)} (\sqrt{a_1 a_2} x)^{(b_1+b_2)/2} K_{|\nu|} (2\sqrt{a_1 a_2} x) \quad (271)$$

with  $\nu = (b_2 - b_1)/2$  and for  $X = X_1/X_2$  by

$$p(x) = \frac{2\Gamma(b+1)}{\Gamma(b_1/2 + 1/2)\Gamma(b_2/2 + 1/2)} a^{1/2} \frac{(a x^2)^{b_1/2}}{(1 + a x^2)^{b+1}} \quad (272)$$

with  $a = a_1/a_2$  and  $b = (b_1 + b_2)/2$ .

**Problem 2.4:** Show that if  $X_{1,2} \sim Un(x|0, 1)$ , then for  $X = X_1^{X_2}$  we have that  $p(x) = -x^{-1} \text{Ei}(\ln x)$ , with  $\text{Ei}(z)$  the exponential integral, and  $E[X^m] = m^{-1} \ln(1 + m)$ .

Hint: Consider  $Z = \log X = X_1 \log X_2 = -X_1 W_2$  and the Mellin Transform for the Uniform and Exponential densities.

#### 2.5.2.4 Distributions with support in $\mathcal{R}$

The Mellin Transform is defined for integrable functions with non-negative support. To deal with the more general case  $X \sim p(x)$  with  $\text{supp}\{X\} = \Omega_{x \geq 0} + \Omega_{x < 0} \subseteq \mathcal{R}$  we have to

- 1) Express the density as  $p(x) = \underbrace{p(x) \mathbf{1}_{x \geq 0}(x)}_{p^+(x)} + \underbrace{p(x) \mathbf{1}_{x < 0}(x)}_{p^-(x)}$ ;
- 2) Define  $Y_1 = X$  when  $x \geq 0$  and  $Y_2 = -X$  when  $x < 0$  so  $\text{supp}\{Y_2\}$  is positive and find  $M_{Y_1}(s)$  and  $M_{Y_2}(s)$ ;
- 3) Get from the inverse transform the corresponding densities  $p_1(z)$  for the quantity of interest  $Z_1 = Z(Y_1, X_2, \dots)$  with  $M_{Y_1}(s)$  and  $p_2(z)$  for  $Z_2 = Z(Y_2, X_2, \dots)$  with  $M_{Y_2}(s)$  and at the end for  $p_2(z)$  make the corresponding change for  $X \rightarrow -X$ .

This is usually quite messy and for most cases of interest it is far easier to find the distribution for the product and ratio of random quantities with a simple change of variables.

• **Ratio of Normal and  $\chi^2$  distributed random quantities** Let's study the random quantity  $X = X_1(X_2/n)^{-1/2}$  where  $X_1 \sim N(x_1|0, 1)$  with  $\text{sup}\{X_1\} = \mathcal{R}$  and  $X_2 \sim \chi^2(x_2|n)$  with  $\text{sup}\{X_2\} = \mathcal{R}^+$ . Then, for  $X_1$  we have

$$p(x_1) = \underbrace{p(x_1) \mathbf{1}_{[0, \infty)}(x_1)}_{p^+(x_1)} + \underbrace{p(x_1) \mathbf{1}_{(-\infty, 0)}(x_1)}_{p^-(x_1)} \quad (273)$$

and therefore for  $X$

$$X \sim p(x) = p(x) \mathbf{1}_{[0, \infty)}(x) + p(x) \mathbf{1}_{(-\infty, 0)}(x) = p^+(x) + p^-(x) \quad (274)$$

Since

$$M_2(s) = \frac{2^{s-1} \Gamma(n/2 + s - 1)}{\Gamma(n/2)} \quad (275)$$

we have for  $Z = (X_2/n)^{-1/2}$  that

$$M_Z(s) = n^{(s-1)/2} M_2((3-s)/2) = \left(\frac{n}{2}\right)^{(s-1)/2} \frac{\Gamma((n+1-s)/2)}{\Gamma(n/2)} \quad (276)$$

for  $0 < \Re(s) < n+1$ . For  $X_1 \in [0, \infty)$  we have that

$$M_1^+(s) = \frac{2^{s/2} \Gamma(s/2)}{2\sqrt{2\pi}}; \quad 0 < \Re(s) \quad (277)$$

and therefore

$$M_X^+(s) = M_1^+(s) M_Z(s) = \frac{n^{s/2} \Gamma(s/2) \Gamma((n+1-s)/2)}{2\sqrt{n\pi}} \quad (278)$$

with holomorphy stripe  $0 < \Re(s) < n+1$ . There are poles at  $s_m = -2m$  with  $m = 0, 1, 2, \dots$  on the negative real axis and  $s_k = n+1+2k$  with  $k = 0, 1, 2, \dots$  on the positive real axis. Closing the contour on the left we include only  $s_m$  so

$$p^+(x) = \frac{1}{\sqrt{n\pi} \Gamma(n/2)} \sum_{m=0}^{\infty} \frac{(-1)^m}{\Gamma(m+1)} \left(\frac{x^2}{n}\right)^m \Gamma\left(m + \frac{n+1}{2}\right) = \quad (279)$$

$$= \frac{\Gamma((n+1)/2)}{\sqrt{n\pi} \Gamma(n/2)} \left(1 + \frac{x^2}{n}\right)^{-(n+1)/2} \mathbf{1}_{[0, \infty)}(x) \quad (280)$$

For  $X_1 \in (-\infty, 0)$  we should in principle define  $Y = -X_1$  with support in  $(0, \infty)$ , find  $M_Y(s)$ , obtain the density for  $X' = Y/Z$  and then obtain the corresponding one for  $X = -X'$ . However, in this case it is clear by symmetry that  $p^+(x) = p^-(x)$  and therefore

$$X \sim p(x) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi} \Gamma(n/2)} \left(1 + \frac{x^2}{n}\right)^{-(n+1)/2} \mathbf{1}_{(-\infty, \infty)}(x) = St(x|n) \quad (281)$$

• **Ratio and product of Normal distributed random quantities** Consider  $X_1 \sim N(x_1|\mu_1, \sigma_1)$  and  $X_2 \sim N(x_2|\mu_2, \sigma_2)$ . The Mellin Transform is

$$M_Y(s) = \frac{e^{-\mu^2/4\sigma^2}}{\sqrt{2\pi}} \sigma^{s-1} \Gamma(s) D_{-s}(\mp\mu/\sigma) \quad (282)$$

with  $D_a(x)$  the Whittaker Parabolic Cylinder Functions. The upper sign  $(-)$  of the argument corresponds to  $X \in [0, \infty)$  and the lower one  $(+)$  to the quantity  $Y = -X \in (0, \infty)$ . Again, the problem is considerably simplified if  $\mu_1 = \mu_2 = 0$  because

$$M_Y(z) = \frac{2^{z/2}}{2\sqrt{2\pi}} \sigma^{z-1} \Gamma(z/2) \quad (283)$$

with  $S = \langle 0, \infty \rangle$  and, due to symmetry, all contributions are the same. Thus, summing over the poles at  $z_n = -2n$  for  $n = 0, 1, 2, \dots$  we have that for  $X = X_1 X_2$  and  $a^{-1} = 4\sigma_1^2 \sigma_2^2$ :

$$p(x) = \frac{2\sqrt{a}}{\pi} \sum_{n=0}^{\infty} \frac{(\sqrt{a}|x|)^{2n}}{\Gamma(n+1)^2} (2\Psi(1+n) - \ln(\sqrt{a}|x|)) = \frac{2\sqrt{a}}{\pi} K_0(2\sqrt{a}|x|) \quad (284)$$

Dealing with the general case of  $\mu_i \neq 0$  it is much more messy to get compact expressions and life is easier with a simple change of variables. Thus, for instance for  $X = X_1/X_2$  we have that

$$p(x) = \frac{\sqrt{a_1 a_2}}{\pi} \int_{-\infty}^{\infty} e^{-\{a_1(xw - \mu_1)^2 + a_2(w - \mu_2)^2\}} |w| dw \quad (285)$$

where  $a_i = 1/(2\sigma_i^2)$  and if we define:

$$w_0 = a_2 + a_1x^2; \quad w_1 = a_1a_2(x\mu_2 - \mu_1)^2 \quad \text{and} \quad w_2 = (a_1\mu_1x + a_2\mu_2)/\sqrt{w_0} \quad (286)$$

one has:

$$p(x) = \frac{\sqrt{a_1a_2}}{\pi} \frac{1}{w_0} e^{-w_1/w_0} \left( e^{-w_2^2} + \sqrt{\pi} w_2 \operatorname{erf}(w_2) \right) \mathbf{1}_{(-\infty, \infty)}(x) \quad (287)$$

## 2.6 Ordered Samples

Let  $X \sim p(x|\theta)$  be a one-dimensional random quantity and the experiment  $e(n)$  that consists on  $n$  independent observations and results in the exchangeable sequence  $\{x_1, x_2, \dots, x_n\}$  are equivalent to an observation of the  $n$ -dimensional random quantity  $\mathbf{X} \sim p(\mathbf{x}|\theta)$  where

$$p(\mathbf{x}|\theta) = p(x_1, x_2, \dots, x_n|\theta) = \prod_{i=1}^n p(x_i|\theta) \quad (288)$$

Consider now a monotonic non-decreasing ordering of the observations

$$\underbrace{x_1 \leq x_2 \leq \dots \leq x_{k-1}}_{k-1} \leq x_k \leq \underbrace{x_{k+1} \leq \dots \leq x_{n-1} \leq x_n}_{n-k}$$

and the *Statistic of Order  $k$* ; that is, the random quantity  $X_{(k)}$  associated with the  $k^{\text{th}}$  observation ( $1 \leq k \leq n$ ) of the *ordered sample* such that there are  $k-1$  observations smaller than  $x_k$  and  $n-k$  above  $x_k$ . Since

$$P(X \leq x_k|\theta) = \int_{-\infty}^{x_k} p(x|\theta) dx = F(x_k|\theta) \quad \text{and} \quad P(X > x_k|\theta) = 1 - F(x_k|\theta) \quad (289)$$

we have that

$$\begin{aligned} X_{(k)} \sim p(x_k|\theta, n, k) &= C_{n,k} p(x_k|\theta) [F(x_k|\theta)]^{k-1} [1 - F(x_k|\theta)]^{n-k} \\ &= C_{n,k} p(x_k|\theta) \underbrace{\left[ \int_{-\infty}^{x_k} p(x|\theta) dx \right]^{k-1}}_{[P(X \leq x_k)]^{k-1}} \underbrace{\left[ \int_{x_k}^{\infty} p(x|\theta) dx \right]^{n-k}}_{[P(X > x_k)]^{n-k}} \end{aligned} \quad (290)$$

The normalization factor

$$C_{n,k} = k \binom{n}{k} \quad (291)$$

is given by combinatorial analysis although in general it is easier to get by normalization of the final density. With a similar reasoning we have that the density function of the two dimensional random quantity  $X_{(ij)} = (X_i, X_j); j > i$ , associated to the observations  $x_i$  and  $x_j$  (*Statistic of Order  $i, j; i < j$* )<sup>14</sup> will be:

$$X_{(ij)} \sim p(x_i, x_j|\theta, i, j, n) = C_{n,i,j} \underbrace{\left[ \int_{-\infty}^{x_i} p(x|\theta) dx \right]^{i-1}}_{[P(X < x_i)]^{i-1}} p(x_i|\theta) \underbrace{\left[ \int_{x_i}^{x_j} p(x|\theta) dx \right]^{j-i-1}}_{[P(x_i < X \leq x_j)]^{j-i-1}}$$

<sup>14</sup>If the random quantities  $X_i$  are not identically distributed the idea is the same but one has to deal with permutations and the expressions are more involved



$$p(x_j|\theta) \underbrace{\left[ \int_{x_j}^{\infty} p(x|\theta) dx \right]^{n-j}}_{[P(x_j < X)]^{n-j}} \quad (292)$$

where  $(x_i, x_j) \in (-\infty, x_j] \times (-\infty, \infty)$  or  $(x_i, x_j) \in (-\infty, \infty) \times [x_i, \infty)$ . Again by combinatorial analysis or integration we have that

$$C_{n,i,j} = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} \quad (293)$$

The main *Order Statistics* we are usually interested in are

- **Maximum**  $X_{(n)} = \max\{X_1, X_2, \dots, X_n\}$ :

$$p(x_n|\cdot) = n p(x_n|\theta) \left[ \int_{-\infty}^{x_n} p(x|\theta) dx \right]^{n-1} \quad (294)$$

- **Minimum**  $X_{(1)} = \min\{X_1, X_2, \dots, X_n\}$ :

$$p(x_1|\cdot) = n p(x_1|\theta) \left[ \int_{x_1}^{\infty} p(x|\theta) dx \right]^{n-1} \quad (295)$$

- **Range**  $R = X_{(n)} - X_{(1)}$

$$p(x_1, x_n|\cdot) = n(n-1) p(x_1|\theta) p(x_n|\theta) \left[ \int_{x_1}^{x_n} p(x|\theta) dx \right]^{n-2} \quad (296)$$

If  $\text{supp}(X) = [a, b]$ , then  $R \in (0, b-a)$  and

$$p(r) = n(n-1) \left\{ \int_a^{b-r} p(w+r) p(w) [F(w+r) - F(w)]^{n-2} dw \right\} \quad (297)$$

There is no explicit form unless we specify the Distribution Function  $F(x|\theta)$ .

- **Difference**  $S = X_{(i+1)} - X_{(i)}$ . If  $\text{supp}(X) = [a, b]$ , then  $S \in (0, b-a)$  and

$$p(s) = \frac{\Gamma(n+1)}{\Gamma(i)\Gamma(n-i)} \left\{ \int_a^{b-s} p(w+s) p(w) [F(w)]^{i-1} [1 - F(w+s)]^{n-i-1} dw \right\} \quad (298)$$

In the case of discrete random quantities, the idea is the same but a bit more messy because one has to watch for the discontinuities of the Distribution Function. Thus, for instance:

- **Maximum**  $X_{(n)} = \max\{X_1, X_2, \dots, X_n\}$ :  
 $X_{(n)} \leq x$  iff **all**  $x_i$  are **less or equal**  $x$  and this happens with probability

$$P(x_n \leq x) = [F(x)]^n \quad (299)$$

$X_{(n)} < x$  iff **all**  $x_i$  are **less than**  $x$  and this happens with probability

$$P(x_n < x) = [F(x-1)]^n \quad (300)$$

Therefore

$$P(x_n = x) = P(x_n \leq x) - P(x_n < x) = [F(x)]^n - [F(x-1)]^n \quad (301)$$

- **Minimum**  $X_{(1)} = \min\{X_1, X_2, \dots, X_n\}$ :  
 $X_{(1)} \geq x$  iff **all**  $x_i$  are **grater or equal**  $x$  and this happens with probability

$$P(x_1 \geq x) = 1 - P(x_1 < x) = [1 - F(x - 1)]^n \quad (302)$$

$X_{(1)} > x$  iff **all**  $x_i$  are **greater than**  $x$  and this happens with probability

$$P(x_1 > x) = 1 - P(x_1 \leq x) = [1 - F(x)]^n \quad (303)$$

Therefore

$$\begin{aligned} P(x_1 = x) &= P(x_1 \leq x) - P(x_1 < x) = [1 - P(x_1 > x)] - [1 - P(x_1 \geq x)] = \\ &= [1 - F(x - 1)]^n - [1 - F(x)]^n \end{aligned} \quad (304)$$

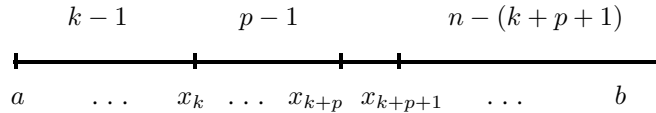
**Example 2.26:** Let  $X \sim Un(x|a, b)$  and an iid sample of size  $n$ . Then, if  $L = b - a$ :

- **Maximum:**  $p(x_n) = n \frac{(x_n - a)^{n-1}}{(b - a)^n} \mathbf{1}_{(a,b)}(x_n)$
- **Minimum:**  $p(x_1) = n \frac{(b - x_1)^{n-1}}{(b - a)^n} \mathbf{1}_{(a,b)}(x_1)$
- **Range:**  $R = X_{(n)} - X_{(1)} : p(r) = \frac{n(n-1)}{L} \left(\frac{r}{L}\right)^{n-2} \left(1 - \frac{r}{L}\right) \mathbf{1}_{(0,L)}(r)$
- **Difference:**  $S = X_{(k+1)} - X_{(k)} : p(s) = \frac{n}{L} \left(1 - \frac{s}{L}\right)^{n-1} \mathbf{1}_{(0,L)}(s)$

**Example 2.27:** Let's look at the Uniform distribution in more detail. Consider a random quantity  $X \sim Un(x|a, b)$ , the experiment  $e(n)$  that provides a sample of  $n$  independent events and the ordered sample

$$\mathbf{X}_n = \{x_1 \leq x_2 \leq \dots \leq x_k \leq \dots \leq x_{k+p} \leq \dots \leq x_{n-1} \leq x_n\} \quad (305)$$

Then, for the ordered statistics  $X_k, X_{k+p}$  and  $X_{k+p+1}$  with  $k, p \in \mathcal{N}, 1 \leq k \leq n - 1$  and  $p \leq n - k - 1$  we have that



$$p(x_k, x_{k+p} | a, b, n, p) \propto \left[ \int_a^{x_k} ds_1 \right]^{k-1} \left[ \int_{x_k}^{x_{k+p}} ds_2 \right]^{p-1} \left[ \int_{x_{k+p+1}}^b ds_3 \right]^{n-(k+p+1)} \quad (306)$$

Let's think for instance that those are the arrival times of  $n$  events collected with a detector in a time window  $[a = 0, b = T]$ . If we define  $w_1 = x_{k+p} - x_k$  and  $w_2 = x_{k+p+1} - x_k$  we have that

$$p(x_k, w_1, w_2 | T, n, p) = x_k^{k-1} w_1^{p-1} (T - x_k - w_2)^{n-k-p-1} \mathbf{1}_{[0, T-w_2]}(x_k) \mathbf{1}_{[0, w_2]}(w_1) \mathbf{1}_{[0, T]}(w_2) \quad (307)$$

and, after integration of  $x_k$ :

$$p(w_1, w_2 | T, n, p) = \binom{n}{p} \frac{p(n-p)}{T^n} w_1^{p-1} (T - w_2)^{n-p-1} \mathbf{1}_{[0, w_2]}(w_1) \mathbf{1}_{[0, T]}(w_2) \quad (308)$$

Observe that the support can be expressed also as  $\mathbf{1}_{[0, T]}(w_1) \mathbf{1}_{[w_1, T]}(w_2)$  and that the distribution of  $(W_1, W_2)$  does not depend on  $k$ . The marginal densities are given by:

$$p(w_1 | T, n, p) = \binom{n}{p} \frac{p}{T^n} w_1^{p-1} (T - w_1)^{n-p} \mathbf{1}_{[0, T]}(w_1) \quad (309)$$

$$p(w_2|T, n, p) = \binom{n}{p} \frac{n-p}{T^n} w_2^p (T - w_2)^{n-p-1} \mathbf{1}_{[0, T]}(w_2) \quad (310)$$

and if we take the limit  $T \rightarrow \infty$  and  $n \rightarrow \infty$  keeping the the rate  $\lambda = n/T$  constant we have

$$\lim_{T, n \rightarrow \infty} p(w_1, w_2|T, n, p) = p(w_1, w_2|\lambda, p) = \frac{\lambda^{p+1}}{\Gamma(p)} e^{-\lambda w_2} w_1^{p-1} \mathbf{1}_{[0, w_2]}(w_1) \mathbf{1}_{[0, \infty)}(w_2) \quad (311)$$

and

$$p(w_1|\lambda, p) = \frac{\lambda^p}{\Gamma(p)} e^{-\lambda w_1} w_1^{p-1} \mathbf{1}_{[0, \infty)}(w_1) \quad (312)$$

In consequence, under the stated conditions the time difference between two consecutive events ( $p = 1$ ) tends to an exponential distribution. Let's consider for simplicity this limiting behaviour in what follows and leave as an exercise the more involved case of finite time window  $T$ .

Suppose now that after having observed one event, say  $x_k$ , we have a dead-time of size  $a$  in the detector during which we can not process any data. All the events that fall in  $(x_k, x_k + a)$  are lost (unless we play with buffers). If the next observed event is at time  $x_{k+p+1}$ , we have lost  $p$  events and the probability for this to happen is

$$\mathcal{P}(w_1 \leq a, w_2 \geq a|\lambda, p) = e^{-\lambda a} \frac{(\lambda a)^p}{\Gamma(p+1)} \quad (313)$$

that is,  $N_{\text{lost}} \sim Po(p|\lambda a)$  regardless the position of the last recorded time ( $x_k$ ) in the ordered sequence. As one could easily have intuited, the expected number of events lost for each observed one is  $E[N_{\text{lost}}] = \lambda a$ . Last, it is clear that the density for the time difference between two consecutive observed events when  $p$  are lost due to the dead-time is

$$p(w_2|w_1 \leq a, \lambda, p) = \lambda e^{-\lambda(w_2-a)} \mathbf{1}_{[a, \infty)}(w_2) \quad (314)$$

Note that it depends on the dead-time window  $a$  and not on the number of events lost.

**Example 2.28:** Let  $X \sim Ex(x|\lambda)$  and an iid sample of size  $n$ . Then:

- *Maximum:*  $p(x_n) = n \lambda e^{-\lambda x_n} (1 - e^{-\lambda x_n})^{n-1} \mathbf{1}_{(0, \infty)}(x_n)$
- *Minimum:*  $p(x_1) = n \lambda e^{-\lambda n x_1} \mathbf{1}_{(0, \infty)}(x_1)$
- *Range:*  $R = X_{(n)} - X_{(1)} : p(r) = (n-1) \lambda^{n-1} e^{-\lambda r} [1 - e^{-\lambda r}]^{n-2} \mathbf{1}_{(0, \infty)}(r)$
- *Difference:*  $S = X_{(k+1)} - X_{(k)} : p(s) = (n-k) \lambda e^{-\lambda s} [1 - e^{-\lambda s}]^{n-k-1} \mathbf{1}_{(0, \infty)}(s)$

## 2.7 Limit Theorems and Convergence

In Probability, the Limit Theorems are statements that, under the conditions of applicability, describe the behavior of a sequence of random quantities or of Distribution Functions. In principle, whenever we can define a distance (or at least a positive defined set function) we can establish a convergence criteria and, obviously, some will be stronger than others so, for instance, a sequence of random quantities  $\{X_i\}_{i=1}^{\infty}$  may converge according to one criteria and not to other. The most usual types of convergence, their relation and the Theorems derived from them are:

Distribution	$\implies$ Central Limit Theorem
$\uparrow$	$\implies$ Glivenko-Cantelly Theorem (weak form)
$\uparrow$	
Probability	$\implies$ Weak Law of Large Numbers
$\uparrow$ $\uparrow$	
$\uparrow$ Almost Sure	$\implies$ Strong Law of Large Numbers
$\uparrow$	
$L_p(\mathcal{R})$ Norm	$\implies$ Convergence in Quadratic Mean
$\uparrow$	
Uniform	$\implies$ Glivenko-Cantelly Theorem

so *Convergence in Distribution* is the weakest of all since does not imply any of the others. In principle, there will be no explicit mention to statistical independence of the random quantities of the sequence nor to an specific Distribution Function. In most cases we shall just state the different criteria for convergence and refer to the literature, for instance [?], for further details and demonstrations. Let's start with the very useful Chebyshev's Theorem.

**2.7.1 Chebyshev's Theorem**

Let  $X$  be a random quantity that takes values in  $\Omega \subset \mathcal{R}$  with Distribution Function  $F(x)$  and consider the random quantity  $Y = g(X)$  with  $g(X)$  a non-negative single valued function for all  $X \in \Omega$ . Then, for  $\alpha \in \mathcal{R}^+$

$$P(g(X) \geq \alpha) \leq \frac{E[g(X)]}{\alpha} \tag{315}$$

In fact, given a measure space  $(\Omega, \mathcal{B}_\Omega, \mu)$ , for any  $\mu$ -integrable function  $f(x)$  and  $c > 0$  we have for  $A = \{x : |f(x)| \geq c\}$  that  $c\mathbf{1}_A(x) \leq |f(x)|$  for all  $x$  and therefore

$$c\mu(A) = \int c\mathbf{1}_A(x) d\mu \leq \int |f(x)| d\mu \tag{316}$$

Let's see two particular cases. First, consider  $g(X) = (X - \mu)^{2n}$  where  $\mu = E[X]$  and  $n$  a positive integer such that  $g(X) \geq 0 \forall X \in \Omega$ . Applying Chebishev's Theorem:

$$P((X - \mu)^{2n} \geq \alpha) = P(|X - \mu| \geq \alpha^{1/2n}) \leq \frac{E[(X - \mu)^{2n}]}{\alpha} = \frac{\mu_{2n}}{\alpha} \tag{317}$$

For  $n = 1$ , if we take  $\alpha = k^2\sigma^2$  we get the *Bienaymé-Chebyshev's* inequality

$$P(|X - \mu| \geq k\sigma) \leq 1/k^2 \tag{318}$$

that is, whatever the Distribution Function of the random quantity  $X$  is, the probability that  $X$  differs from its expected value  $\mu$  more than  $k$  times its standard deviation is less or equal than  $1/k^2$ . As a second case, assume  $X$  takes only positive real values and has a first order moment  $E[X] = \mu$ . Then (Markov's inequality):

$$P(X \geq \alpha) \leq \frac{\mu}{\alpha} \quad \xrightarrow{\alpha=k\mu} \quad P(X \geq k\mu) \leq 1/k \tag{319}$$

The Markov and Bienaymé-Chebyshev's inequalities provide upper bounds for the probability knowing just mean value and the variance although they are usually very conservative. They can be considerably improved if we have more information about the Distribution Function but, as we shall see, the main interest of Chebishev's inequality lies on its importance to prove Limit Theorems.

### 2.7.2 Convergence in Probability

The sequence of random quantities  $\{X_n(w)\}_{n=1}^{\infty}$  converges in probability to  $X(w)$  iff:

$$\lim_{n \rightarrow \infty} P(|X_n(w) - X(w)| \geq \epsilon) = 0 \quad ; \quad \forall \epsilon > 0 \quad ; \quad (320)$$

or, equivalently, iff:

$$\lim_{n \rightarrow \infty} P(|X_n(w) - X(w)| < \epsilon) = 1 \quad \forall \epsilon > 0 \quad ; \quad (321)$$

Note that  $P(|X_n(w) - X(w)| \geq \epsilon)$  is a real number so this is the usual limit for a sequence of real numbers and, in consequence, for all  $\epsilon > 0$  and  $\delta > 0 \exists n_0(\epsilon, \delta)$  such that for all  $n > n_0(\epsilon, \delta)$  it holds that  $P(|X_n(w) - X(w)| \geq \epsilon) < \delta$ . For a sequence of n-dimensional random quantities, this can be generalized to  $\lim_{n \rightarrow \infty} P(\|X_n(w), X(w)\|) < \delta$  and, as said earlier, Convergence in Probability implies Convergence in Distribution but the converse is not true. An important consequence of the Convergence in Probability is the

• **Weak Law of Large Numbers:** Consider a sequence of independent random quantities  $\{X_i(w)\}_{i=1}^{\infty}$ , all with the same Distribution Function and first order moment  $E[X_i(w)] = \mu$ , and define a new random quantity

$$Z_n(w) = \frac{1}{n} \sum_{i=1}^n X_i(w) \quad (322)$$

The, the sequence  $\{Z_n(w)\}_{n=1}^{\infty}$  converges in probability to  $\mu$ ; that is:

$$\lim_{n \rightarrow \infty} P(|Z_n(w) - \mu| \geq \epsilon) = 0 \quad ; \quad \forall \epsilon > 0 \quad ; \quad (323)$$

The Law of Large Numbers was stated first by J. Bernoulli in 1713 for the Binomial Distribution, generalized (and named *Law of Large Numbers*) by S.D. Poisson and shown in the general case by A. Khinchin in 1929. In the case  $X_i(w)$  have variance  $V(X_i) = \sigma^2$  it is straight forward from Chebishev's inequality:

$$P(|Z_n - \mu| \geq \epsilon) = P((Z_n - \mu)^2 \geq \epsilon^2) \leq \frac{E[(Z_n - \mu)^2]}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \quad (324)$$

Intuitively, Convergence in Probability means that when n is very large, the probability that  $Z_n(w)$  differs from  $\mu$  by a small amount is very small; that is,  $Z_n(w)$  gets more concentrated around  $\mu$ . But “very small” is not zero and it may happen that for some  $k > n$   $Z_k$  differs from  $\mu$  by more than  $\epsilon$ . An stronger criteria of convergence is the *Almost Sure Convergence*.

### 2.7.3 Almost Sure Convergence

A sequence  $\{X_n(w)\}_{n=1}^{\infty}$  of random quantities converges *almost sure* to  $X(w)$  if, and only if:

$$\lim_{n \rightarrow \infty} X_n(w) = X(w) \quad (325)$$

for all  $w \in \Omega$  except at most on a set  $W \subset \Omega$  of zero measure ( $P(W) = 0$  so it is also referred to as *convergence almost everywhere*). This means that for all  $\epsilon > 0$  and all  $w \in W^c = \Omega - W$ ,  $\exists n_0(\epsilon, w) > 0$  such that  $|X_n(w) - X(w)| < \epsilon$  for all  $n > n_0(\epsilon, w)$ . Thus, we have the equivalent forms:

$$P \left[ \lim_{n \rightarrow \infty} |X_n(w) - X(w)| \geq \epsilon \right] = 0 \quad \text{or} \quad P \left[ \lim_{n \rightarrow \infty} |X_n(w) - X(w)| < \epsilon \right] = 1 \quad (326)$$

for all  $\epsilon > 0$ . Needed less to say that the random quantities  $X_1, X_2 \dots$  and  $X$  are defined on the same probability space. Again, Almost Sure Convergence implies Convergence in Probability but the converse is not true. An important consequence of the Almost Sure Convergence is the:

- **Strong Law of Large Numbers** (E. Borel 1909, A.N. Kolmogorov,...) : Let  $\{X_i(w)\}_{i=1}^{\infty}$  be a sequence of independent random quantities all with the same Distribution Function and first order moment  $E[X_i(w)] = \mu$ . Then the sequence  $\{Z_n(w)\}_{n=1}^{\infty}$  with

$$Z_n(w) = \frac{1}{n} \sum_{i=1}^n X_i(w) \tag{327}$$

converges almost sure to  $\mu$ ; that is:

$$P \left[ \lim_{n \rightarrow \infty} |Z_n(w) - \mu| \geq \epsilon \right] = 0 \quad \forall \epsilon > 0 \tag{328}$$

Intuitively, Almost Sure Convergence means that the probability that for some  $k > n$ ,  $Z_k$  differs from  $\mu$  by more than  $\epsilon$  becomes smaller as  $n$  grows.

### 2.7.4 Convergence in Distribution

Consider the sequence of random quantities  $\{X_n(w)\}_{n=1}^{\infty}$  and of their corresponding Distribution Functions  $\{F_n(x)\}_{n=1}^{\infty}$ . In the limit  $n \rightarrow \infty$ , the random quantity  $X_n(w)$  tends to be distributed as  $X(w) \sim F(x)$  iff

$$\lim_{n \rightarrow \infty} F_n(x) = F(x) \iff \lim_{n \rightarrow \infty} P(X_n \leq x) = P(X \leq x) \quad ; \quad \forall x \in C(F) \tag{329}$$

with  $C(F)$  the set of points of continuity of  $F(x)$ . Expressed in a different manner, the sequence  $\{X_n(w)\}_{n=1}^{\infty}$  Converges in Distribution to  $X(w)$  if, and only if, for all  $\epsilon > 0$  and  $x \in C(F)$ ,  $\exists n_0(\epsilon, x)$  such that  $|F_n(x) - F(x)| < \epsilon$ ,  $\forall n > n_0(\epsilon, x)$ . Note that, in general,  $n_0$  depends on  $x$  so it is possible that, given an  $\epsilon > 0$ , the value of  $n_0$  for which the condition  $|F_n(x) - F(x)| < \epsilon$  is satisfied for certain values of  $x$  may not be valid for others. It is important to note also that we have not made any statement about the statistical independence of the random quantities and that the Convergence in Distribution is determined only by the Distribution Functions so the corresponding random quantities do not have to be defined on the same probability space. To study the Convergence in Distribution, the following theorem it is very useful:

- **Theorem** (Lévy 1937; Cramèr 1937) : Consider a sequence of Distribution Functions  $\{F_n(x)\}_{n=1}^{\infty}$  and of the corresponding Characteristic Functions  $\{\Phi_n(t)\}_{n=1}^{\infty}$ . Then

- ▷ if  $\lim_{n \rightarrow \infty} F_n(x) = F(x)$ , then  $\lim_{n \rightarrow \infty} \Phi_n(t) = \Phi(t)$  for all  $t \in \mathcal{R}$  with  $\Phi(t)$  the Characteristic Function of  $F(x)$ .
- ▷ Conversely, if  $\Phi_n(t) \xrightarrow{n \rightarrow \infty} \Phi(t) \forall t \in \mathcal{R}$  and  $\Phi(t)$  is continuous at  $t = 0$ , then  $F_n(x) \xrightarrow{n \rightarrow \infty} F(x)$

This criteria of convergence is weak in the sense that if there is convergence if probability or almost sure or in quadratic mean then there is convergence in distribution but the converse is not necessarily true. However, there is a very important consequence of the Convergence in Distribution:

- **Central Limit Theorem** (Lindberg-Levy) : Let  $\{X_i(w)\}_{i=1}^{\infty}$  be a sequence of independent random quantities all with the same Distribution Function and with second order moments so  $E[X_i(w)] = \mu$  and  $V[X_i(w)] = \sigma^2$ . Then the sequence  $\{Z_n(w)\}_{n=1}^{\infty}$  of random quantities

$$Z_n(w) = \frac{1}{n} \sum_{i=1}^n X_i(w) \tag{330}$$

with

$$E[Z_n] = \frac{1}{n} \sum_{i=1}^n E[X_i] = \mu \quad \text{and} \quad V[Z_n] = \frac{1}{n^2} \sum_{i=1}^n V[X_i] = \frac{\sigma^2}{n} \quad (331)$$

tends, in the limit  $n \rightarrow \infty$ , to be distributed as  $N(z|\mu, \sigma/\sqrt{n})$  or, what is the same, the *standardized* random quantity

$$\tilde{Z}_n = \frac{Z_n - \mu}{\sqrt{V[Z_n]}} = \frac{\frac{1}{n} \sum_{i=1}^n X_i - \mu}{\sigma/\sqrt{n}} \quad (332)$$

tends to be distributed as  $N(x|0, 1)$ .

Consider, without loss of generality, the random quantity  $W_i = X_i - \mu$  so that  $E[W_i] = E[X_i] - \mu = 0$  and  $V[W_i] = V[X_i] = \sigma^2$ . Then,

$$\Phi_W(t) = 1 - \frac{1}{2}t^2\sigma^2 + \mathcal{O}(t^k) \quad (333)$$

Since we require that the random quantities  $X_i$  have at least moments of order two, the remaining terms  $\mathcal{O}(t^k)$  are either zero or powers of  $t$  larger than 2. Then,

$$Z_n = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{n} \sum_{i=1}^n W_i + \mu \quad ; \quad E[Z_n] = \mu \quad ; \quad V[Z_n] = \sigma_{Z_n}^2 = \frac{\sigma^2}{n} \quad (334)$$

so

$$\Phi_{Z_n}(t) = e^{it\mu} [\Phi_W(t/n)]^n \quad \longrightarrow \quad \lim_{n \rightarrow \infty} \Phi_{Z_n}(t) = e^{it\mu} \lim_{n \rightarrow \infty} [\Phi_W(t/n)]^n \quad (335)$$

Now, since:

$$\Phi_W(t/n) = 1 - \frac{1}{2} \left(\frac{t}{n}\right)^2 \sigma^2 + \mathcal{O}(t^k/n^k) = 1 - \frac{1}{2} \frac{t^2}{n} \sigma_{Z_n}^2 + \mathcal{O}(t^k/n^k) \quad (336)$$

we have that:

$$\lim_{n \rightarrow \infty} [\Phi_W(t/n)]^n = \lim_{n \rightarrow \infty} \left[ 1 - \frac{1}{2} \frac{t^2}{n} \sigma_{Z_n}^2 + \mathcal{O}(t^k/n^k) \right]^n = \exp \left\{ -\frac{1}{2} t^2 \sigma_{Z_n}^2 \right\} \quad (337)$$

and therefore:

$$\lim_{n \rightarrow \infty} \Phi_{Z_n}(t) = e^{it\mu} e^{-\frac{1}{2} t^2 \sigma^2 / n} \quad (338)$$

so,  $\lim_{n \rightarrow \infty} Z_n \sim N(x|\mu, \sigma/\sqrt{n})$ .

The first indications about the Central Limit Theorem are due to A. De Moivre (1733). Later, C.F. Gauss and P.S. Laplace enunciated the behavior in a general way and, in 1901, A. Lyapunov gave the first rigorous demonstration under more restrictive conditions. The theorem in the form we have presented here is due to Lindeberg and Lévy and requires that the random quantities  $X_i$  are:

- i) Statistically Independent;
- ii) have the same Distribution Function;
- iii) First and Second order moments exist (i.e. they have mean value and variance).

In general, there is a set of Central Limit Theorems depending on which of the previous conditions are satisfied and justify the empirical fact that many natural phenomena are adequately described by the Normal Distribution. To quote E. T. Whittaker and G. Robinson (*Calculus of Observations*):

"Everybody believes in the exponential law of errors;  
 The experimenters because they think that it can be proved by mathematics;  
 and the mathematicians because they believe it has been established by observation"

**Example 2.29:** From the limiting behavior of the Characteristic Function, show that:

- If  $X \sim Bi(r|n, p)$ , in the limit  $p \rightarrow 0$  with  $np$  constant tends to a Poisson Distribution  $Po(r|\mu = np)$ ;
- If  $X \sim Bi(r|n, p)$ , in the limit  $n \rightarrow \infty$  the standardized random quantity

$$Z = \frac{X - \mu_X}{\sigma_X} = \frac{X - np}{\sqrt{npq}} \stackrel{n \rightarrow \infty}{\sim} N(x|0, 1) \tag{339}$$

- If  $X \sim Po(r|\mu)$ , then

$$Z = \frac{X - \mu_X}{\sigma_X} = \frac{X - \mu}{\sqrt{\mu}} \stackrel{\mu \rightarrow \infty}{\sim} N(x|0, 1) \tag{340}$$

- $X \sim \chi^2(x|n)$ , then  $n \rightarrow \infty$  the standardized random quantity

$$Z = \frac{X - \mu_X}{\sigma_X} = \frac{X - \nu}{\sqrt{2\nu}} \stackrel{n \rightarrow \infty}{\sim} N(x|0, 1) \tag{341}$$

- The Student's Distribution  $St(x|0, 1, \nu)$  converges to  $N(x|0, 1)$  in the limit  $\nu \rightarrow \infty$ ;
- The Snedecor's Distribution  $Sn(x|\nu_1, \nu_2)$  converges to  $\chi^2(x|\nu_1)$  in the limit  $\nu_2 \rightarrow \infty$ , to  $St(x|0, 1, \nu_2)$  in the limit  $\nu_1 \rightarrow \infty$  and to  $N(x|0, 1)$  in the limit  $\nu_1, \nu_2 \rightarrow \infty$ .

**Example 2.30:** It is interesting to see the Central Limit Theorem at work. For this, we have done a Monte Carlo sampling of the random quantity  $X \sim Un(x|0, 1)$ . The sampling distribution is shown in the figure ?? (1) and the following ones (2-6) show the sample mean of  $n = 2$  (2), 5 (3), 10 (4), 20 (5) y 50 (6) consecutive values. Each histogram has 500000 events and, as you can see, as  $n$  grows the distribution "looks" more Normal. For  $n = 20$  and  $n = 50$  the Normal distribution is superimposed.

The same behavior is observed in figure ?? where we have generated a sequence of values from a parabolic distribution with minimum at  $x = 1$  and support on  $\Omega = [0, 2]$ .

Last, figure ?? shows the results for a sampling from the Cauchy Distribution  $X \sim Ca(x|0, 1)$ . As you can see, the sampling averages follow a Cauchy Distribution regardless the value of  $n$ . For  $n = 20$  and  $n = 50$  a Cauchy and a Normal distributions have been superimposed. In this case, since the Cauchy Distribution has no moments the Central Limit Theorem does not apply.

**Example 2.31:** Let  $\{X_i(w)\}_{i=1}^{\infty}$  be a sequence of independent random quantities all with the same Distribution Function, mean value  $\mu$  and variance  $\sigma^2$  and consider the random quantity

$$Z(w) = \frac{1}{n} \sum_{i=1}^n X_i(w) \tag{342}$$

What is the value of  $n$  such that the probability that  $Z$  differs from  $\mu$  more than  $\epsilon$  is less than  $\delta = 0.01$ ?

From the Central Limit Theorem we know that in the limit  $n \rightarrow \infty$ ,  $Z \sim N(x|\mu, \sigma/\sqrt{n})$  so we may consider that, for large  $n$ :

$$P(|Z - \mu| \geq \epsilon) = P(\mu - \epsilon \geq Z \geq \mu + \epsilon) \simeq \tag{343}$$

$$\simeq \int_{-\infty}^{\mu - \epsilon} N(x|\mu, \sigma) dx + \int_{\mu + \epsilon}^{+\infty} N(x|\mu, \sigma) dx = 1 - \operatorname{erf} \left[ \frac{\sqrt{n}\epsilon}{\sigma\sqrt{2}} \right] < \delta \tag{344}$$



For  $\delta = 0.01$  we have that

$$\frac{\sqrt{n}\epsilon}{\sigma} \geq 2.575 \quad \longrightarrow \quad n \geq \frac{6.63 \sigma^2}{\epsilon^2} \quad (345)$$

### 2.7.5 Convergence in $L_p$ Norm

A sequence of random quantities  $\{X_n(w)\}_{n=1}^{\infty}$  converges to  $X(w)$  in  $L_p(\mathcal{R})$  ( $p \geq 1$ ) norm iff,

$$X(w) \in L_p(\mathcal{R}), \quad X_n(w) \in L_p(\mathcal{R}) \quad \forall n \quad \text{and} \quad \lim_{n \rightarrow \infty} E[|X_n(w) - X(w)|^p] = 0 \quad (346)$$

that is, iff for any real  $\epsilon > 0$  there exists a natural  $n_0(\epsilon) > 0$  such that for all  $n \geq n_0(\epsilon)$  it holds that  $E[|X_n(w) - X(w)|^p] < \epsilon$ . In the particular case that  $p = 2$  it is called *Convergence in Quadratic Mean*.

From Chebyshev's Theorem

$$P(|X_n(w) - X(w)| \geq \alpha^{1/p}) \leq \frac{E[(X_n(w) - X(w))^p]}{\alpha} \quad (347)$$

so, taking  $\alpha = \epsilon^p$ , if there is convergence in  $L_p(\mathcal{R})$  norm:

$$\lim_{n \rightarrow \infty} P(|X_n(w) - X(w)| \geq \epsilon) \leq \lim_{n \rightarrow \infty} \frac{E[(X_n(w) - X(w))^p]}{\epsilon^p} = 0 \quad \forall \epsilon > 0 \quad (348)$$

and, in consequence, we have convergence in probability.

### 2.7.6 Uniform Convergence

In some cases, point-wise convergence of Distribution Functions is not strong enough to guarantee the desired behavior and we require a stronger type of convergence. To some extent one may think that, more than a criteria of convergence, Uniform Convergence refers to the way in which it is achieved. Point-wise convergence requires the existence of an  $n_0$  that may depend on  $\epsilon$  and on  $x$  so that the condition  $|f_n(x) - f(x)| < \epsilon$  for  $n \geq n_0$  may be satisfied for some values of  $x$  and not for others, for which a different value of  $n_0$  is needed. The idea behind uniform convergence is that we can find a value of  $n_0$  for which the condition is satisfied regardless the value of  $x$ . Thus, we say that a sequence  $\{f_n(x)\}_{n=1}^{\infty}$  converges uniformly to  $f(x)$  iff:

$$\forall \epsilon > 0, \quad \exists n_0 \in \mathcal{N} \quad \text{such that} \quad |f_n(x) - f(x)| < \epsilon \quad \forall n > n_0 \quad \text{and} \quad \forall x \quad (349)$$

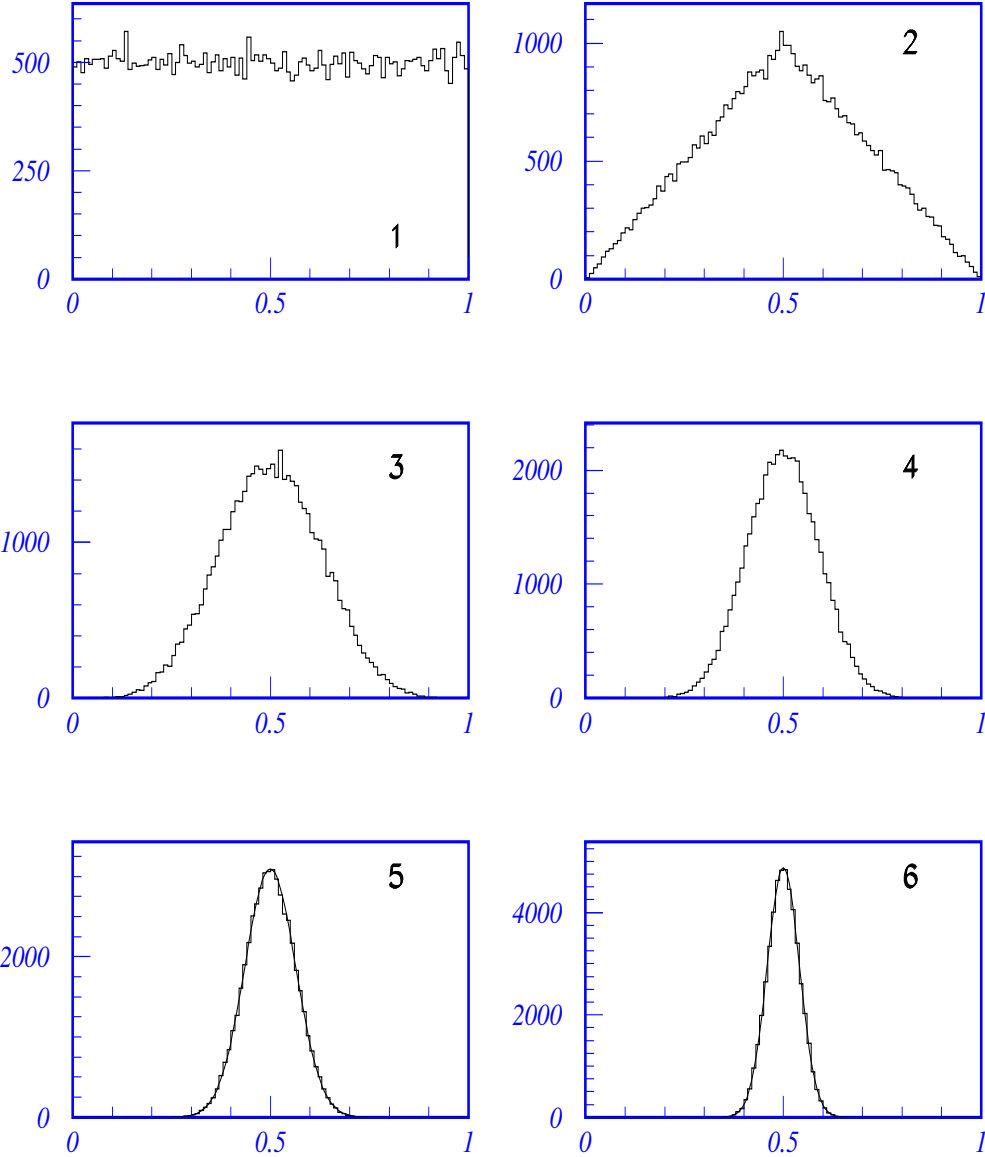
or, in other words, iff:

$$\sup_x |f_n(x) - f(x)| \xrightarrow{n \rightarrow \infty} 0 \quad (350)$$

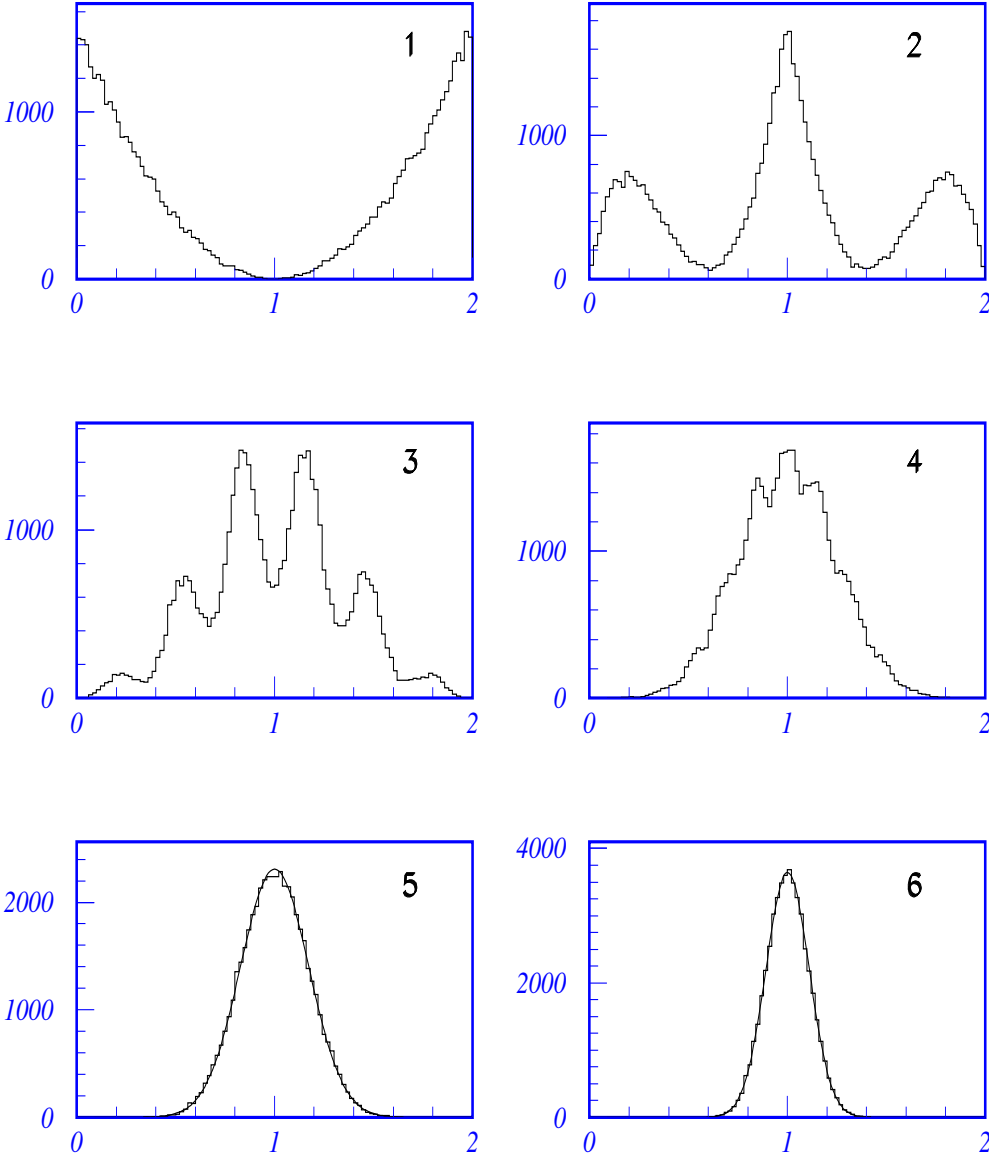
Thus, it is a stronger type of convergence that implies point-wise convergence. Intuitively, one may visualize the uniform convergence of  $f_n(x)$  to  $f(x)$  if one can draw a band  $f(x) \pm \epsilon$  that contains all  $f_n(x)$  for any  $n$  sufficiently large. Look for instance at the sequence of functions  $f_n(x) = x(1 + 1/n)$  with  $n = 1, 2, \dots$  and  $x \in \mathcal{R}$ . It is clear that converges point-wise to  $f(x) = x$  because  $\lim_{n \rightarrow \infty} f_n(x) = f(x)$  for all  $x \in \mathcal{R}$ ; that is, if we take  $n_0(x, \epsilon) = x/\epsilon$ , for all  $n > n_0(x, \epsilon)$  it is true that  $|f_n(x) - f(x)| < \epsilon$  but for larger values of  $x$  we need larger values of  $n$ . Thus, the the convergence is not uniform because

$$\sup_x |f_n(x) - f(x)| = \sup_x |x/n| = \infty \quad \forall n \in \mathcal{N} \quad (351)$$

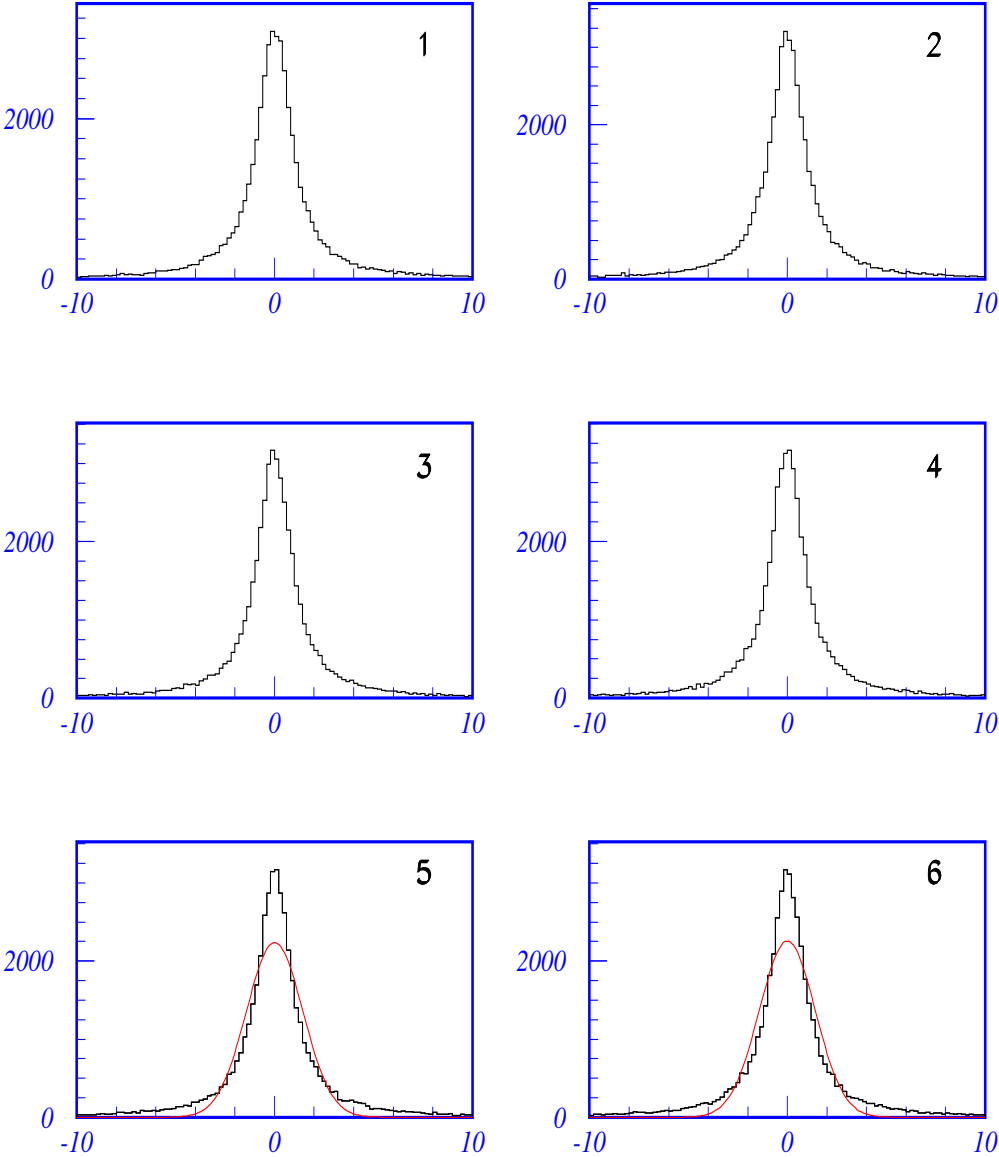
Intuitively, for whatever small a given  $\epsilon$  is, the band  $f(x) \pm \epsilon = x \pm \epsilon$  does not contain  $f_n(x)$  for all  $n$  sufficiently large. As a second example, take  $f_n(x) = x^n$  with  $x \in (0, 1)$ . We have that  $\lim_{n \rightarrow \infty} f_n(x) = 0$  but  $\sup_x |g_n(x)| = 1$  so the convergence is not uniform. For the cases we shall be interested in, if a



**Fig. 2:** Generated sample from  $Un(x|0, 1)$  (1) and sampling distribution of the mean of 2 (2), 5 (3), 10 (4), 20 (5) y 50 (6) generated values.



**Fig. 3:** Generated sample from a parabolic distribution with minimum at  $x = 1$  and support on  $\Omega = [0, 2]$  (1) and sampling distribution of the mean of 2 (2), 5 (3), 10 (4), 20 (5) y 50 (6) generated values.



**Fig. 4:** Generated sample from a Cauchy distribution  $Ca(x|0, 1)$  (1) and sampling distribution of the mean of 2 (2), 5 (3), 10 (4), 20 (5) y 50 (6) generated values.

Distribution Function  $F(x)$  is continuous and the sequence of  $\{F_n(x)\}_{n=1}^{\infty}$  converges in distribution to  $F(x)$  (i.e. point-wise) then it does *uniformly* too. An important case of uniform convergence is the (sometimes called *Fundamental Theorem of Statistics*):

• **Glivenko-Cantelli Theorem** (V. Glivenko-F.P. Cantelli; 1933) : Consider the random quantity  $X \sim F(x)$  and a statistically independent (essential point) sampling of size  $n$   $\{x_1, x_2, \dots, x_n\}$ . The *empirical Distribution Function*

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(-\infty, x]}(x_i) \quad (352)$$

converges uniformly to  $F(x)$ ; that is (Kolmogorov-Smirnov Statistic):

$$\lim_{n \rightarrow \infty} \sup_x |F_n(x) - F(x)| = 0 \quad (353)$$

Let's see the convergence in probability, in quadratic mean and, in consequence, in distribution. For a fixed value  $x = x_0$ ,  $Y = \mathbf{1}_{(-\infty, x_0]}(X)$  is a random quantity that follows a Bernoulli distribution with probability

$$p = P(Y = 1) = P(\mathbf{1}_{(-\infty, x_0]}(x) = 1) = P(X \leq x_0) = F(x_0) \quad (354)$$

$$P(Y = 0) = P(\mathbf{1}_{(-\infty, x_0]}(x) = 0) = P(X > x_0) = 1 - F(x_0) \quad (355)$$

and Characteristic Function

$$\Phi_Y(t) = E[e^{itY}] = e^{it} p + (1 - p) = e^{it} F(x_0) + (1 - F(x_0)) \quad (356)$$

Then, for a fixed value of  $x$  we have for the random quantity

$$Z_n(x) = \sum_{i=1}^n \mathbf{1}_{(-\infty, x]}(x_i) = n F_n(x) \quad \longrightarrow \quad \Phi_{Z_n}(t) = (e^{it} F(x) + (1 - F(x)))^n \quad (357)$$

and therefore  $Z_n(x) \sim Bi(k|n, F(x))$  so, if  $W = n F_n(x)$ , then

$$P(W = k|n, F(x)) = \binom{n}{k} F(x)^k (1 - F(x))^{n-k} \quad (358)$$

with

$$\begin{aligned} E[W] &= n F(x) & \longrightarrow & E[F_n(x)] = F(x) \\ V[W] &= n F(x) (1 - F(x)) & \longrightarrow & V[F_n(x)] = \frac{1}{n} F(x) (1 - F(x)) \end{aligned} \quad (359)$$

From Chebishev's Theorem

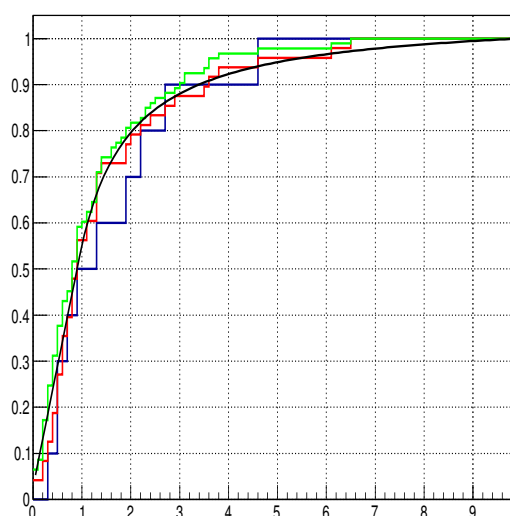
$$P(|F_n(x) - F(x)| \geq \epsilon) \leq \frac{1}{n \epsilon^2} F(x) (1 - F(x)) \quad (360)$$

and therefore

$$\lim_{n \rightarrow \infty} P[|F_n(x) - F(x)| \geq \epsilon] = 0 \quad ; \quad \forall \epsilon > 0 \quad (361)$$

so the empirical Distribution Function  $F_n(x)$  converges in probability to  $F(x)$ . In fact, since

$$\lim_{n \rightarrow \infty} E[|F_n(x) - F(x)|^2] = \lim_{n \rightarrow \infty} \frac{F(x) (1 - F(x))}{n} = 0 \quad (362)$$



**Fig. 5:** Empirical Distribution Function of example 2.32 for sample sizes 10 (blue), 50 (green) and 100 (red) together with the Distribution Function (black).

converges also in quadratic mean and therefore in distribution.

**Example 2.32:** Let  $X = X_1/X_2$  with  $X_i \sim Un(x|0, 1)$ ;  $i = 1, 2$  and Distribution Function

$$F(x) = \frac{x}{2} \mathbf{1}_{(0,1]}(x) + \left(1 - \frac{x}{2}\right) \mathbf{1}_{(1,\infty)}(x) \quad (363)$$

that you can get (exercise) from the Mellin Transform. This is depicted in black in figure ???. There are no moments for this distribution; that is  $E[X^n]$  does not exist for  $n \geq 1$ . We have done Monte Carlo samplings of size  $n = 10, 50$  and  $100$  and the corresponding empirical Distribution Functions

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(-\infty, x]}(x_i) \quad (364)$$

are shown in blue, red and green respectively.

"... some rule could be found, according to which we ought to estimate the chance that the probability for the happening of an event perfectly unknown, should lie between any two named degrees of probability, antecedently to any experiments made about it;..."

*An Essay towards solving a Problem in the Doctrine of Chances.  
By the late Rev. Mr. Bayes ...*

### 3 Bayesian Inference

The goal of *statistical inference* is to get information from experimental observations about quantities (parameters, models,...) on which we want to learn something, be them directly observable or not. Bayesian inference<sup>15</sup> is based on the *Bayes rule* and considers probability as a measure of the the degree of knowledge we have on the quantities of interest. Bayesian methods provide a framework with enough freedom to analyze different models, as complex as needed, using in a natural and conceptually simple way all the information available from the experimental data within a scheme that allows to understand the different steps of the learning process:

<sup>15</sup>For a gentle reading on the subject see [?]

- 1) state the knowledge we have before we do the experiment;
- 2) how the knowledge is modified after the data is taken;
- 3) how to incorporate new experimental results.
- 4) predict what shall we expect in a future experiment from the knowledge acquired.

It was Sir R.A Fisher, one of the greatest statisticians ever, who said that "The Theory of Inverse Probability (that is how Bayesianism was called at the beginning of the XX<sup>th</sup> century) is founded upon an error and must be wholly rejected" although, as time went by, he became a little more acquiescent with Bayesianism. You will see that that Bayesianism is great, rational, coherent, conceptually simple,... "even useful",... and worth to, at least, take a look at it and at the more detailed references on the subject given along the section. At the end, to quote Lindley, "Inside every non-Bayesian there is a Bayesian struggling to get out". For a more classical approach to Statistical Inference see [?] where most of what you will need in Experimental Physics is covered in detail.

### 3.1 Elements of Parametric Inference

Consider an experiment designed to provide information about the set of parameters  $\theta = \{\theta_1, \dots, \theta_k\} \in \Theta \subseteq R^k$  and whose realization results in the random sample  $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ . The inferential process entails:

- 1) Specification of the probabilistic model for the random quantities of interest; that is, state the joint density:

$$p(\theta, \mathbf{x}) = p(\theta_1, \theta_2, \dots, \theta_k, x_1, x_2, \dots, x_n); \quad \theta \in \Theta \subseteq R^k; \quad \mathbf{x} \in X \quad (365)$$

- 2) Conditioning the observed data ( $\mathbf{x}$ ) to the parameters ( $\theta$ ) of the model:

$$p(\theta, \mathbf{x}) = p(\mathbf{x}|\theta) p(\theta) \quad (366)$$

- 3) Last, since  $p(\theta, \mathbf{x}) = p(\mathbf{x}|\theta) p(\theta) = p(\theta|\mathbf{x}) p(\mathbf{x})$  and

$$p(\mathbf{x}) = \int_{\Theta} p(\mathbf{x}, \theta) d\theta = \int_{\Theta} p(\mathbf{x}|\theta) p(\theta) d\theta \quad (367)$$

we have ( *Bayes Rule*) that:

$$p(\theta|\mathbf{x}) = \frac{p(\mathbf{x}|\theta) p(\theta)}{\int_{\Theta} p(\mathbf{x}|\theta) p(\theta) d\theta} \quad (368)$$

This is the basic equation for parametric inference. The integral of the denominator does not depend on the parameters ( $\theta$ ) of interest; is just a normalization factor so we can write in a general way;

$$p(\theta|\mathbf{x}) \propto p(\mathbf{x}|\theta) p(\theta) \quad (369)$$

Let's see these elements in detail:

$p(\theta|\mathbf{x})$  : This is the *Posterior Distribution* that quantifies the knowledge we have on the parameters of interest  $\theta$  conditioned to the observed data  $\mathbf{x}$  (that is, after the experiment has been done) and will allow to perform inferences about the parameters;

$p(\mathbf{x}|\theta)$  : The *Likelihood*; the sampling distribution considered as a function of the parameters  $\theta$  for the *fixed* values (already observed)  $\mathbf{x}$ . Usually, it is written as  $\ell(\theta; \mathbf{x})$  to stress the fact that it is a function of the parameters.

The experimental results modify the prior knowledge we have on the parameters  $\theta$  only through the likelihood so, for the inferential process, we can consider the likelihood function defined up to multiplicative factors provided they do not depend on the parameters.

$p(\theta)$  : This is a *reference function*, independent of the results of the experiment, that quantifies or expresses, in a sense to be discussed later, the knowledge we have on the parameters  $\theta$  *before* the experiment is done. It is termed *Prior Density* although, in many cases, it is an improper function and therefore not a probability density.

### 3.2 Exchangeable sequences

The inferential process to obtain information about a set of parameters  $\theta \in \Theta$  of a model  $X \sim p(x|\theta)$  with  $X \in \Omega_X$  is based on the realization of an experiment  $e(1)$  that provides an observation  $\{x_1\}$ . The  $n$ -fold repetition of the experiment under the same conditions,  $e(n)$ , will provide the random sample  $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$  and this can be considered as a draw of the  $n$ -dimensional random quantity  $\mathbf{X} = (X_1, X_2, \dots, X_n)$  where each  $X_i \sim p(x|\theta)$ .

In Classical Statistics, the inferential process makes extensive use of the idea that the observed sample is originated from a sequence of *independent and identically distributed* (iid) random quantities while Bayesian Inference rests on the less restrictive idea of *exchangeability* [?]. An infinite sequence of random quantities  $\{X_i\}_{i=1}^{\infty}$  is said to be *exchangeable* if *any* finite sub-sequence  $\{X_1, X_2, \dots, X_n\}$  is *exchangeable*; that is, if the joint density  $p(x_1, x_2, \dots, x_n)$  is invariant under *any* permutation of the indices.

The hypothesis of *exchangeability* assumes a symmetry of the experimental observations  $\{x_1, x_2, \dots, x_n\}$  such that the subscripts which identify a particular observation (for instance the order in which they appear) are irrelevant for the inferences. Clearly, if  $\{X_1, X_2, \dots, X_n\}$  are iid then the conditional joint density can be expressed as:

$$p(x_1, x_2, \dots, x_n) = \prod_{i=1}^n p(x_i) \quad (370)$$

and therefore, since the product is invariant to reordering, is an *exchangeable* sequence. The converse is not necessarily true<sup>16</sup> so the hypothesis of exchangeability is weaker than the hypothesis of independence. Now, if  $\{X_i\}_{i=1}^{\infty}$  is an exchangeable sequence of real-valued random quantities it can be shown that, for *any* finite subset, there exists a parameter  $\theta \in \Theta$ , a parametric model  $p(x|\theta)$  and measure  $d\mu(\theta)$  such that<sup>17</sup>:

$$p(x_1, x_2, \dots, x_n) = \int_{\Theta} \prod_{i=1}^n p(x_i|\theta) d\mu(\theta) \quad (371)$$

Thus, any finite sequence of exchangeable observations is described by a model  $p(x|\theta)$  and, if  $d\mu(\theta) = p(\theta)d\theta$ , there is a prior density  $p(\theta)$  that we may consider as describing the available information on the

<sup>16</sup>It is easy to check for instance that if  $X_0$  is a non-trivial random quantity independent of the  $X_i$ , the sequence  $\{X_0 + X_1, X_0 + X_2, \dots, X_0 + X_n\}$  is exchangeable but not iid.

<sup>17</sup>This is referred as *De Finetti's Theorem* after B. de Finetti (1930s) and was generalized by E. Hewitt and L.J. Savage in the 1950s. See [?].



parameter  $\theta$  before the experiment is done. This justifies and, in fact, leads to the Bayesian approach in which, by formally applying Bayes Theorem

$$p(\mathbf{x}, \theta) = p(\mathbf{x}|\theta) p(\theta) = p(\theta|\mathbf{x}) p(\mathbf{x}) \quad (372)$$

we obtain the *posterior density*  $p(\theta|\mathbf{x})$  that accounts for the degree of knowledge we have on the parameter after the experiment has been performed. Note that the random quantities of the exchangeable sequence  $\{X_1, X_2, \dots, X_n\}$  are *conditionally independent given  $\theta$  but not iid* because

$$p(x_j) = \int_{\Theta} p(x_j|\theta) d\mu(\theta) \left( \prod_{i(\neq j)=1}^n \int_{\Omega_X} p(x_i|\theta) dx_i \right) \quad (373)$$

and

$$p(x_1, x_2, \dots, x_n) \neq \prod_{i=1}^n p(x_i) \quad (374)$$

There are situations for which the hypothesis of exchangeability can not be assumed to hold. That is the case, for instance, when the data collected by an experiment depends on the running conditions that may be different for different periods of time, for data provided by two different experiments with different acceptances, selection criteria, efficiencies,... or the same medical treatment when applied to individuals from different environments, sex, ethnic groups,... In these cases, we shall have different *units of observation* and it may be more sound to assume *partial exchangeability* within each unit (data taking periods, detectors, hospitals,...) and design a *hierarchical structure* with parameters that account for the relevant information from each unit analyzing all the data in a more global framework.

**NOTE 5:** Suppose that we have a parametric model  $p_1(x|\theta)$  and the exchangeable sample  $\mathbf{x}_1 = \{x_1, x_2, \dots, x_n\}$  provided by the experiment  $e_1(n)$ . The inferences on the parameters  $\theta$  will be drawn from the posterior density  $p(\theta|\mathbf{x}_1) \propto p_1(\mathbf{x}_1|\theta)p(\theta)$ . Now, we do a second experiment  $e_2(m)$ , statistically independent of the first, that provides the exchangeable sample  $\mathbf{x}_2 = \{x_{n+1}, x_{n+2}, \dots, x_{n+m}\}$  from the model  $p_2(x|\theta)$ . It is sound to take as prior density for this second experiment the posterior of the first including therefore the information that we already have about  $\theta$  so

$$p(\theta|\mathbf{x}_2) \propto p_2(\mathbf{x}_2|\theta)p(\theta|\mathbf{x}_1) \propto p_2(\mathbf{x}_2|\theta)p_1(\mathbf{x}_1|\theta)p(\theta). \quad (375)$$

Being the two experiments statistically independent and their sequences exchangeable, if they have the same sampling distribution  $p(x|\theta)$  we have that  $p_1(\mathbf{x}_1|\theta)p_2(\mathbf{x}_2|\theta) = p(\mathbf{x}|\theta)$  where  $\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2\} = \{x_1, \dots, x_n, x_{n+1}, \dots, x_{n+m}\}$  and therefore  $p(\theta|\mathbf{x}_2) \propto p(\mathbf{x}|\theta)p(\theta)$ . Thus, the knowledge we have on  $\theta$  including the information provided by the experiments  $e_1(n)$  and  $e_2(m)$  is determined by the likelihood function  $p(\mathbf{x}|\theta)$  and, in consequence, under the aforementioned conditions the realization of  $e_1(n)$  first and  $e_2(m)$  after is equivalent, from the inferential point of view, to the realization of the experiment  $e(n+m)$ .

### 3.3 Predictive Inference

Consider the realization of the experiment  $e_1(n)$  that provides the sample  $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$  drawn from the model  $p(x|\theta)$ . Inferences about  $\theta \in \Theta$  are determined by the posterior density

$$p(\theta|\mathbf{x}) \propto p(\mathbf{x}|\theta) \pi(\theta) \quad (376)$$

Now suppose that, under the same model and the same experimental conditions, we think about doing a new independent experiment  $e_2(m)$ . What will be the distribution of the random sample  $\mathbf{y} = \{y_1, y_2, \dots, y_m\}$  not yet observed? Consider the experiment  $e(n+m)$  and the sampling density

$$p(\theta, \mathbf{x}, \mathbf{y}) = p(\mathbf{x}, \mathbf{y}|\theta) \pi(\theta) \quad (377)$$

Since both experiments are independent and iid, we have the joint density

$$p(\mathbf{x}, \mathbf{y}|\theta) = p(\mathbf{x}|\theta) p(\mathbf{y}|\theta) \quad \longrightarrow \quad p(\theta, \mathbf{x}, \mathbf{y}) = p(\mathbf{x}|\theta) p(\mathbf{y}|\theta) \pi(\theta) \quad (378)$$

and integrating the parameter  $\theta \in \Theta$ :

$$p(\mathbf{y}, \mathbf{x}) = p(\mathbf{y}|\mathbf{x}) p(\mathbf{x}) = \int_{\Theta} p(\mathbf{y}|\theta) p(\mathbf{x}|\theta) \pi(\theta) d\theta = p(\mathbf{x}) \int_{\Theta} p(\mathbf{y}|\theta) p(\theta|\mathbf{x}) d\theta \quad (379)$$

Thus, we have that

$$p(\mathbf{y}|\mathbf{x}) = \int_{\Theta} p(\mathbf{y}|\theta) p(\theta|\mathbf{x}) d\theta \quad (380)$$

This is the basic expression for the *predictive inference* and allows us to predict the results  $\mathbf{y}$  of a future experiment from the results  $\mathbf{x}$  observed in a previous experiment within the same parametric model. Note that  $p(\mathbf{y}|\mathbf{x})$  is the density of the quantities not yet observed conditioned to the observed sample. Thus, even though the experiments  $e(\mathbf{y})$  and  $e(\mathbf{x})$  are statistically independent, the realization of the first one ( $e(\mathbf{x})$ ) modifies the knowledge we have on the parameters  $\theta$  of the model and therefore affect the prediction on future experiments for, if we do not consider the results of the first experiment or just don't do it, the predictive distribution for  $e(\mathbf{y})$  would be

$$p(\mathbf{y}) = \int_{\Theta} p(\mathbf{y}|\theta) \pi(\theta) d\theta \quad (381)$$

It is then clear from the expression of *predictive inference* that in practice it is equivalent to consider as prior density for the second experiment the proper density  $\pi(\theta) = p(\theta|\mathbf{x})$ . If the first experiment provides very little information on the parameters, then  $p(\theta|\mathbf{x}) \simeq \pi(\theta)$  and

$$p(\mathbf{y}|\mathbf{x}) \simeq \int_{\Theta} p(\mathbf{y}|\theta) \pi(\theta) d\theta \simeq p(\mathbf{y}) \quad (382)$$

On the other hand, if after the first experiment we know the parameters with high accuracy then, in distributional sense,  $\langle p(\theta|\mathbf{x}), \cdot \rangle \simeq \langle \delta(\theta_0), \cdot \rangle$  and

$$p(\mathbf{y}|\mathbf{x}) \simeq \langle \delta(\theta_0), p(\mathbf{y}|\theta) \rangle = p(\mathbf{y}|\theta_0) \quad (383)$$

### 3.4 Sufficient Statistics

Consider  $m$  random quantities  $\{X_1, X_2, \dots, X_m\}$  that take values in  $\Omega_1 \times \dots \times \Omega_m$  and a random vector

$$\mathbf{T} : \Omega_1 \times \dots \times \Omega_m \longrightarrow \mathcal{R}^{k(m)} \quad (384)$$

whose  $k(m) \leq m$  components are functions of the random quantities  $\{X_i\}_{i=1}^m$ . Given the sample  $\{x_1, x_2, \dots, x_m\}$ , the vector  $\mathbf{t} = \mathbf{t}(x_1, \dots, x_m)$  is a  $k(m)$ -dimensional statistic. The practical interest lies in the existence of statistics that contain all the relevant information about the parameters so we don't have to work with the whole sample and simplify considerably the expressions. Thus, of special relevance are the *sufficient statistics*. Given the model  $p(x_1, x_2, \dots, x_n|\theta)$ , the set of statistics  $\mathbf{t} = \mathbf{t}(x_1, \dots, x_m)$  is *sufficient* for  $\theta$  if, and only if,  $\forall m \geq 1$  and any prior distribution  $\pi(\theta)$  it holds that

$$p(\theta|x_1, x_2, \dots, x_m) = p(\theta|\mathbf{t}) \quad (385)$$

Since the data act in the Bayes formula only through the likelihood, it is clear that to specify the posterior density of  $\theta$  we can consider

$$p(\theta|x_1, x_2, \dots, x_m) = p(\theta|\mathbf{t}) \propto p(\mathbf{t}|\theta) \pi(\theta) \quad (386)$$

and all other aspects of the data but  $\mathbf{t}$  are irrelevant. It is obvious however that  $\mathbf{t} = \{x_1, \dots, x_m\}$  is sufficient and, in principle, gives no simplification in the modeling. For this we should have  $k(m) = \dim(\mathbf{t}) < m$  (*minimal sufficient statistics*) and, in the ideal case, we would like that  $k(m) = k$  does not depend on  $m$ . Except some irregular cases, the only distributions that admit a fixed number of sufficient statistics independently of the sample size (that is,  $k(m) = k < m \forall m$ ) are those that belong to the exponential family.

### Example 3.1:

**1)** Consider the exponential model  $X \sim Ex(x|\theta)$ : and the iid experiment  $e(m)$  that provides the sample  $\mathbf{x} = \{x_1, \dots, x_m\}$ . The likelihood function is:

$$p(\mathbf{x}|\theta) = \theta^m e^{-\theta(x_1 + \dots + x_m)} = \theta^{t_1} e^{-\theta t_2} \quad (387)$$

and therefore we have the sufficient statistic  $\mathbf{t} = (m, \sum_{i=1}^m x_i) : \Omega_1 \times \dots \times \Omega_m \longrightarrow \mathcal{R}^{k(m)=2}$

**2)** Consider the Normal model  $X \sim N(x|\mu, \sigma)$  and the iid experiment  $e(m)$  again with  $\mathbf{x} = \{x_1, \dots, x_m\}$ . The likelihood function is:

$$p(\mathbf{x}|\mu, \sigma) \propto \sigma^{-m} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^m (x_i - \mu)^2 \right\} = \sigma^{-t_1} \exp \left\{ -\frac{1}{2\sigma^2} (t_3 - 2\mu t_2 + \mu^2 t_1) \right\} \quad (388)$$

and  $\mathbf{t} = (m, \sum_{i=1}^m x_i, \sum_{i=1}^m x_i^2) : \Omega_1 \times \dots \times \Omega_m \longrightarrow \mathcal{R}^{k(m)=3}$  a sufficient statistic. Usually we shall consider  $\mathbf{t} = \{m, \bar{x}, s^2\}$  with

$$\bar{x} = \frac{1}{m} \sum_{i=1}^m x_i \quad \text{and} \quad s^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \bar{x})^2 \quad (389)$$

the sample mean and the sample variance. Inferences on the parameters  $\mu$  and  $\sigma$  will depend on  $\mathbf{t}$  and all other aspects of the data are irrelevant.

**3)** Consider the Uniform model  $X \sim Un(x|0, \theta)$  and the iid sampling  $\{x_1, x_2, \dots, x_m\}$ . Then  $\mathbf{t} = (m, \max\{x_i, i = 1, \dots, m\}) : \Omega_1 \times \dots \times \Omega_m \longrightarrow \mathcal{R}^{k(m)=2}$  is a sufficient statistic for  $\theta$ .

## 3.5 Exponential Family

A probability density  $p(x|\theta)$ , with  $x \in \Omega_X$  and  $\theta \in \Theta \subseteq \mathcal{R}^k$  belongs to the *k-parameter exponential family* if it has the form:

$$p(x|\theta) = f(x) g(\theta) \exp \left\{ \sum_{i=1}^k c_i \phi_i(\theta) h_i(x) \right\} \quad (390)$$

with

$$g(\theta)^{-1} = \int_{\Omega_x} f(x) \prod_{i=1}^k \exp \{c_i \phi_i(\theta) h_i(x)\} dx \leq \infty \quad (391)$$

The family is called *regular* if  $\text{supp}\{X\}$  is independent of  $\theta$ ; *irregular* otherwise.

If  $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$  is an exchangeable random sampling from the k-parameter regular exponential family, then

$$p(\mathbf{x}|\theta) = \left[ \prod_{i=1}^n f(x_i) \right] [g(\theta)]^n \exp \left\{ \sum_{i=1}^k c_i \phi_i(\theta) \left( \sum_{j=1}^n h_i(x_j) \right) \right\} \quad (392)$$

and therefore  $\mathbf{t}(\mathbf{x}) = \{n, \sum_{i=1}^n h_1(x_i), \dots, \sum_{i=1}^n h_k(x_i)\}$  will be a set of *sufficient statistics*.

**Example 3.2:** Several distributions of interest, like Poisson and Binomial, belong to the exponential family:

- 1) Poisson  $Po(n|\mu)$ :  $P(n|\mu) = \frac{e^{-\mu} \mu^n}{\Gamma(n+1)} = \frac{e^{-(\mu-n \ln \mu)}}{\Gamma(n+1)}$
- 2) Binomial  $Bi(n|N, \theta)$ :  $P(n|N, \theta) = \binom{N}{n} \theta^n (1-\theta)^{N-n} = \binom{N}{n} e^{n \ln \theta + (N-n) \ln(1-\theta)}$

However, the Cauchy  $Ca(x|\alpha, \beta)$  distribution, for instance, does not because

$$p(x_1, \dots, x_m | \alpha, \beta) \propto \prod_{i=1}^m (1 + \beta(x_i - \alpha)^2)^{-1} = \exp \left\{ \sum_{i=1}^m \log(1 + \beta(x_i - \alpha)^2) \right\} \quad (393)$$

can not be expressed as the exponential family form. In consequence, there are no sufficient *minimal* statistics (in other words  $\mathbf{t} = \{n, x_1, \dots, x_n\}$  is the sufficient statistic) and we will have to work with the whole sample.

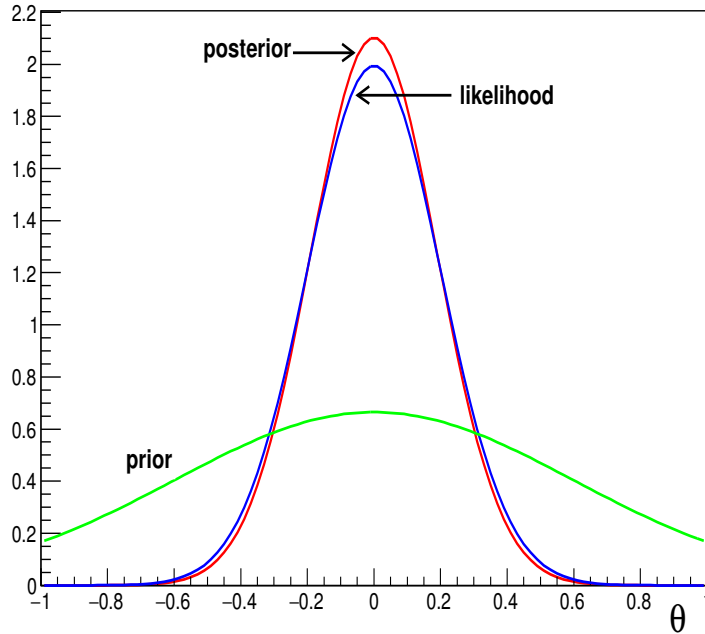
### 3.6 Prior functions

In the *Bayes rule*,  $p(\theta|\mathbf{x}) \propto p(\mathbf{x}|\theta) p(\theta)$ , the *prior function*  $p(\theta)$  represents the knowledge ( *degree of credibility*) that we have about the parameters before the experiment is done and it is a necessary element to obtain the *posterior density*  $p(\theta|\mathbf{x})$  from which we shall make inferences. If we have faithful information on them before we do the experiment, it is reasonable to incorporate that in the specification of the prior density ( *informative prior*) so the new data will provide additional information that will update and improve our knowledge. The specific form of the prior can be motivated, for instance, by the results obtained in previous experiments. However, it is usual that before we do the experiment, either we have a vague knowledge of the parameters compared to what we expect to get from the experiment or simply we do not want to include previous results to perform an independent analysis. In this case, all the new information will be contained in the likelihood function  $p(\mathbf{x}|\theta)$  of the experiment and the prior density ( *non-informative prior*) will be merely a mathematical element needed for the inferential process. Being this the case, we expect that the whole weight of the inferences rests on the likelihood and the prior function has the smallest possible influence on them. To learn something from the experiment it is then desirable to have a situation like the one shown in figure ?? where the posterior distribution  $p(\theta|\mathbf{x})$  is dominated by the likelihood function. Otherwise, the experiment will provide little information compared to the one we had before and, unless our previous knowledge is based on suspicious observations, it will be wise to design a better experiment.

A considerable amount of effort has been put to obtain reasonable *non-informative priors* that can be used as a standard reference function for the Bayes rule. Clearly, *non-informative* is somewhat misleading because we are never in a state of absolute ignorance about the parameters and the specification of a mathematical model for the process assumes some knowledge about them (masses and life-times take non-negative real values, probabilities have support on  $[0, 1], \dots$ ). On the other hand, it doesn't make sense to think about a function that represents ignorance in a formal and objective way so *knowing little a priori* is relative to what we may expect to learn from the experiment. Whatever prior we use will certainly have some effect on the posterior inferences and, in some cases, it would be wise to consider a reasonable set of them to see what is the effect.

The ultimate task of this section is to present the most usual approaches to derive a non-informative prior function to be used as a standard reference that contains little information about the parameters compared to what we expect to get from the experiment<sup>18</sup>. In many cases, these priors will not be Lebesgue integrable ( *improper functions*) and, obviously, can not be considered as probability density functions that quantify any knowledge on the parameters (although, with little rigor, sometimes we still

<sup>18</sup>For a comprehensive discussion see [?]



**Fig. 6:** Prior, likelihood and posterior as function of the the parameter  $\theta$ . In this case, the prior is a smooth function and the posterior is dominated by the likelihood.

talk about prior *densities*). If one is reluctant to use them right the way one can, for instance, define them on a sufficiently large compact support that contains the region where the likelihood is dominant. However, since

$$p(\theta|\mathbf{x}) d\theta \propto p(\mathbf{x}|\theta) p(\theta) d\theta = p(\mathbf{x}|\theta) d\mu(\theta) \quad (394)$$

in most cases it will be sufficient to consider them simply as what they really are: a measure. In any case, what is mandatory is that the posterior is a well defined proper density.

### 3.6.1 Principle of Insufficient Reason

The *Principle of Insufficient Reason*<sup>19</sup> dates back to J. Bernoulli and P.S. Laplace and, originally, it states that if we have  $n$  exclusive and exhaustive hypothesis and there is no special reason to prefer one over the other, it is reasonable to consider them equally likely and assign a prior probability  $1/n$  to each of them. This certainly sounds reasonable and the idea was right the way extended to parameters taking countable possible values and to those with continuous support that, in case of compact sets, becomes a uniform density. It was extensively used by P.S. Laplace and T. Bayes, being he the first to use a uniform prior density for making inferences on the parameter of a Binomial distribution, and is usually referred to as the “*Bayes-Laplace Postulate*”. However, a uniform prior density is obviously not invariant under reparameterizations. If prior to the experiment we have a very vague knowledge about the parameter  $\theta \in [a, b]$ , we certainly have a vague knowledge about  $\phi = 1/\theta$  or  $\zeta = \log\theta$  and a uniform distribution for  $\theta$ :

$$\pi(\theta) d\theta = \frac{1}{b-a} d\theta \quad (395)$$

<sup>19</sup>Apparently, “*Insufficient Reason*” was coined by Laplace in reference to the Leibniz’s *Principle of Sufficient Reason* stating essentially that every fact has a sufficient reason for why it is the way it is and not other way.

implies that:

$$\pi(\phi) d\phi = \frac{1}{\phi^2} d\phi \quad \text{and} \quad \pi(\zeta) d\zeta = e^\zeta d\zeta \quad (396)$$

Shouldn't we take as well a uniform density for  $\phi$  or  $\zeta$ ?

Nevertheless, we shall see that a uniform density, that is far from representing ignorance on a parameter, may be a reasonable choice in many cases even though, if the support of the parameter is infinite, it is an improper function.

### 3.6.2 Parameters of position and scale

An important class of parameters we are interested in are those of position and scale. Let's treat them separately and leave for a forthcoming section the argument behind that. Start with a random quantity  $X \sim p(x|\mu)$  with  $\mu$  a *location parameter*. The density has the form  $p(x|\mu) = f(x - \mu)$  so, taking a prior function  $\pi(\mu)$  we can write

$$p(x, \mu) dx d\mu = [p(x|\mu) dx] [\pi(\mu) d\mu] = [f(x - \mu) dx] [\pi(\mu) d\mu] \quad (397)$$

Now, consider random quantity  $X' = X + a$  with  $a \in \mathcal{R}$  a known value. Defining the new parameter  $\mu' = \mu + a$  we have

$$p(x', \mu') dx' d\mu' = [p(x'|\mu') dx'] [\pi'(\mu') d\mu'] = [f(x' - \mu') dx'] [\pi(\mu' - a) d\mu'] \quad (398)$$

In both cases the models have the same structure so making inferences on  $\mu$  from the sample  $\{x_1, x_1, \dots, x_n\}$  is formally equivalent to making inferences on  $\mu'$  from the shifted sample  $\{x'_1, x'_2, \dots, x'_n\}$ . Since we have the same prior degree of knowledge on  $\mu$  and  $\mu'$ , it is reasonable to take the same functional form for  $\pi(\cdot)$  and  $\pi'(\cdot)$  so:

$$\pi(\mu' - a) d\mu' = \pi(\mu') d\mu' \quad \forall a \in \mathcal{R} \quad (399)$$

and, in consequence:

$$\pi(\mu) = \text{constant} \quad (400)$$

If  $\theta$  is a *scale parameter*, the model has the form  $p(x|\theta) = \theta f(x\theta)$  so taking a prior function  $\pi(\theta)$  we have that

$$p(x, \theta) dx d\theta = [p(x|\theta) dx] [\pi(\theta) d\theta] = [\theta f(x\theta) dx] [\pi(\theta) d\theta] \quad (401)$$

For the scaled random quantity  $X' = aX$  with  $a \in \mathcal{R}^+$  known, we have that:

$$p(x', \theta') dx' d\theta' = [p(x'|\theta') dx'] [\pi'(\theta') d\theta'] = [\theta' f(x'\theta') dx'] [\pi(a\theta') a d\theta'] \quad (402)$$

where we have defined the new parameter  $\theta' = \theta/a$ . Following the same argument as before, it is sound to assume the same functional form for  $\pi(\cdot)$  and  $\pi'(\cdot)$  so:

$$\pi(a\theta') a d\theta' = \pi(\theta') d\theta' \quad \forall a \in \mathcal{R}$$

and, in consequence:

$$\pi(\theta) = \frac{1}{\theta} \quad (403)$$

Both prior functions are improper so they may be explicitated as

$$\pi(\mu, \theta) \propto \frac{1}{\theta} \mathbf{1}_\Theta(\theta) \mathbf{1}_M(\mu) \quad (404)$$

with  $\Theta, M$  an appropriate sequence of compact sets or considered as prior measures provided that the posterior densities are well defined. Let's see some examples.

**Example 3.3: The Exponential Distribution.** Consider the sequence of independent observations  $\{x_1, x_2, \dots, x_n\}$  of the random quantity  $X \sim Ex(x|\theta)$  drawn under the same conditions. The joint density is

$$p(x_1, x_2, \dots, x_n|\theta) = \theta^n e^{-\theta(x_1 + x_2 + \dots + x_n)} \quad (405)$$

The statistic  $t = n^{-1} \sum_{i=1}^n x_i$  is sufficient for  $\theta$  and is distributed as

$$p(t|\theta) = \frac{(n\theta)^n}{\Gamma(n)} t^{n-1} \exp\{-n\theta t\} \quad (406)$$

It is clear that  $\theta$  is a scale parameter so we shall take the prior function  $\pi(\theta) = 1/\theta$ . Note that if we make the change  $z = \log t$  and  $\phi = \log \theta$  we have that

$$p(z|\phi) = \frac{n^n}{\Gamma(n)} \exp\{n((\phi + z) - e^{\phi+z})\} \quad (407)$$

In this parameterization,  $\phi$  is a position parameter and therefore  $\pi(\phi) = \text{const}$  in consistency with  $\pi(\theta)$ . Then, we have the proper posterior for inferences:

$$p(\theta|t, n) = \frac{(nt)^n}{\Gamma(n)} \exp\{-nt\theta\} \theta^{n-1} \quad ; \quad \theta > 0 \quad (408)$$

Consider now the sequence of compact sets  $C_k = [1/k, k]$  covering  $R^+$  as  $k \rightarrow \infty$ . Then, with support on  $C_k$  we have the proper prior density

$$\pi_k(\theta) = \frac{1}{2 \log k} \frac{1}{\theta} \mathbf{1}_{C_k}(\theta) \quad (409)$$

and the sequence of posteriors:

$$p_k(\theta|t, n) = \frac{(nt)^n}{\gamma(n, ntk) - \gamma(n, nt/k)} \exp\{-nt\theta\} \theta^{n-1} \mathbf{1}_{C_k}(\theta) \quad (410)$$

with  $\gamma(a, x)$  the Incomplete Gamma Function. It is clear that

$$\lim_{k \rightarrow \infty} p_k(\theta|t, n) = p(\theta|t, n) \quad (411)$$

**Example 3.4: The Uniform Distribution.** Consider the random quantity  $X \sim Un(x|0, \theta)$  and the independent sampling  $\{x_1, x_2, \dots, x_n\}$ . To draw inferences on  $\theta$ , the statistics  $x_M = \max\{x_1, x_2, \dots, x_n\}$  is sufficient and is distributed as (show that):

$$p(x_M|\theta) = n \frac{x_M^{n-1}}{\theta^n} \mathbf{1}_{[0, \theta]}(x_M) \quad (412)$$

As in the previous case,  $\theta$  is a scale parameter and with the change  $t_M = \log x_M$ ,  $\phi = \log \theta$  is a position parameter. Then, we shall take  $\pi(\theta) \propto \theta^{-1}$  and get the posterior density (Pareto):

$$p(\theta|x_M, n) = n \frac{x_M^n}{\theta^{n+1}} \mathbf{1}_{[x_M, \infty)}(\theta) \quad (413)$$

**Example 3.5: The one-dimensional Normal Distribution.** Consider the random quantity  $X \sim N(x|\mu, \sigma)$  and the experiment  $e(n)$  that provides the independent and exchangeable sequence  $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$  of observations. The likelihood function will then be:

$$p(\mathbf{x}|\mu, \sigma) = \prod_{i=1}^n p(x_i|\mu, \sigma) \propto \frac{1}{\sigma^n} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right\} \quad (414)$$

There is a three-dimensional sufficient statistic  $\mathbf{t} = \{n, \bar{x}, s^2\}$  where

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{and} \quad s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (415)$$

so we can write

$$p(\mathbf{x}|\mu, \sigma) \propto \frac{1}{\sigma^n} \exp\left\{-\frac{n}{2\sigma^2} (s^2 + (\bar{x} - \mu)^2)\right\} \quad (416)$$

In this case we have both position and scale parameters so we take  $\pi(\mu, \sigma) = \pi(\mu)\pi(\sigma) = \sigma^{-1}$  and get the proper posterior

$$p(\mu, \sigma|\mathbf{x}) \propto p(\mathbf{x}|\mu, \sigma) \pi(\mu, \sigma) \propto \frac{1}{\sigma^{n+1}} \exp\left\{-\frac{n}{2\sigma^2} [s^2 + (\bar{x} - \mu)^2]\right\} \quad (417)$$

• **Marginal posterior density of  $\sigma$ :** Integrating the parameter  $\mu \in \mathcal{R}$  we have that:

$$p(\sigma|\mathbf{x}) = \int_{-\infty}^{+\infty} p(\mu, \sigma|\mathbf{x}) d\mu \propto \sigma^{-n} \exp\left\{-\frac{n s^2}{2\sigma^2}\right\} \mathbf{1}_{(0, \infty)}(\sigma) \quad (418)$$

and therefore, the random quantity

$$Z = \frac{n s^2}{\sigma^2} \sim \chi^2(z|n-1) \quad (419)$$

• **Marginal posterior density of  $\mu$ :** Integrating the parameter  $\sigma \in [0, \infty)$  we have that:

$$p(\mu|\mathbf{x}) = \int_0^{+\infty} p(\mu, \sigma|\mathbf{x}) d\sigma \propto \left(1 + \frac{(\mu - \bar{x})^2}{s^2}\right)^{-n/2} \mathbf{1}_{(-\infty, \infty)}(\mu) \quad (420)$$

so the random quantity

$$T = \frac{\sqrt{n-1}(\mu - \bar{x})}{s} \sim St(t|n-1) \quad (421)$$

It is clear that  $p(\mu, \sigma|\mathbf{x}) \neq p(\mu|\mathbf{x})p(\sigma|\mathbf{x})$  and, in consequence, are not independent.

• **Distribution of  $\mu$  conditioned to  $\sigma$ :** Since  $p(\mu, \sigma|\mathbf{x}) = p(\mu|\sigma, \mathbf{x})p(\sigma|\mathbf{x})$  we have that

$$p(\mu|\sigma, \mathbf{x}) \propto \frac{1}{\sigma} \exp\left\{-\frac{n}{2\sigma^2} (\mu - \bar{x})^2\right\} \quad (422)$$

so  $\mu|\sigma \sim N(\mu|\bar{x}, \sigma/\sqrt{n})$ .

**Example 3.6: Contrast of parameters of Normal Densities.** Consider two independent random quantities  $X_1 \sim N(x_1, |\mu_1, \sigma_1)$  and  $X_2 \sim N(x_2, |\mu_2, \sigma_2)$  and the random samplings  $\mathbf{x}_1 = \{x_{11}, x_{12}, \dots, x_{1n_1}\}$  and  $\mathbf{x}_2 = \{x_{21}, x_{22}, \dots, x_{2n_2}\}$  of sizes  $n_1$  and  $n_2$  under the usual conditions. From the considerations of the previous example, we can write

$$p(\mathbf{x}_i|\mu_i, \sigma_i) \propto \frac{1}{\sigma_i^{n_i}} \exp\left\{-\frac{n_i}{2\sigma_i^2} (s_i^2 + (\bar{x}_i - \mu_i)^2)\right\}; \quad i = 1, 2 \quad (423)$$

Clearly,  $(\mu_1, \mu_2)$  are position parameters and  $(\sigma_1, \sigma_2)$  scale parameters so, in principle, we shall take the improper prior function

$$\pi(\mu_1, \sigma_1, \mu_2, \sigma_2) = \pi(\mu_1)\pi(\mu_2)\pi(\sigma_1)\pi(\sigma_2) \propto \frac{1}{\sigma_1 \sigma_2} \quad (424)$$



However, if we have know that both distributions have the same variance, then we may set  $\sigma = \sigma_1 = \sigma_2$  and, in this case, the prior function will be

$$\pi(\mu_1, \mu_2, \sigma) = \pi(\mu_1)\pi(\mu_2)\pi(\sigma) \propto \frac{1}{\sigma} \quad (425)$$

Let's analyze both cases.

• **Marginal Distribution of  $\sigma_1$  and  $\sigma_2$ :** In this case we assume that  $\sigma_1 \neq \sigma_2$  and we shall take the prior  $\pi(\mu_1, \sigma_1, \mu_2, \sigma_2) \propto (\sigma_1 \sigma_2)^{-1}$ . Integrating  $\mu_1$  and  $\mu_2$  we get:

$$p(\sigma_1, \sigma_2 | \mathbf{x}_1, \mathbf{x}_2) = p(\sigma_1 | \mathbf{x}_1) p(\sigma_2 | \mathbf{x}_2) \propto \sigma_1^{-n_1} \sigma_2^{-n_2} \exp \left\{ -\frac{1}{2} \left( \frac{n_1 s_1^2}{\sigma_1^2} + \frac{n_2 s_2^2}{\sigma_2^2} \right) \right\} \quad (426)$$

Now, if we define the new random quantities

$$Z = \frac{s_2^2}{w^2 s_1^2} = \frac{(\sigma_1/s_1)^2}{(\sigma_2/s_2)^2} \quad \text{and} \quad W = \frac{n_1 s_1^2}{\sigma_1^2} \quad (427)$$

both with support in  $(0, +\infty)$ , and integrate the last we get we get that  $Z$  follows a Snedecor Distribution  $Sn(z | n_2 - 1, n_1 - 1)$  whose density is

$$p(z | \mathbf{x}_1, \mathbf{x}_2) = \frac{(\nu_1/\nu_2)^{\nu_1/2}}{\text{Be}(\nu_1/2, \nu_2/2)} z^{(\nu_1/2)-1} \left( 1 + \frac{\nu_1}{\nu_2} z \right)^{-(\nu_1+\nu_2)/2} \mathbf{1}_{(0, \infty)}(z) \quad (428)$$

• **Marginal Distribution of  $\mu_1$  and  $\mu_2$ :** In this case, it is different whether we assume that, although unknown, the variances are the same or not. In the first case, we set  $\sigma_1 = \sigma_2 = \sigma$  and take the reference prior  $\pi(\mu_1, \mu_2, \sigma) = \sigma^{-1}$ . Defining

$$A = n_1 [s_1^2 + (\bar{x}_1 - \mu_1)^2] + n_2 [s_2^2 + (\bar{x}_2 - \mu_2)^2] \quad (429)$$

we can write

$$p(\mu_1, \mu_2, \sigma | \mathbf{x}, \mathbf{y}) \propto \frac{1}{\sigma^{n_1+n_2+1}} \exp \left\{ -\frac{1}{2} A/\sigma^2 \right\} \quad (430)$$

It is left as an exercise to show that if we make the transformation

$$w = \mu_1 - \mu_2 \in (-\infty, +\infty) \quad ; \quad u = \mu_2 \in (-\infty, +\infty) \quad \text{and} \quad z = \sigma^{-2} \in (0, +\infty) \quad (431)$$

and integrate the last two, we get

$$p(w | \mathbf{x}_1, \mathbf{x}_2) \propto \left( 1 + \frac{n_1 n_2}{n_1 + n_2} \frac{[(\bar{x}_1 - \bar{x}_2) - w]^2}{n_1 s_1^2 + n_2 s_2^2} \right)^{-(n_1+n_2-1)/2} \quad (432)$$

Introducing the more usual terminology

$$s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2} \quad (433)$$

we have that

$$p(w | \mathbf{x}_1, \mathbf{x}_2) \propto \left( 1 + \frac{n_1 n_2}{n_1 + n_2} \frac{[w - (\bar{x}_1 - \bar{x}_2)]^2}{s^2 (n_1 + n_2 - 2)} \right)^{-[(n_1+n_2-2)+1]/2} \quad (434)$$

and therefore the random quantity

$$T = \frac{(\mu_1 - \mu_2) - (\bar{x}_1 - \bar{x}_2)}{s(1/n_1 + 1/n_2)^{1/2}} \quad (435)$$

follows a Student's Distribution  $St(t|\nu)$  with  $\nu = n_1 + n_2 - 2$  degrees of freedom.

Let's see now the case where we can not assume that the variances are equal. Taking the prior reference function  $\pi(\mu_1, \mu_2, \sigma_1, \sigma_2) = (\sigma_1 \sigma_2)^{-1}$  we get

$$p(\mu_1, \mu_2, \sigma_1, \sigma_2 | \mathbf{x}_1, \mathbf{x}_2) \propto \sigma_1^{-(n_1+1)} \sigma_2^{-(n_2+1)} \exp \left\{ -\frac{1}{2} \sum_{i=1}^2 \frac{s_i^2 + (\bar{x}_i - \mu_i)^2}{\sigma_i^2/n_i} \right\} \quad (436)$$

After the appropriate integrations (left as exercise), defining  $w = \mu_1 - \mu_2$  and  $u = \mu_2$  we end up with the density

$$p(w, u | \mathbf{x}_1, \mathbf{x}_2) \propto \left( 1 + \frac{(\bar{x}_1 - w - u)^2}{s_1^2} \right)^{-n_1/2} \left( 1 + \frac{(\bar{x}_2 - u)^2}{s_2^2} \right)^{-n_2/2} \quad (437)$$

where integral over  $u \in \mathcal{R}$  can not be expressed in a simple way. The density

$$p(w | \mathbf{x}_1, \mathbf{x}_2) \propto \int_{-\infty}^{+\infty} p(w, u | \mathbf{x}_1, \mathbf{x}_2) du \quad (438)$$

is called the *Behrens-Fisher Distribution*. Thus, to make statements on the difference of Normal means, we should analyze first the sample variances and decide how shall we treat them.

### 3.6.3 Covariance under reparameterizations

The question of how to establish a reasonable criteria to obtain a prior for a given model  $p(\mathbf{x}|\theta)$  that can be used as a standard reference function was studied by Harold Jeffreys [?] in the mid XX<sup>th</sup> century. The rationale behind the argument is that if we have the model  $p(\mathbf{x}|\theta)$  with  $\theta \in \Omega_\theta \subseteq R^n$  and make a reparameterizations  $\phi = \phi(\theta)$  with  $\phi(\cdot)$  a one-to-one differentiable function, the statements we make about  $\theta$  should be consistent with those we make about  $\phi$  and, in consequence, priors should be related by

$$\pi_\theta(\theta) d\theta = \pi_\phi(\phi(\theta)) \left| \det \left[ \frac{\partial \phi_i(\theta)}{\partial \theta_j} \right] \right| d\theta \quad (439)$$

Now, assume that the Fisher's matrix

$$\mathbf{I}_{ij}(\theta) = E_{\mathbf{X}} \left[ \frac{\partial \log p(\mathbf{x}|\theta)}{\partial \theta_i} \frac{\partial \log p(\mathbf{x}|\theta)}{\partial \theta_j} \right] \quad (440)$$

exists for this model. Under a differentiable one-to-one transformation  $\phi = \phi(\theta)$  we have that

$$\mathbf{I}_{ij}(\phi) = \frac{\partial \theta_k}{\partial \phi_i} \frac{\partial \theta_l}{\partial \phi_j} \mathbf{I}_{kl}(\theta) \quad (441)$$

so it behaves as a covariant symmetric tensor of second order (left as exercise). Then, since

$$\det [\mathbf{I}(\phi)] = \left| \det \left[ \frac{\partial \theta_i}{\partial \phi_j} \right] \right|^2 \det [\mathbf{I}(\theta)] \quad (442)$$

Jeffreys proposed to consider the prior

$$\pi(\theta) \propto [\det[\mathbf{I}(\theta)]]^{1/2} \quad (443)$$

In fact, if we consider the parameter space as a Riemannian manifold ([?]) the Fisher's matrix is the metric tensor (Fisher-Rao metric) and this is just the invariant volume element. Intuitively, if we make a transformation such that at a particular value  $\phi_0 = \phi(\theta_0)$  the Fisher's tensor is constant and diagonal,

the metric in a neighborhood of  $\phi_0$  is Euclidean and we have location parameters for which a constant prior is appropriate and therefore

$$\pi(\phi)d\phi \propto d\phi = \left[ \det [\mathbf{I}(\theta)]^{1/2} \right] d\theta = \pi(\theta)d\theta \quad (444)$$

It should be pointed out that there may be other priors that are also invariant under reparameterizations and that, as usual, we talk loosely about *prior densities* although they usually are improper functions.

For one-dimensional parameter, the density function expressed in terms of

$$\phi \sim \int [\mathbf{I}(\theta)]^{1/2} d\theta \quad (445)$$

may be reasonably well approximated by a Normal density (at least in the parametric region where the likelihood is dominant) because  $\mathbf{I}(\phi)$  is constant and then, due to translation invariance, a constant prior for  $\phi$  is justified. Let's see some examples.

**Example 3.7: The Binomial Distribution.** Consider the random quantity  $X \sim Bi(x|\theta, n)$ :

$$p(x|n, \theta) = \binom{n}{x} \theta^x (1 - \theta)^{n-x} \quad ; \quad n, x \in N_0; x \leq n \quad (446)$$

with  $0 < \theta < 1$ . Since  $E[X] = n\theta$  we have that:

$$\mathbf{I}(\theta) = E_X \left[ \left( - \frac{\partial^2 \log p(x|n, \theta)}{\partial \theta^2} \right) \right] = \frac{n}{\theta(1 - \theta)} \quad (447)$$

so the Jeffreys prior (proper in this case) for the parameter  $\theta$  is

$$\pi(\theta) \propto [\theta(1 - \theta)]^{-1/2} \quad (448)$$

and the posterior density will therefore be

$$p(\theta|x, n) \propto \theta^{x-1/2} (1 - \theta)^{n-x-1/2} \quad (449)$$

that is; a  $Be(x|k + 1/2, n - k + 1/2)$  distribution. Since

$$\phi = \int \frac{d\theta}{\sqrt{\theta(1 - \theta)}} = 2 \operatorname{asin}(\theta^{1/2}) \quad (450)$$

we have that  $\theta = \sin^2 \phi/2$  and, parameterized in terms of  $\phi$ ,  $\mathbf{I}(\phi)$  is constant so the distribution “looks” more Normal (see figure ??).

**Example 3.8: The Poisson Distribution.** Consider the random quantity  $X \sim Po(x|\mu)$ :

$$p(x|\mu) = e^{-\mu} \frac{\mu^x}{\Gamma(x + 1)} \quad ; \quad x \in N; \mu \in R^+ \quad (451)$$

Then, since  $E[X] = \mu$  we have

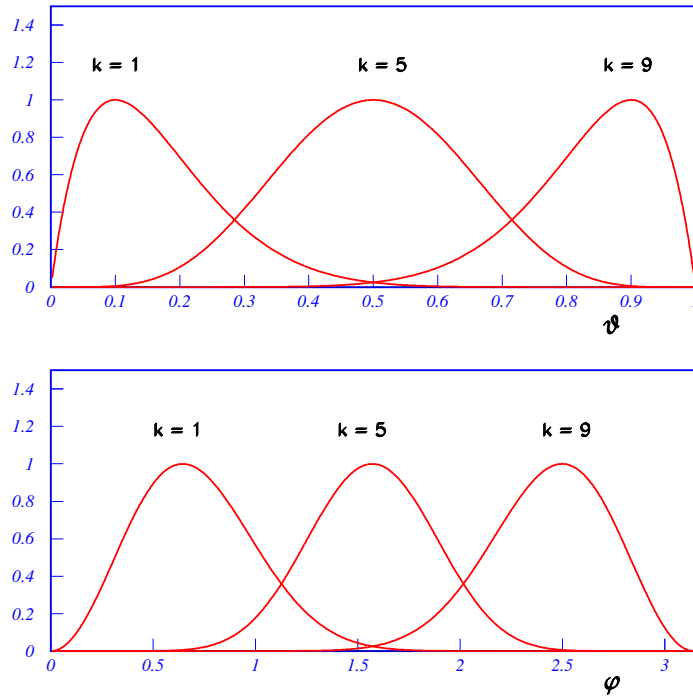
$$\mathbf{I}(\mu) = E_X \left[ \left( - \frac{\partial^2 \log p(x|\mu)}{\partial \mu^2} \right) \right] = \frac{1}{\mu} \quad (452)$$

so we shall take as *prior* (improper):

$$\pi(\mu) = [\mathbf{I}(\mu)]^{1/2} = \mu^{-1/2} \quad (453)$$

and make inferences on  $\mu$  from the proper posterior density

$$p(\mu|x) \propto e^{-\mu} \mu^{x-1/2} \quad (454)$$



**Fig. 7:** Dependence of the likelihood function with the parameter  $\theta$  (upper) and with  $\phi = 2 \operatorname{asin}(\theta^{1/2})$  (lower) for a Binomial process with  $n = 10$  and  $k = 1, 5$  and  $9$ .

that is, a  $Ga(x|1, x + 1/2)$  distribution.

**Example 3.9: The Pareto Distribution.** Consider the random quantity  $X \sim Pa(x|\theta, x_0)$  with  $x_0 \in R^+$  known and density

$$p(x|\theta, x_0) = \frac{\theta}{x_0} \left(\frac{x_0}{x}\right)^{\theta+1} \mathbf{1}_{(x_0, \infty)}(x) \quad ; \quad \theta \in R^+ \tag{455}$$

Then,

$$\mathbf{I}(\theta) = E_X \left[ \left( -\frac{\partial^2 \log p(x|\theta, x_0)}{\partial \theta^2} \right) \right] = \frac{1}{\theta^2} \tag{456}$$

so we shall take as *prior* (improper):

$$\pi(\theta) \propto [\mathbf{I}(\mu)]^{1/2} = \theta^{-1} \tag{457}$$

and make inferences from the posterior density (proper)

$$p(\theta|x, x_0) = x^{-\theta} \log x \tag{458}$$

Note that if we make the transformation  $t = \log x$ , the density becomes

$$p(t|\theta, x_0) = \theta x_0^\theta e^{-\theta t} \mathbf{1}_{(\log x_0, \infty)}(t) \tag{459}$$

for which  $\theta$  is a scale parameter and, from previous considerations, we should take  $\pi(\theta) \propto \theta^{-1}$  in consistency with Jeffreys's prior.

**Example 3.10: The Gamma Distribution.** Consider the random quantity  $X \sim Ga(x|\alpha, \beta)$  with  $\alpha, \beta \in \mathbb{R}^+$  and density

$$p(x|\alpha, \beta) = \frac{\alpha^\beta}{\Gamma(\beta)} e^{-\alpha x} x^{\beta-1} \mathbf{1}_{(0, \infty)}(x) \quad (460)$$

Show that the Fisher's matrix is

$$\mathbf{I}(\alpha, \beta) = \begin{pmatrix} \beta\alpha^{-2} & -\alpha^{-1} \\ -\alpha^{-1} & \Psi'(\beta) \end{pmatrix} \quad (461)$$

with  $\Psi'(x)$  the first derivative of the Digamma Function and, following Jeffreys' rule, we should take the prior

$$\pi(\alpha, \beta) \propto \alpha^{-1} [\beta\Psi'(\beta) - 1]^{1/2} \quad (462)$$

Note that  $\alpha$  is a scale parameter so, from previous considerations, we should take  $\pi(\alpha) \propto \alpha^{-1}$ . Furthermore, if we consider  $\alpha$  and  $\beta$  independently, we shall get

$$\pi(\alpha, \beta) = \pi(\alpha)\pi(\beta) \propto \alpha^{-1} [\Psi'(\beta)]^{1/2} \quad (463)$$

**Example 3.11: The Beta Distribution.**

Show that for the  $Be(x|\alpha, \beta)$  distribution with density

$$p(x|\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} (x^{\alpha-1} (1-x)^{\beta-1}) \mathbf{1}_{[0,1]}(x) \quad ; \quad \alpha, \beta \in \mathbb{R}^+ \quad (464)$$

the Fisher's matrix is given by

$$\mathbf{I}(\alpha, \beta) = \begin{pmatrix} \Psi'(\alpha) - \Psi'(\alpha + \beta) & -\Psi'(\alpha + \beta) \\ -\Psi'(\alpha + \beta) & \Psi'(\beta) - \Psi'(\alpha + \beta) \end{pmatrix} \quad (465)$$

with  $\Psi'(x)$  the first derivative of the Digamma Function.

**Example 3.12: The Normal Distribution.**

**Univariate:** The Fisher's matrix is given by

$$\mathbf{I}(\mu, \sigma) = \begin{pmatrix} \sigma^{-2} & 0 \\ 0 & 2\sigma^{-2} \end{pmatrix} \quad (466)$$

so

$$\pi_1(\mu, \sigma) \propto [\det[\mathbf{I}(\mu, \sigma)]]^{1/2} \propto \frac{1}{\sigma^2} \quad (467)$$

However, had we treated the two parameters independently, we should have obtained

$$\pi_2(\mu, \sigma) = \pi(\mu)\pi(\sigma) \propto \frac{1}{\sigma} \quad (468)$$

The prior  $\pi_2 \propto \sigma^{-1}$  is the one we had used in example 3.5 where the problem was treated as two one-dimensional independent problems and, as we saw:

$$T = \frac{\sqrt{n-1}(\mu - \bar{x})}{s} \sim St(t|n-1) \quad \text{and} \quad Z = \frac{n s^2}{\sigma^2} \sim \chi^2(z|n-1) \quad (469)$$

with  $E[Z] = n-1$ . Had we used prior  $\pi_1 \propto \sigma^{-2}$ , we would have obtained that  $Z \sim \chi^2(z|n)$  and therefore  $E[Z] = n$ . This is not reasonable. On the one hand, we know from the sampling distribution  $N(x|\mu, \sigma)$  that  $E[n s^2 \sigma^{-2}] = n-1$ . On the other hand, we have two parameters  $(\mu, \sigma)$  and integrate on one  $(\sigma)$  so the number of degrees of freedom should be  $n-1$ .

**Bivariate:** The Fisher's matrix is given by

$$\mathbf{I}(\mu_1, \mu_2) = (1 - \rho^2)^{-1} \begin{pmatrix} \sigma_1^{-2} & -\rho(\sigma_1\sigma_2)^{-1} \\ -\rho(\sigma_1\sigma_2)^{-1} & \sigma_2^{-2} \end{pmatrix} \quad (470)$$

$$\mathbf{I}(\sigma_1, \sigma_2, \rho) = (1 - \rho^2)^{-1} \begin{pmatrix} (2 - \rho^2)\sigma_1^{-2} & -\rho^2(\sigma_1\sigma_2)^{-1} & -\rho\sigma_1^{-1} \\ -\rho^2(\sigma_1\sigma_2)^{-1} & (2 - \rho^2)\sigma_2^{-2} & -\rho\sigma_2^{-1} \\ -\rho\sigma_1^{-1} & -\rho\sigma_2^{-1} & (1 + \rho^2)(1 - \rho^2)^{-1} \end{pmatrix} \quad (471)$$

$$\mathbf{I}(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho) = \begin{pmatrix} \mathbf{I}(\mu_1, \mu_2) & \mathbf{0} \\ \mathbf{0} & \mathbf{I}(\sigma_1, \sigma_2, \rho) \end{pmatrix} \quad (472)$$

Form this,

$$\pi(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho) \propto |\det \mathbf{I}(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)|^{1/2} = \frac{1}{\sigma_1^2 \sigma_2^2 (1 - \rho^2)^2} \quad (473)$$

while if we consider  $\pi(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho) = \pi(\mu_1, \mu_2)\pi(\sigma_1, \sigma_2, \rho)$  we get

$$\pi(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho) \propto \frac{1}{\sigma_1 \sigma_2 (1 - \rho^2)^{3/2}} \quad (474)$$

**Problem 3.1:** Show that for the density  $p(\mathbf{x}|\theta)$ ;  $\mathbf{x} \in \Omega \subseteq R^n$ , the Fisher's matrix (if exists)

$$\mathbf{I}_{ij}(\theta) = E_{\mathbf{X}} \left[ \frac{\partial \log p(\mathbf{x}|\theta)}{\partial \theta_i} \frac{\partial \log p(\mathbf{x}|\theta)}{\partial \theta_j} \right] \quad (475)$$

transforms under a differentiable one-to-one transformation  $\phi = \phi(\theta)$  as a covariant symmetric tensor of second order; that is

$$\mathbf{I}_{ij}(\phi) = \frac{\partial \theta_k}{\partial \phi_i} \frac{\partial \theta_l}{\partial \phi_j} \mathbf{I}_{kl}(\theta) \quad (476)$$

**Problem 3.2:** Show that for  $X \sim Po(x|\mu + b)$  with  $b \in R^+$  known (Poisson model with known background), we have that  $\mathbf{I}(\mu) = (\mu + b)^{-1}$  and therefore the posterior (proper) is given by:

$$p(\mu|x, b) \propto e^{-(\mu+b)} (\mu + b)^{x-1/2} \quad (477)$$

**Problem 3.3:** Show that for the one parameter mixture model  $p(x|\lambda) = \lambda p_1(x) + (1 - \lambda)p_2(x)$  with  $p_1(x) \neq p_2(x)$  properly normalized and  $\lambda \in (0, 1)$ ,

$$\mathbf{I}(\lambda) = \frac{1}{\lambda(1 - \lambda)} \left\{ 1 - \int_{-\infty}^{\infty} \frac{p_1(x)p_2(x)}{p(x|\lambda)} dx \right\} \quad (478)$$

When  $p_1(x)$  and  $p_2(x)$  are "well separated", the integral is  $\ll 1$  and therefore  $\mathbf{I}(\lambda) \sim [\lambda(1 - \lambda)]^{-1}$ . On the other hand, when they "get closer" we can write  $p_2(x) = p_1(x) + \eta(x)$  with  $\int_{-\infty}^{\infty} \eta(x) dx = 0$  and, after a Taylor expansion for  $|\eta(x)| \ll 1$  get to first order that

$$\mathbf{I}(\lambda) \simeq \int_{-\infty}^{\infty} \frac{(p_1(x) - p_2(x))^2}{p_1(x)} dx + \dots \quad (479)$$

independent of  $\lambda$ . Thus, for this problem it will be sound to consider the prior  $\pi(\lambda|a, b) = Be(\lambda|a, b)$  with parameters between  $(1/2, 1/2)$  and  $(1, 1)$ .

**3.6.4 Invariance under a Group of Transformations**

Some times, we may be interested to provide the prior with invariance under some transformations of the parameters (or a subset of them) considered of interest for the problem at hand. As we have stated, from a formal point of view the prior can be treated as an absolute continuous measure with respect to Lebesgue so  $p(\theta|\mathbf{x}) d\theta \propto p(\mathbf{x}|\theta) \pi(\theta) d\theta = p(\mathbf{x}|\theta) d\mu(\theta)$ . Now, consider the probability space  $(\Omega, B, \mu)$  and a measurable homeomorphism  $T : \Omega \rightarrow \Omega$ . A measure  $\mu$  on the Borel algebra  $B$  would be invariant by the mapping  $T$  if for any  $A \subset B$ , we have that  $\mu(T^{-1}(A)) = \mu(A)$ . We know, for instance, that there is a unique measure  $\lambda$  on  $R^n$  that is invariant under translations and such that for the unit cube  $\lambda([0, 1]^n) = 1$ : the Lebesgue measure (in fact, it could have been defined that way). This is consistent with the constant prior specified already for position parameters. The Lebesgue measure is also the unique measure in  $R^n$  that is invariant under the rotation group  $SO(n)$  (see problem 3.5). Thus, when expressed in spherical polar coordinates, it would be reasonable for the spherical surface  $S^{n-1}$  the rotation invariant prior

$$d\mu(\phi) = \prod_{k=1}^{n-1} (\sin \phi_k)^{(n-1)-k} d\phi_k \tag{480}$$

with  $\phi_{n-1} \in [0, 2\pi)$  and  $\phi_j \in [0, \pi]$  for the rest. We shall use this prior function in a later problem.

In other cases, the group of invariance is suggested by the model

$$M : \{p(\mathbf{x}|\theta), \mathbf{x} \in \Omega_X, \theta \in \Omega_\Theta\} \tag{481}$$

in the sense that we can make a transformation of the random quantity  $\mathbf{X} \rightarrow \mathbf{X}'$  and absorb the change in a redefinition of the parameters  $\theta \rightarrow \theta'$  such that the expression of the probability density remains unchanged. Consider a group of transformations <sup>20</sup>  $G$  that acts

$$\text{on the Sample Space: } x \rightarrow x' = g \circ x; \quad g \in G; x, x' \in \Omega_X$$

$$\text{on the Parametric Space: } \theta \rightarrow \theta' = g \circ \theta; \quad g \in G; \theta, \theta' \in \Omega_\Theta$$

The model  $M$  is said to be invariant under  $G$  if  $\forall g \in G$  and  $\forall \theta \in \Omega_\Theta$  the random quantity  $X' = g \circ X$  is distributed as  $p(x'|\theta') \equiv p(g \circ x | g \circ \theta)$ . Therefore, transformations of data under  $G$  will make no difference on the inferences if we assign consistent ‘‘prior beliefs’’ to the original and transformed parameters. Note that the action of the group on the sample and parameter spaces will, in general, be different. The essential point is that, as Alfred Haar showed in 1933, for the action of the group  $G$  of transformations there is an invariant measure  $\mu$  (*Haar measure*; [?]) such that

$$\int_{\Omega_X} f(g \circ x) d\mu(x) = \int_{\Omega_X} f(x') d\mu(x') \tag{482}$$

for any Lebesgue integrable function  $f(x)$  on  $\Omega_X$ . Shortly after, it was shown (Von Neumann (1934); Weil and Cartan (1940)) that this measure is unique up to a multiplicative constant. In our case, the function will be  $p(\cdot|\theta) 1_\Theta(\theta)$  and the invariant measure we are looking for is  $d\mu(\theta) \propto \pi(\theta) d\theta$ . Furthermore, since the group may be non-abelian, we shall consider the action on the right and on the left of the parameter space. Thus, we shall have:

$$\int_{\Theta} p(\cdot | g \circ \theta) \pi_L(\theta) d\theta = \int_{\Theta} p(\cdot | \theta') \pi_L(\theta') d\theta' \tag{483}$$

if the group acts on the left and

$$\int_{\Theta} p(\cdot | \theta \circ g) \pi_R(\theta) d\theta = \int_{\Theta} p(\cdot | \theta') \pi_R(\theta') d\theta' \tag{484}$$

---

<sup>20</sup>In this context, the use of Transformation Groups arguments was pioneered by E.T. Jaynes [?].

if the action is on the right. Then, we should start by identifying the group of transformations under which the model is invariant (if any; in many cases, either there is no invariance or at least not obvious) work in the parameter space. The most interesting cases for us are:

$$\text{Affine Transformations: } x \rightarrow x' = g \circ x = a + b x$$

$$\text{Matrix Transformations: } x \rightarrow x' = g \circ x = R x$$

Translations and scale transformations are a particular case of the first and rotations of the second. Let's start with the location and scale parameters; that is, a density

$$p(x|\mu, \sigma) dx = \frac{1}{\sigma} f\left(\frac{x - \mu}{\sigma}\right) dx \quad (485)$$

the Affine group  $G = \{g \equiv (a, b); a \in R; b \in R^+\}$  so  $x' = g \circ x = a + b x$  and the model will be invariant if

$$(\mu', \sigma') = g \circ (\mu, \sigma) = (a, b) \circ (\mu, \sigma) = (a + b\mu, b\sigma) \quad (486)$$

Now,

$$\begin{aligned} \int p(\cdot|\mu', \sigma') \pi_L(\mu', \sigma') d\mu' d\sigma' &= \int p(\cdot|g \circ (\mu, \sigma)) \pi_L(\mu, \sigma) d\mu d\sigma = \\ &= \int p(\cdot|\mu', \sigma') \{ \pi_L[g^{-1}(\mu', \sigma')] J(\mu', \sigma'; \mu, \sigma) \} d\mu' d\sigma' = \\ &= \int p(\cdot|\mu', \sigma') \left\{ \pi_L\left(\frac{\mu' - a}{b}, \frac{\sigma'}{b}\right) \frac{1}{b^2} \right\} d\mu' d\sigma' \end{aligned} \quad (487)$$

and this should hold for all  $(a, b) \in R \times R^+$  so, in consequence:

$$d\mu_L(\mu, \sigma) = \pi_L(\mu, \sigma) d\mu d\sigma \propto \frac{1}{\sigma^2} d\mu d\sigma \quad (488)$$

However, the group of Affine Transformations is non-abelian so if we study the the action on the left, there is no reason why we should not consider also the action on the right. Since

$$(\mu', \sigma') = (\mu, \sigma) \circ g = (\mu, \sigma) \circ (a, b) = (\mu + a\sigma, b\sigma) \quad (489)$$

the same reasoning leads to (left as exercise):

$$d\mu_R(\mu, \sigma) = \pi_R(\mu, \sigma) d\mu d\sigma \propto \frac{1}{\sigma} d\mu d\sigma \quad (490)$$

The first one ( $\pi_L$ ) is the one we obtain using Jeffrey's rule in two dimensions while  $\pi_R$  is the one we get for position and scale parameters or Jeffrey's rule treating both parameters independently; that is, as two one-dimensional problems instead a one two-dimensional problem. Thus, although from the invariance point of view there is no reason why one should prefer one over the other, the right invariant Haar prior gives more consistent results. In fact ([?], [?]), a necessary and sufficient condition for a sequence of posteriors based on proper priors to converge in probability to an invariant posterior is that the prior is the right Haar measure.

**Problem 3.4:** As a remainder, given a measure space  $(\Omega, \mathcal{B}, \mu)$  a mapping  $T : \Omega \rightarrow \Omega$  is measurable if  $T^{-1}(A) \in \mathcal{B}$  for all  $A \in \mathcal{B}$  and the measure  $\mu$  is invariant under  $T$  if  $\mu(T^{-1}(A)) = \mu(A)$  for all  $A \in \mathcal{B}$ . Show that the measure  $d\mu(\theta) = [\theta(1 - \theta)]^{-1/2} d\theta$  is invariant under the mapping  $T : [0, 1] \rightarrow [0, 1]$  such that  $T : \theta \rightarrow \theta' = T(\theta) = 4\theta(1 - \theta)$ . This is the Jeffrey's prior for the Binomial model  $Bi(x|N, \theta)$ .



**Problem 3.5:** Consider the n-dimensional spherical surface  $S_n$  of unit radius,  $\mathbf{x} \in S_n$  and the transformation  $\mathbf{x}' = \mathbf{R}\mathbf{x} \in S_n$  where  $\mathbf{R} \in SO(n)$ . Show that the Haar invariant measure is the Lebesgue measure on the sphere.

Hint: Recall that  $\mathbf{R}$  is an orthogonal matrix so  $\mathbf{R}^t = \mathbf{R}^{-1}$ ; that  $|\det \mathbf{R}| = 1$  so  $J(\mathbf{x}'; \mathbf{x}) = |\partial \mathbf{x} / \partial \mathbf{x}'| = |\partial \mathbf{R}^{-1} \mathbf{x}' / \partial \mathbf{x}'| = |\det \mathbf{R}| = 1$  and that  $\mathbf{x}'^t \mathbf{x}' = \mathbf{x}^t \mathbf{x} = 1$ .

**Example 3.12: Bivariate Normal Distribution.** Let  $\mathbf{X} = (X_1, X_2) \sim N(\mathbf{x} | \mathbf{0}, \phi)$  with  $\phi = \{\sigma_1, \sigma_2, \rho\}$ ; that is:

$$p(\mathbf{x} | \phi) = (2\pi)^{-1} |\det[\Sigma]|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{x}^t \Sigma^{-1} \mathbf{x}) \right\} \quad (491)$$

with the covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \quad \text{and} \quad \det[\Sigma] = \sigma_1^2\sigma_2^2(1 - \rho^2) \quad (492)$$

Using the Cholesky decomposition we can express  $\Sigma^{-1}$  as the product of two lower (or upper) triangular matrices:

$$\Sigma^{-1} = \frac{1}{\det[\Sigma]} \begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix} = \mathbf{A}^t \mathbf{A} \quad \text{with} \quad \mathbf{A} = \begin{pmatrix} \frac{1}{\sigma_1} & 0 \\ \frac{-\rho}{\sigma_1\sqrt{1-\rho^2}} & \frac{1}{\sigma_2\sqrt{1-\rho^2}} \end{pmatrix} \quad (493)$$

For the action on the left:

$$\mathbf{M} = \mathbf{T} = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix} ; a, b > 0 \quad \longrightarrow \quad J(\mathbf{A}'; \mathbf{A}) = a^2c \quad (494)$$

and, in consequence

$$\pi(aa'_{11}, aa'_{21} + ba'_{22}, ca'_{22}) ac^2 = \pi(a'_{11}, a'_{21}, a'_{22}) \quad \longrightarrow \quad \pi(a'_{11}, a'_{21}, a'_{22}) \propto \frac{1}{a'^2_{11} a'_{22}} \quad (495)$$

and  $\det[\Sigma] = (\det[\Sigma^{-1}])^{-1} = (\det[\mathbf{A}])^{-2}$ . Thus, in the new parameterization  $\theta = \{a_{11}, a_{21}, a_{22}\}$

$$p(\mathbf{x} | \theta) = (2\pi)^{-1} |\det[\mathbf{A}]| \exp \left\{ -\frac{1}{2} (\mathbf{x}^t \mathbf{A}^t \mathbf{A} \mathbf{x}) \right\} \quad (496)$$

Consider now the group of lower triangular 2x2 matrices

$$G_l = \{\mathbf{T} \in LT_{2x2}; T_{ii} > 0\} \quad (497)$$

Since  $\mathbf{T}^{-1} \in G_l$ , inserting the identity matrix  $\mathbf{I} = \mathbf{T}\mathbf{T}^{-1} = \mathbf{T}^{-1}\mathbf{T}$  we have: action

On the Left	On the Right
$\mathbf{T} \circ \mathbf{x} \rightarrow \mathbf{T}\mathbf{x} = \mathbf{x}'$	$\mathbf{x} \circ \mathbf{T} \rightarrow \mathbf{T}^{-1}\mathbf{x} = \mathbf{x}'$
$[\mathbf{x}^t (\mathbf{T}^t (\mathbf{T}^t)^{-1}) \mathbf{A}^t \mathbf{A} (\mathbf{T}^{-1} \mathbf{T}) \mathbf{x}]$	$[\mathbf{x}^t ((\mathbf{T}^t)^{-1} \mathbf{T}^t) \mathbf{A}^t \mathbf{A} (\mathbf{T} \mathbf{T}^{-1}) \mathbf{x}]$
$\mathbf{M} = \mathbf{T}$	$\mathbf{M} = \mathbf{T}^{-1}$

Then

$$\mathbf{M}\mathbf{x} = \mathbf{x}' \quad ; \quad \mathbf{x} = \mathbf{M}^{-1}\mathbf{x}' \quad ; \quad \mathbf{x}'^t = \mathbf{x}^t \mathbf{M}^t \quad \text{and} \quad d\mathbf{x} = \frac{1}{|\det[\mathbf{M}]|} d\mathbf{x}' \quad (498)$$

so

$$p(\mathbf{x}' | \theta) = (2\pi)^{-1} \frac{|\det[\mathbf{A}]|}{|\det[\mathbf{M}]|} \exp \left\{ -\frac{1}{2} (\mathbf{x}'^t (\mathbf{A}\mathbf{M}^{-1})^t (\mathbf{A}\mathbf{M}^{-1}) \mathbf{x}') \right\} \quad (499)$$

and the model is invariant under  $G_l$  if the action on the parameter space is

$$G_l : \mathbf{A} \longrightarrow \mathbf{A}' = \mathbf{A}\mathbf{M}^{-1} \quad ; \quad \mathbf{A} = \mathbf{A}'\mathbf{M} \quad ; \quad \det[\mathbf{A}] = \det[\mathbf{A}'] \det[\mathbf{M}] \quad (500)$$

so

$$p(\mathbf{x}'|\theta') = (2\pi)^{-1} |\det[\mathbf{A}']| \exp \left\{ -\frac{1}{2} \left( \mathbf{x}'^t \mathbf{A}'^t \mathbf{A}' \mathbf{x}' \right) \right\} \quad (501)$$

Then, the Haar equation reads

$$\int_{\Theta} p(\bullet|\mathbf{A}') \pi(\mathbf{A}') d\mathbf{A}' = \int_{\Theta} p(\bullet|g\circ\mathbf{A}) \pi(\mathbf{A}) d\mathbf{A} = \int_{\Theta} p(\bullet|\mathbf{A}') \pi(\mathbf{A}'\mathbf{M}) J(\mathbf{A}'; \mathbf{A}) d\mathbf{A}$$

and, in consequence,  $\forall \mathbf{M} \in G$

$$\pi(\mathbf{A}'\mathbf{M}) J(\mathbf{A}'; \mathbf{A}) da'_{11} da'_{21} da'_{22} = \pi(\mathbf{A}') da'_{11} da'_{21} da'_{22} \quad (502)$$

For the action on the left:

$$\mathbf{M} = \mathbf{T} = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix} \quad ; \quad a, b > 0 \quad \longrightarrow \quad J(\mathbf{A}'; \mathbf{A}) = a^2 c \quad (503)$$

and, in consequence

$$\pi(aa'_{11}, aa'_{21} + ba'_{22}, ca'_{22}) a^2 c = \pi(a'_{11}, a'_{21}, a'_{22}) \longrightarrow \pi(a'_{11}, a'_{21}, a'_{22}) \propto \frac{1}{a'^2_{11} a'_{22}} \quad (504)$$

For the action on the right:

$$\mathbf{M} = \mathbf{T}^{-1} = \begin{pmatrix} a^{-1} & 0 \\ -b(ac)^{-1} & c^{-1} \end{pmatrix} \quad \longrightarrow \quad J(\mathbf{A}'; \mathbf{A}) = (ac^2)^{-1} \quad (505)$$

and, in consequence

$$\pi\left(\frac{a'_{11}}{a}, \frac{ca'_{21} - ba'_{22}}{ac}, \frac{a'_{22}}{c}\right) \frac{1}{ac^2} = \pi(a'_{11}, a'_{21}, a'_{22}) \longrightarrow \pi(a'_{11}, a'_{21}, a'_{22}) \propto \frac{1}{a'_{11} a'^2_{22}} \quad (506)$$

In terms of the parameters of interest  $\{\sigma_1, \sigma_2, \rho\}$ , since

$$da_{11} da_{21} da_{22} = \frac{1}{\sigma_1^2 \sigma_2^2 (1 - \rho^2)^2} d\sigma_1 d\sigma_2 d\rho \quad (507)$$

we have finally that for invariance under  $G_l$ :

$$\pi_L^l(\sigma_1, \sigma_2, \rho) = \frac{1}{\sigma_1 \sigma_2 (1 - \rho^2)^{3/2}} \quad \text{and} \quad \pi_R^l(\sigma_1, \sigma_2, \rho) = \frac{1}{\sigma_2^2 (1 - \rho^2)} \quad (508)$$

The same analysis with decomposition in upper triangular matrices leads to

$$\pi_L^u(\sigma_1, \sigma_2, \rho) = \frac{1}{\sigma_1 \sigma_2 (1 - \rho^2)^{3/2}} \quad \text{and} \quad \pi_R^u(\sigma_1, \sigma_2, \rho) = \frac{1}{\sigma_1^2 (1 - \rho^2)} \quad (509)$$

As we see, in both cases the left Haar invariant prior coincides with Jeffrey's prior when  $\{\mu_1, \mu_2\}$  and  $\{\sigma_1, \sigma_2, \rho\}$  are decoupled.

At this point, one may be tempted to use a right Haar invariant prior where the two parameters  $\sigma_1$  and  $\sigma_2$  are treated on equal footing

$$\pi(\sigma_1, \sigma_2, \rho) = \frac{1}{\sigma_1 \sigma_2 (1 - \rho^2)} \quad (510)$$

Under this prior, since the sample correlation

$$r = \frac{\sum_i (x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2)}{(\sum_i (x_{1i} - \bar{x}_1)^2 \sum_i (x_{2i} - \bar{x}_2)^2)^{1/2}} \quad (511)$$

is a sufficient statistics for  $\rho$ , we have that the posterior for inferences on the correlation coefficient will be

$$p(\rho|\mathbf{x}) \propto (1 - \rho^2)^{(n-3)/2} F(n-1, n-1, n-1/2; (1+r\rho)/2) \quad (512)$$

with  $F(a, b, c; z)$  the Hypergeometric Function.

**Example 3.13:** If  $\theta \in \Theta \rightarrow g \circ \theta = \phi(\theta) = \theta' \in \Theta$  with  $\phi(\theta)$  is a one-to-one differentiable mapping, then

$$\begin{aligned} \int_{\Theta} p(\bullet|\theta') d\mu(\theta) &= \int_{\Theta} p(\bullet|\theta') \pi(\theta) d\theta = \int_{\Theta} p(\bullet|\theta') \pi(\phi^{-1}(\theta')) \left| \frac{\partial \phi^{-1}(\theta')}{\partial \theta} \right| d\theta = \\ &= \int_{\Theta} p(\bullet|\theta') \pi(\theta') d\theta' = \int_{\Theta} p(\bullet|\theta') d\mu(\theta') \end{aligned} \quad (513)$$

and therefore, Jeffreys' prior defines a Haar invariant measure.

### 3.6.5 Conjugated Distributions

In as much as possible, we would like to consider reference priors  $\pi(\theta|a, b, \dots)$  versatile enough such that by varying some of the parameters  $a, b, \dots$  we get diverse forms to analyze the effect on the final results and, on the other hand, to simplify the evaluation of integrals like:

$$p(x) = \int p(x|\theta) \cdot p(\theta) d\theta \quad \text{and} \quad p(y|x) = \int p(y|\theta) \cdot p(\theta|x) d\theta \quad (514)$$

This leads us to consider as reference priors the *Conjugated Distributions* [?].

Let  $\mathcal{S}$  be a class of sampling distributions  $p(x|\theta)$  and  $\mathcal{P}$  the class of prior densities for the parameter  $\theta$ . If

$$p(\theta|x) \in \mathcal{P} \quad \text{for all} \quad p(x|\theta) \in \mathcal{S} \quad \text{and} \quad p(\theta) \in \mathcal{P} \quad (515)$$

we say that the class  $\mathcal{P}$  is conjugated to  $\mathcal{S}$ . We are mainly interested in the class of priors  $\mathcal{P}$  that have the same functional form as the likelihood. In this case, since both the prior density and the posterior belong to the same family of distributions, we say that they are *closed under sampling*. It should be stressed that the criteria for taking conjugated reference priors is eminently practical and, in many cases, they do not exist. In fact, only the *exponential family* of distributions has conjugated prior densities. Thus, if  $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$  is an exchangeable random sampling from the k-parameter regular exponential family, then

$$p(\mathbf{x}|\theta) = f(\mathbf{x}) g(\theta) \exp \left\{ \sum_{j=1}^k c_j \phi_j(\theta) \left( \sum_{i=1}^n h_j(x_i) \right) \right\} \quad (516)$$

and the *conjugated prior* will have the form:

$$\pi(\theta|\tau) = \frac{1}{K(\tau)} [g(\theta)]^{\tau_0} \exp \left\{ \sum_{j=1}^k c_j \phi_j(\theta) \tau_j \right\} \quad (517)$$

where  $\theta \in \Theta$ ,  $\tau = \{\tau_0, \tau_1, \dots, \tau_k\}$  the *hyperparameters* and  $K(\tau) < \infty$  the normalization factor so  $\int_{\Theta} \pi(\theta|\tau) d\theta = 1$ . Then, the general scheme will be <sup>21</sup>:

- 1) Choose the class of priors  $\pi(\theta|\tau)$  that reflect the structure of the model;
- 2) Choose a prior function  $\pi(\tau)$  for the *hyperparameters*;

<sup>21</sup>We can go an step upwards and assign a prior to the hyperparameters with hyper-hyperparameters....

- 3) Express the posterior density as  $p(\theta, \tau|\mathbf{x}) \propto p(\mathbf{x}|\theta)\pi(\theta|\tau)\pi(\tau)$ ;  
 4) Marginalize for the parameters of interest:

$$p(\theta|\mathbf{x}) \propto \int_{\Phi} p(\mathbf{x}|\theta)\pi(\theta|\tau)\pi(\tau)d\tau \quad (518)$$

or, if desired, get the conditional density

$$p(\theta|\mathbf{x}, \tau) = \frac{p(\mathbf{x}, \theta, \tau)}{p(\mathbf{x}, \tau)} = \frac{p(\mathbf{x}|\theta)\pi(\theta|\tau)}{p(\mathbf{x}|\tau)} \quad (519)$$

The obvious question that arises is how do we choose the prior  $\pi(\phi)$  for the hyperparameters. Besides *reasonableness*, we may consider two approaches. Integrating the parameters  $\theta$  of interest, we get

$$p(\tau, \mathbf{x}) = \pi(\tau) \int_{\Theta} p(\mathbf{x}|\theta)\pi(\theta|\tau) d\theta = \pi(\tau) p(\mathbf{x}|\tau) \quad (520)$$

so we may use any of the procedures under discussion to take  $\pi(\tau)$  as the prior for the model  $p(\mathbf{x}|\tau)$  and then obtain

$$\pi(\theta) = \int_{\Omega_{\tau}} \pi(\theta|\tau) \pi(\tau) d\tau \quad (521)$$

The beauty of Bayes rule but not very practical in complicated situations. A second approach, more ugly and practical, is the so called *Empirical Method* where we assign numeric values to the hyperparameters suggested by  $p(\mathbf{x}|\tau)$  (for instance, moments, maximum-likelihood estimation,...); that is, setting, in a distributional sense,  $\pi(\tau) = \delta_{\tau^*}$  so  $\langle \pi(\tau), p(\theta, \mathbf{x}, \tau) \rangle = p(\theta, \mathbf{x}, \tau^*)$ . Thus,

$$p(\theta|\mathbf{x}, \tau^*) \propto p(\mathbf{x}|\theta)\pi(\theta|\tau^*) \quad (522)$$

Obviously, fixing the hyperparameters assumes a perfect knowledge of them and does not allow for variations but the procedure may be useful to guess at least were to go.

Last, it may happen that a single conjugated prior does not represent sufficiently well our beliefs. In this case, we may consider a k-mixture of conjugated priors

$$\pi(\theta|\tau_1, \dots, \tau_k) = \sum_{i=1}^k w_i \pi(\theta|\tau_i) \quad (523)$$

In fact [?], any prior density for a model that belongs to the exponential family can be approximated arbitrarily close by a mixture of conjugated priors.

**Example 3.14:** Let's see the conjugated prior distributions for some models:

- **Poisson model**  $Po(n|\mu)$ : Writing

$$p(n|\mu) = \frac{e^{-\mu} \mu^n}{\Gamma(n+1)} = \frac{e^{-(\mu-n\log \mu)}}{\Gamma(n+1)} \quad (524)$$

it is clear that the Poisson distribution belongs to the exponential family and the conjugated prior density for the parameter  $\mu$  is

$$\pi(\mu|\tau_1, \tau_2) \propto e^{-\tau_1\mu + \tau_2 \log \mu} \propto Ga(\mu|\tau_1, \tau_2) \quad (525)$$

If we set a prior  $\pi(\tau_1, \tau_2)$  for the hyperparameters we can write

$$p(n, \mu, \tau_1, \tau_2) p(n|\mu) \pi(\mu|\tau_1, \tau_2) = \pi(\tau_1, \tau_2) \quad (526)$$

and integrating  $\mu$ :

$$p(n, \tau_1, \tau_2) = \left[ \frac{\Gamma(n + \tau_2)}{\Gamma(\tau_1)} \frac{\tau_1^{\tau_2}}{(1 + \tau_1)^{n + \tau_2}} \right] \pi(\tau_1, \tau_2) = p(n|\tau_1, \tau_2) \pi(\tau_1, \tau_2) \quad (527)$$

• **Binomial model**  $Bi(n|N, \theta)$ : Writing

$$P(n|N, \theta) = \binom{N}{n} \theta^n (1 - \theta)^{N-n} = \binom{N}{n} e^{n \log \theta + (N-n) \log (1-\theta)} \quad (528)$$

it is clear that it belong to the exponential family and and the conjugated prior density for the parameter  $\theta$  will be:

$$\pi(\theta|\tau_1, \tau_2) = Be(\tau|\tau_1, \tau_2) \quad (529)$$

• **Multinomial model** Let  $\mathbf{X} = (X_1, X_2, \dots, X_k) \sim Mn(\mathbf{x}|\theta)$ ; that is:

$$\mathbf{X} \sim p(\mathbf{x}|\theta) = \Gamma(n + 1) \prod_{i=1}^k \frac{\theta_i^{x_i}}{\Gamma(x_i + 1)} \quad \begin{cases} X_i \in N, & \sum_{i=1}^k X_i = n \\ \theta_i \in [0, 1], & \sum_{i=1}^k \theta_i = 1 \end{cases} \quad (530)$$

The Dirichlet distribution  $Di(\theta|\alpha)$ :

$$\pi(\theta|\alpha) = D(\alpha) \prod_{i=1}^k \theta_i^{\alpha_i - 1} \quad \begin{cases} \alpha = (\alpha_1, \alpha_2, \dots, \alpha_k), & \alpha_i > 0, \quad \sum_{i=1}^k \alpha_i = \alpha_0 \\ D(\alpha) = \Gamma(\alpha_0) \left[ \prod_{i=1}^k \Gamma(\alpha_i) \right]^{-1} \end{cases} \quad (531)$$

is the natural conjugated prior for this model. It is a degenerated distribution in the sense that

$$\pi(\theta|\alpha) = D(\alpha) \left[ \prod_{i=1}^{k-1} \theta_i^{\alpha_i - 1} \right] \left[ 1 - \sum_{i=1}^{k-1} \theta_i \right]^{\alpha_k - 1} \quad (532)$$

The posterior density will then be  $\theta \sim Di(\theta|\mathbf{x} + \alpha)$  with

$$E[\theta_i] = \frac{x_i + \alpha_i}{n + \alpha_0} \quad \text{and} \quad V[\theta_i, \theta_j] = \frac{E[\theta_i](\delta_{ij} - E[\theta_j])}{n + \alpha_0 + 1} \quad (533)$$

The parameters  $\alpha$  of the Dirichlet distribution  $Di(\theta|\alpha)$  determine the expected values  $E[\theta_i] = \alpha_i/\alpha_0$ . In practice, it is more convenient to control also the variances and use the *Generalized Dirichlet Distribution*  $GDi(\theta|\alpha, \beta)$ :

$$\pi(\theta|\alpha, \beta) = \prod_{i=1}^{k-1} \frac{\Gamma(\alpha_i + \beta_i)}{\Gamma(\alpha_i)\Gamma(\beta_i)} \theta_i^{\alpha_i - 1} \left[ 1 - \sum_{j=1}^i \theta_j \right]^{\gamma_i} \quad (534)$$

where:

$$0 < \theta_i < 1, \quad \sum_{i=1}^{k-1} \theta_i < 1, \quad \theta_n = 1 - \sum_{i=1}^{k-1} \theta_i \quad (535)$$

$$\alpha_i > 0, \quad \beta_i > 0, \quad \text{and} \quad \gamma_i \begin{cases} \beta_i - \alpha_{i+1} - \beta_{i+1}; & i = 1, 2, \dots, k-2 \\ \beta_{k-1} - 1; & i = k-1 \end{cases} \quad (536)$$

When  $\beta_i = \alpha_{i+1} + \beta_{i+1}$  it becomes the Dirichlet distribution. For this prior we have that

$$E[\theta_i] = \frac{\alpha_i}{\alpha_i + \beta_i} S_i \quad \text{and} \quad V[\theta_i, \theta_j] = E[\theta_j] \left( \frac{\alpha_i + \delta_{ij}}{\alpha_i + \beta_i + 1} T_i - E[\theta_i] \right) \quad (537)$$

where

$$S_i = \prod_{j=1}^{i-1} \frac{\beta_j}{\alpha_j + \beta_j} \quad \text{and} \quad T_i = \prod_{j=1}^{i-1} \frac{\beta_j + 1}{\alpha_j + \beta_j + 1} \quad (538)$$

with  $S_1 = T_1 = 1$  and we can have control over the prior means and variances.

### 3.6.6 Probability Matching Priors

A *probability matching prior* is a prior function such that the one sided credible intervals derived from the posterior distribution coincide, to a certain level of accuracy, with those derived by the classical approach. This condition leads to a differential equation for the prior distribution [?], [?]. We shall illustrate in the following lines the rationale behind for the simple one parameter case assuming that the needed regularity conditions are satisfied. Consider then a random quantity  $X \sim p(x|\theta)$  and an iid sampling  $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$  with  $\theta$  the parameter of interest. The classical approach for inferences is based on the the likelihood

$$p(\mathbf{x}|\theta) = p(x_1, x_2, \dots, x_n|\theta) = \prod_{i=1}^n p(x_i|\theta) \quad (539)$$

and goes through the following reasoning:

- 1) Assumes that the parameter  $\theta$  has the *true* but unknown value  $\theta_0$  so the sample is actually drawn from  $p(x|\theta_0)$ ;
- 2) Find the estimator  $\theta_m(\mathbf{x})$  of  $\theta_0$  as the value of  $\theta$  that maximizes the likelihood; that is:

$$\theta_m = \max_{\theta} \{p(\mathbf{x}|\theta)\} \quad \rightarrow \quad \left( \frac{\partial \ln p(\mathbf{x}|\theta)}{\partial \theta} \right)_{\theta_m} = 0 \quad (540)$$

- 3) Given the model  $X \sim p(x|\theta_0)$ , after the appropriate change change of variables get the distribution  $p(\theta_m|\theta_0)$  of the random quantity  $\theta_m(X_1, X_2, \dots, X_n)$  and draw inferences from it.

The Bayesian inferential process considers a prior distribution  $\pi(\theta)$  and draws inferences on  $\theta$  from the posterior distribution of the quantity of interest

$$p(\theta|\mathbf{x}) \propto p(\mathbf{x}|\theta) \pi(\theta) \quad (541)$$

Let's start with the Bayesian and expand the term on the right around  $\theta_m$ . On the one hand:

$$\ln \frac{p(\mathbf{x}|\theta)}{p(\mathbf{x}|\theta_m)} = \frac{1}{2!} \left( \frac{\partial^2 \ln p(\mathbf{x}|\theta)}{\partial \theta^2} \right)_{\theta_m} (\theta - \theta_m)^2 + \frac{1}{3!} \left( \frac{\partial^3 \ln p(\mathbf{x}|\theta)}{\partial \theta^3} \right)_{\theta_m} (\theta - \theta_m)^3 + \dots \quad (542)$$

Now,

$$-\frac{1}{n} \frac{\partial^2 \ln p(\mathbf{x}|\theta)}{\partial \theta^2} = \frac{1}{n} \sum_{i=1}^n \frac{\partial^2 (-\ln p(x_i|\theta))}{\partial \theta^2} \xrightarrow{n \rightarrow \infty} E_X \left[ \frac{\partial^2 (-\ln p(x|\theta))}{\partial \theta^2} \right] = I(\theta) \quad (543)$$

so we can substitute:

$$\left( \frac{\partial^2 \ln p(\mathbf{x}|\theta)}{\partial \theta^2} \right)_{\theta_m} = -n I(\theta_m) \quad \text{and} \quad \left( \frac{\partial^3 \ln p(\mathbf{x}|\theta)}{\partial \theta^3} \right)_{\theta_m} = -n \left( \frac{\partial I(\theta)}{\partial \theta} \right)_{\theta_m} \quad (544)$$

to get

$$p(\mathbf{x}|\theta) = e^{\ln p(\mathbf{x}|\theta)} \propto e^{-\frac{nI(\theta_m)}{2} (\theta - \theta_m)^2} \left( 1 - \frac{n}{3!} \left( \frac{\partial I(\theta)}{\partial \theta} \right)_{\theta_m} (\theta - \theta_m)^3 + \dots \right) \quad (545)$$

On the other hand:

$$\pi(\theta) = \pi(\theta_m) \left( 1 + \left( \frac{1}{\pi(\theta)} \frac{\partial \pi(\theta)}{\partial \theta} \right)_{\theta_m} (\theta - \theta_m) + \dots \right) \quad (546)$$

so If we define the random quantity  $T = \sqrt{nI(\theta_m)}(\theta - \theta_m)$  and consider that

$$I^{-3/2}(\theta) \frac{\partial I(\theta)}{\partial \theta} = -2 \frac{\partial I^{-1/2}}{\partial \theta} \quad (547)$$

we get finally:

$$p(t|\mathbf{x}) = \frac{\exp(-t^2/2)}{\sqrt{2\pi}} \left( 1 + \frac{1}{\sqrt{n}} \left[ \left( \frac{I^{-1/2}(\theta)}{\pi(\theta)} \frac{\partial \pi(\theta)}{\partial \theta} \right)_{\theta_m} t + \frac{1}{3} \left( \frac{\partial I^{-1/2}}{\partial \theta} \right)_{\theta_m} t^3 \right] + O\left(\frac{1}{n}\right) \right) \quad (548)$$

Let's now find

$$P(T \leq z|\mathbf{x}) = \int_{-\infty}^z p(t|\mathbf{x}) dt \quad (549)$$

Defining

$$Z(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad \text{and} \quad P(x) = \int_{-\infty}^x Z(t) dt \quad (550)$$

and considering that

$$\int_{-\infty}^z Z(t) t dt = -Z(z) \quad \text{and} \quad \int_{-\infty}^z Z(t) t^3 dt = -Z(z) (z^2 + 2) \quad (551)$$

it is straight forward to get:

$$P(T \leq z|\mathbf{x}) = P(z) - \frac{Z(z)}{\sqrt{n}} \left[ \left( \frac{I^{-1/2}(\theta)}{\pi(\theta)} \frac{\partial \pi(\theta)}{\partial \theta} \right)_{\theta_m} + \frac{z^2 + 2}{3} \left( \frac{\partial I^{-1/2}}{\partial \theta} \right)_{\theta_m} \right] + O\left(\frac{1}{n}\right) \quad (552)$$

From this probability distribution, we can infer what the classical approach will get. Since he will draw inferences from  $p(\mathbf{x}|\theta_0)$ , we can take a sequence of proper priors  $\pi_k(\theta|\theta_0)$  for  $k = 1, 2, \dots$  that induce a sequence of distributions such that

$$\lim_{k \rightarrow \infty} \langle \pi_k(\theta|\theta_0), p(\mathbf{x}|\theta) \rangle = p(\mathbf{x}|\theta_0) \quad (553)$$

In Distributional sense, the sequence of distributions generated by

$$\pi_k(\theta|\theta_0) = \frac{k}{2} \mathbf{1}_{[\theta_0 - 1/k, \theta_0 + 1/k]} \quad ; \quad k = 1, 2, \dots \quad (554)$$

converge to the Delta distribution  $\delta_{\theta_0}$  and, from distributional derivatives, as  $k \rightarrow \infty$ ,

$$\left\langle \frac{d}{d\theta} \pi_k(\theta|\theta_0), I^{-1/2}(\theta) \right\rangle = - \left\langle \pi_k(\theta|\theta_0), \frac{d}{d\theta} I^{-1/2}(\theta) \right\rangle \simeq - \left( \frac{\partial I^{-1/2}(\theta)}{\partial \theta} \right)_{\theta_0} \quad (555)$$

But  $\theta_0 = \theta_m + O(1/\sqrt{n})$  so, for a sequence of priors that shrink to  $\theta_0 \simeq \theta_m$ ,

$$P(T \leq z|\mathbf{x}) = P(z) - \frac{Z(z)}{\sqrt{n}} \left[ \frac{z^2 + 1}{3} \left( \frac{\partial I^{-1/2}}{\partial \theta} \right)_{\theta_m} \right] + O\left(\frac{1}{n}\right) \quad (556)$$

For terms of order  $O(1/\sqrt{n})$  in the equations (1) and (2) to be the same, we need that:

$$\left( \frac{1}{\sqrt{I(\theta)}} \frac{1}{\pi(\theta)} \frac{\partial \pi(\theta)}{\partial \theta} \right)_{\theta_m} = - \left( \frac{\partial I^{-1/2}}{\partial \theta} \right)_{\theta_m} \quad (557)$$

and therefore

$$\pi(\theta) = I^{1/2}(\theta) \quad (558)$$

that is, Jeffrey's prior. In the case of  $n$ -dimensional parameters, the reasoning goes along the same lines but the expressions and the development become much more lengthy and messy and we refer to the literature.

The procedure for a first order *probability matching prior* [?], [?] starts from the likelihood

$$p(x_1, x_2, \dots, x_n | \theta_1, \theta_2, \dots, \theta_p) \quad (559)$$

and then:

- 1) Get the Fisher's matrix  $\mathbf{I}(\theta_1, \theta_2, \dots, \theta_p)$  and the inverse  $\mathbf{I}^{-1}(\theta_1, \theta_2, \dots, \theta_p)$ ;
- 2) Suppose we are interested in the parameter  $t = t(\theta_1, \theta_2, \dots, \theta_p)$  a twice continuous and differentiable function of the parameters. Define the column vector

$$\nabla_t = \left( \frac{\partial t}{\partial \theta_1}, \frac{\partial t}{\partial \theta_2}, \dots, \frac{\partial t}{\partial \theta_p} \right)^T \quad (560)$$

- 3) Define the column vector

$$\eta = \frac{\mathbf{I}^{-1} \nabla_t}{(\nabla_t^T \mathbf{I}^{-1} \nabla_t)^{1/2}} \quad \text{so that} \quad \eta^T \mathbf{I} \eta = 1 \quad (561)$$

- 4) The probability matching prior for the parameter  $t = t(\theta)$  in terms of  $\theta_1, \theta_2, \dots, \theta_p$  is given by the equation:

$$\sum_{k=1}^p \frac{\partial}{\partial \theta_k} [\eta_k(\theta) \pi(\theta)] = 0 \quad (562)$$

Any solution  $\pi(\theta_1, \theta_2, \dots, \theta_p)$  will do the job.

- 5) Introduce  $t = t(\theta)$  in this expression, say, for instance  $\theta_1 = \theta_1(t, \theta_2, \dots, \theta_p)$ , and the corresponding Jacobian  $J(t, \theta_2, \dots, \theta_p)$ . Then we get the prior for the parameter  $t$  of interest and the nuisance parameters  $\theta_2, \dots, \theta_p$  that, eventually, will be integrated out.

**Example 3.15:** Consider two independent random quantities  $X_1$  and  $X_2$  such that

$$P(X_i = n_k) = Po(n_k | \mu_i). \quad (563)$$

We are interested in the parameter  $t = \mu_1/\mu_2$  so setting  $\mu = \mu_2$  we have the ordered parameterization  $\{t, \mu\}$ . The joint probability is

$$P(n_1, n_2 | \mu_1, \mu_2) = P(n_1 | \mu_1) P(n_2 | \mu_2) = e^{-(\mu_1 + \mu_2)} \frac{\mu_1^{n_1} \mu_2^{n_2}}{\Gamma(n_1 + 1) \Gamma(n_2 + 1)} \quad (564)$$

from which we get the Fisher's matrix

$$\mathbf{I}(\mu_1, \mu_2) = \begin{pmatrix} 1/\mu_1 & 0 \\ 0 & 1/\mu_2 \end{pmatrix} \quad \text{and} \quad \mathbf{I}^{-1}(\mu_1, \mu_2) = \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix} \quad (565)$$

We are interested in the parameter  $t = \mu_1/\mu_2$ , a twice continuous and differentiable function of the parameters, so

$$\nabla_t(\mu_1, \mu_2) = \left( \frac{\partial t}{\partial \mu_1}, \frac{\partial t}{\partial \mu_2} \right)^T = (\mu_2^{-1}, -\mu_1 \mu_2^{-2})^T = \begin{pmatrix} \mu_2^{-1} \\ -\mu_1 \mu_2^{-2} \end{pmatrix} \quad (566)$$



Therefore:

$$\mathbf{I}^{-1} \nabla_t = \begin{pmatrix} \mu_1 \mu_2^{-1} \\ -\mu_1 \mu_2^{-1} \end{pmatrix} \quad S = \nabla_t^T \mathbf{I}^{-1} \nabla_t = \frac{\mu_1(\mu_1 + \mu_2)}{\mu_2^3} \quad (567)$$

$$\eta = \frac{\mathbf{I}^{-1} \nabla_t}{(\nabla_t^T \mathbf{I}^{-1} \nabla_t)^{1/2}} = \begin{pmatrix} (\mu_1 \mu_2)^{1/2} (\mu_1 + \mu_2)^{-1/2} \\ -(\mu_1 \mu_2)^{1/2} (\mu_1 + \mu_2)^{-1/2} \end{pmatrix} \quad (568)$$

so that  $\eta^T \mathbf{I} \eta = 1$ . The probability matching prior for the parameter  $t = \mu_1/\mu_2$  in terms of  $\mu_1$  and  $\mu_2$  is given by the equation:

$$\sum_{k=1}^2 \frac{\partial}{\partial \mu_k} [\eta_k(\mu) \pi(\mu)] = 0 \quad (569)$$

so, if  $f(\mu_1, \mu_2) = (\mu_1 \mu_2)^{1/2} (\mu_1 + \mu_2)^{-1/2}$ , we have to solve

$$\frac{\partial}{\partial \mu_1} f(\mu_1, \mu_2) \pi(\mu_1, \mu_2) = \frac{\partial}{\partial \mu_2} f(\mu_1, \mu_2) \pi(\mu_1, \mu_2) \quad (570)$$

Any solution will do so:

$$\pi(\mu_1, \mu_2) \propto f^{-1}(\mu_1, \mu_2) = \frac{\sqrt{\mu_1 + \mu_2}}{\sqrt{\mu_1 \mu_2}} \quad (571)$$

Substituting  $\mu_1 = t\mu_2$  and including the Jacobian  $J = \mu_2$  we have finally:

$$\pi(t, \mu_2) \propto \sqrt{\mu_2} \sqrt{\frac{1+t}{t}} \quad (572)$$

The posterior density will be:

$$p(t, \mu_2 | n_1, n_2) \propto p(n_1, n_2 | t, \mu_2) \pi(t, \mu_2) \propto e^{-\mu_2(1+t)} t^{n_1-1/2} (1+t)^{1/2} \mu_2^{n_2+3/2-1} \quad (573)$$

and, integrating the nuisance parameter  $\mu_2 \in [0, \infty)$ , we get the posterior density:

$$p(t | n_1, n_2) = N \frac{t^{n_1-1/2}}{(1+t)^{n+1}} \quad (574)$$

with  $N^{-1} = B(n_1 + 1/2, n_2 + 1/2)$ .

**Example 3.16: Gamma distribution.** Show that for  $Ga(x|\alpha, \beta)$ :

$$p(x|\alpha, \beta) = \frac{\alpha^\beta}{\Gamma(\beta)} e^{-\alpha x} x^{\beta-1} \mathbf{1}_{(0, \infty)}(x) \quad (575)$$

the probability matching prior for the ordering

- $\{\beta, \alpha\}$  is  $\pi(\alpha, \beta) = \beta^{-1/2} \left[ \alpha^{-1} \sqrt{\beta \Psi'(\beta)} - 1 \right]$
- $\{\alpha, \beta\}$  is  $\pi(\alpha, \beta) = \left[ \alpha^{-1} \sqrt{\Psi'(\beta)} \right] \sqrt{\beta \Psi'(\beta)} - 1$

to be compared with Jeffrey's prior  $\pi_2^J(\alpha, \beta) = \alpha^{-1} \sqrt{\beta \Psi'(\beta)} - 1$  and Jeffrey's prior when both parameters are treated individually  $\pi_{1+1}^J(\alpha, \beta) = \alpha^{-1} \sqrt{\Psi'(\beta)}$

**Example 3.17: Bivariate Normal Distribution.**

For the ordered parameterization  $\rho, \sigma_1, \sigma_2$ : the Fisher's matrix (see example 3.12) is:

$$\mathbf{I}(\rho, \sigma_1, \sigma_2) = (1 - \rho^2)^{-1} \begin{pmatrix} (1 + \rho^2)(1 - \rho^2)^{-1} & -\rho\sigma_1^{-1} & -\rho\sigma_2^{-1} \\ -\rho\sigma_1^{-1} & (2 - \rho^2)\sigma_1^{-2} & -\rho^2(\sigma_1\sigma_2)^{-1} \\ -\rho\sigma_2^{-1} & -\rho^2(\sigma_1\sigma_2)^{-1} & (2 - \rho^2)\sigma_2^{-2} \end{pmatrix} \quad (576)$$

and the inverse:

$$\mathbf{I}^{-1}(\rho, \sigma_1, \sigma_2) = \frac{1}{2} \begin{pmatrix} 2(1 - \rho^2)^2 & \sigma_1 \rho(1 - \rho^2) & \sigma_2 \rho(1 - \rho^2) \\ \sigma_1 \rho(1 - \rho^2) & \sigma_1^2 & \rho^2 \sigma_1 \sigma_2 \\ \sigma_2 \rho(1 - \rho^2) & \rho^2 \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix} \quad (577)$$

Then

$$\frac{2}{\rho} \frac{\partial}{\partial \rho} [\pi(1 - \rho^2)] + \frac{\partial}{\partial \sigma_1} [\pi \sigma_1] + \frac{\partial}{\partial \sigma_2} [\pi \sigma_2] = 0 \quad (578)$$

for which

$$\pi(\sigma_1, \sigma_2, \rho) = \frac{1}{\sigma_1 \sigma_2 (1 - \rho^2)} \quad (579)$$

is a solution.

**Problem 3.6:** Consider

$$X \sim p(x|a, b, \sigma) = \frac{\sinh[\sigma(b - a)]}{2(b - a)} \frac{1}{\cosh[\sigma(x - a)] \cosh[\sigma(b - x)]} \mathbf{1}_{(-\infty, \infty)}(x) \quad (580)$$

where  $a < b \in \mathcal{R}$  and  $\sigma \in (0, \infty)$ . Show that

$$E[X] = \frac{b + a}{2} \quad \text{and} \quad V[x] = \frac{(b - a)^2}{12} + \frac{\pi^2}{12\sigma^2} \quad (581)$$

and that, for known  $\sigma \gg$ , the probability matching prior for  $a$  and  $b$  tends to  $\pi_{pm}(a, b) \sim (b - a)^{-1/2}$ . Show also that, under the same limit,  $\pi_{pm}(\theta) \sim \theta^{-1/2}$  for  $(a, b) = (-\theta, \theta)$  and  $(a, b) = (0, \theta)$ . Since  $p(x|a, b, \sigma) \xrightarrow{\sigma \gg} Un(x|a, b)$  discuss in this last case what is the difference with the example 3.4.

### 3.6.7 Reference Analysis

The expected amount of information ( *Expected Mutual Information*) on the parameter  $\theta$  provided by  $k$  independent observations of the model  $p(\mathbf{x}|\theta)$  relative to the prior knowledge on  $\theta$  described by  $\pi(\theta)$  is

$$I[e(k), \pi(\theta)] = \int_{\Theta} \pi(\theta) d\theta \int_{\Omega_X} p(\mathbf{z}_k|\theta) \log \frac{p(\theta|\mathbf{z}_k)}{\pi(\theta)} d\mathbf{z}_k \quad (582)$$

where  $\mathbf{z}_k = \{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ . If  $\lim_{k \rightarrow \infty} I[e(k), \pi(\theta)]$  exists, it will quantify the maximum amount of information that we could obtain on  $\theta$  from experiments described by this model relative to the prior knowledge  $\pi(\theta)$ . The central idea of the *reference analysis* [?], [?] is to take as *reference prior for the model*  $p(\mathbf{x}|\theta)$  that which maximizes the maximum amount of information we may get so it will be the *less informative* for this model. From Calculus of Variations, if we introduce the prior  $\pi^*(\theta) = \pi(\theta) + \epsilon \eta(\theta)$  with  $\pi(\theta)$  an extremal of the expected information  $I[e(k), \pi(\theta)]$  and  $\eta(\theta)$  such that

$$\int_{\Theta} \pi(\theta) d\theta = \int_{\Theta} \pi^*(\theta) d\theta = 1 \quad \longrightarrow \quad \int_{\Theta} \eta(\theta) d\theta = 0 \quad (583)$$

it is it is easy to see (left as exercise) that

$$\pi(\theta) \propto \exp \left\{ \int_{\Omega_X} p(\mathbf{z}_k|\theta) \log p(\theta|\mathbf{z}_k) d\mathbf{z}_k \right\} = f_k(\theta) \quad (584)$$

This is a nice but complicated implicit equation because, on the one hand,  $f_k(\theta)$  depends on  $\pi(\theta)$  through the posterior  $p(\theta|\mathbf{z}_k)$  and, on the other hand, the limit  $k \rightarrow \infty$  is usually divergent (intuitively, the more

precision we want for  $\theta$ , the more information is needed and to know the actual value from the experiment requires an infinite amount of information). This can be circumvented regularizing the expression as

$$\pi(\theta) \propto \pi(\theta_0) \lim_{k \rightarrow \infty} \frac{f_k(\theta)}{f_k(\theta_0)} \quad (585)$$

with  $\theta_0$  any interior point of  $\Theta$  (we are used to that in particle physics!). Let's see some examples.

**Example 3.18:** Consider again the exponential model for which  $t = n^{-1} \sum_{i=1}^n x_i$  is sufficient for  $\theta$  and distributed as

$$p(t|\theta) = \frac{(n\theta)^n}{\Gamma(n)} t^{n-1} \exp\{-n\theta t\} \quad (586)$$

Taking  $\pi(\theta) = \mathbf{1}_{(0,\infty)}(\theta)$  we have the proper posterior

$$\pi^*(\theta|t) = \frac{(nt)^{n+1}}{\Gamma(n+1)} \exp\{-n\theta t\} \theta^n \quad (587)$$

Then  $\log \pi^*(\theta|t) = -(n\theta)t + n \log \theta + (n+1) \log t + g_1(n)$  and

$$f_n(\theta) = \exp \left\{ \int_{\Omega_x} p(t|\theta) \log \pi^*(\theta|t) dt \right\} = \frac{g_2(n)}{\theta} \longrightarrow \pi(\theta) \propto \pi(\theta_0) \lim_{n \rightarrow \infty} \frac{f_n(\theta)}{f_n(\theta_0)} \propto \frac{1}{\theta} \quad (588)$$

**Example 3.19:** Prior functions depend on the particular model we are treating. To learn about a parameter, we can do different experimental designs that respond to different models and, even though the parameter is the same, they may have different priors. For instance, we may be interested in the *acceptance*; the probability to accept an event under some conditions. For this, we can generate for instance a sample of  $N$  observed events and see how many ( $x$ ) pass the conditions. This experimental design corresponds to a Binomial distribution

$$p(x|N, \theta) = \binom{N}{x} \theta^x (1-\theta)^{N-x} \quad (589)$$

with  $x = \{0, 1, \dots, N\}$ . For this model, the reference prior (also Jeffrey's and PM) is  $\pi(\theta) = \theta^{-1/2}(1-\theta)^{-1/2}$  and the posterior  $\theta \sim Be(\theta|x+1/2, N-x+1/2)$ . Conversely, we can generate events until  $r$  are accepted and see how many ( $x$ ) have we generated. This experimental design corresponds to a Negative Binomial distribution

$$p(x|r, \theta) = \binom{x-1}{r-1} \theta^r (1-\theta)^{x-r} \quad (590)$$

where  $x = r, r+1, \dots$  and  $r \geq 1$ . For this model, the reference prior (Jeffrey's and PM too) is  $\pi(\theta) = \theta^{-1}(1-\theta)^{-1/2}$  and the posterior  $\theta \sim Be(\theta|r, x-r+1/2)$ .

**Problem 3.7:** Consider

1)  $X \sim Po(x|\theta) = \exp\{-\theta\} \frac{\theta^x}{\Gamma(x+1)}$  and the experiment  $e(k) \xrightarrow{iid} \{x_1, x_2, \dots, x_k\}$ . Take  $\pi^*(\theta) = \mathbf{1}_{(0,\infty)}(\theta)$ , and show that

$$\pi(\theta) \propto \pi(\theta_0) \lim_{k \rightarrow \infty} \frac{f_k(\theta)}{f_k(\theta_0)} \propto \theta^{-1/2} \quad (591)$$

2)  $X \sim Bi(x|N, \theta) = \binom{N}{x} \theta^x (1-\theta)^{N-x}$  and the experiment  $e(k) \xrightarrow{iid} \{x_1, x_2, \dots, x_k\}$ . Take  $\pi^*(\theta) \propto \theta^{a-1}(1-\theta)^{b-1} \mathbf{1}_{(0,1)}(\theta)$  with  $a, b > 0$  and show that

$$\pi(\theta) \propto \pi(\theta_0) \lim_{k \rightarrow \infty} \frac{f_k(\theta)}{f_k(\theta_0)} \propto \theta^{-1/2}(1-\theta)^{-1/2} \quad (592)$$

(Hint: For 1) and 2) consider the Taylor expansion of  $\log \Gamma(z, \cdot)$  around  $E[z]$  and the asymptotic behavior of the Polygamma Function  $\Psi^n(z) = a_n z^{-n} + a_{n+1} z^{-(n+1)} + \dots$ ).

3)  $X \sim Un(x|0, \theta)$  and the iid sample  $\{x_1, x_2, \dots, x_k\}$ . For inferences on  $\theta$ , show that  $f_k = \theta^{-1}g(k)$  and in consequence the posterior is Pareto  $Pa(\theta|x_M, n)$  with  $x_M = \max\{x_1, x_2, \dots, x_k\}$  the sufficient statistic.

A very useful constructive theorem to obtain the *reference prior* is given in [?]. First, a *permissible prior* for the model  $p(\mathbf{x}|\theta)$  is defined as a strictly positive function  $\pi(\theta)$  such that it renders a proper posterior; that is,

$$\forall x \in \Omega_X \quad \int_{\Theta} p(\mathbf{x}|\theta) \pi(\theta) d\theta < \infty \tag{593}$$

and that for some approximating sequence  $\Theta_k \subset \Theta$ ;  $\lim_{k \rightarrow \infty} \Theta_k = \Theta$ , the sequence of posteriors  $p_k(\theta|\mathbf{x}) \propto p(\mathbf{x}|\theta)\pi_k(\theta)$  converges logarithmically to  $p(\theta|\mathbf{x}) \propto p(\mathbf{x}|\theta)\pi(\theta)$ . Then, the *reference prior* is just a *permissible prior* that maximizes the maximum amount of information the experiment can provide for the parameter. The constructive procedure for a one-dimensional parameter consists on:

1) Take  $\pi^*(\theta)$  as a continuous strictly positive function such that the corresponding posterior

$$\pi^*(\theta|\mathbf{z}_k) = \frac{p(\mathbf{z}_k|\theta) \pi^*(\theta)}{\int_{\Theta} p(\mathbf{z}_k|\theta) \pi(\theta) d\theta} \tag{594}$$

is proper and asymptotically consistent.  $\pi^*(\theta)$  is arbitrary so it can be taken for convenience to simplify the integrals.

2) Obtain

$$f_k^*(\theta) = \exp \left\{ \int_{\Omega_X} p(\mathbf{z}_k|\theta) \log \pi^*(\theta|\mathbf{z}_k) d\mathbf{z}_k \right\} \quad \text{and} \quad h_k(\theta; \theta_0) = \frac{f_k^*(\theta)}{f_k^*(\theta_0)} \tag{595}$$

for any interior point  $\theta_0 \in \Theta$ ;

3) If

- 3.1) each  $f_k^*(\theta)$  is continuous;
- 3.2) for any fixed  $\theta$  and large  $k$ , is  $h_k(\theta; \theta_0)$  is either monotonic in  $k$  or bounded from above by  $h(\theta)$  that is integrable on any compact set;
- 3.3)  $\pi(\theta) = \lim_{k \rightarrow \infty} h_k(\theta; \theta_0)$  is a *permissible prior function*

then  $\pi(\theta)$  is a reference prior for the model  $p(\mathbf{x}|\theta)$ . It is important to note that there is no requirement on the existence of the Fisher's information  $\mathbf{I}(\theta)$ . If it exists, a simple Taylor expansion of the densities shows that for a one-dimensional parameter  $\pi(\theta) = [\mathbf{I}(\theta)]^{1/2}$  in consistency with Jeffrey's proposal. Usually, the last is easier to evaluate but not always as we shall see.

In many cases  $\text{supp}(\theta)$  is unbounded and the prior  $\pi(\theta)$  is not a proper density. As we have seen this is not a problem as long as the posterior  $p(\theta|\mathbf{z}_k) \propto p(\mathbf{z}_k|\theta)\pi(\theta)$  is proper although, in any case, one can proceed "*more formally*" considering a sequence of proper priors  $\pi_m(\theta)$  defined on a sequence of compact sets  $\Theta_m \subset \Theta$  such that  $\lim_{m \rightarrow \infty} \Theta_m = \Theta$  and taking the limit of the corresponding sequence of posteriors  $p_m(\theta|\mathbf{z}_k) \propto p(\mathbf{z}_k|\theta)\pi_m(\theta)$ . Usually simple sequences as for example  $\Theta_m = [1/m, m]$ ;  $\lim_{m \rightarrow \infty} \Theta_m = (0, \infty)$ , or  $\Theta_m = [-m, m]$ ;  $\lim_{m \rightarrow \infty} \Theta_m = (-\infty, \infty)$  will suffice.

When the parameter  $\theta$  is n-dimensional, the procedure is more laborious. First, one starts [?] arranging the parameters in decreasing order of importance  $\{\theta_1, \theta_2, \dots, \theta_n\}$  (as we did for the Probability Matching Priors) and then follow the previous scheme to obtain the conditional prior functions

$$\pi(\theta_n|\theta_1, \theta_2, \dots, \theta_{n-1}) \pi(\theta_{n-1}|\theta_1, \theta_2, \dots, \theta_{n-2}) \cdots \pi(\theta_2|\theta_1) \pi(\theta_1) \tag{596}$$

For instance in the case of two parameters and the ordered parameterization  $\{\theta, \lambda\}$ :

- 1) Get the conditional  $\pi(\lambda|\theta)$  as the reference prior for  $\lambda$  keeping  $\theta$  fixed;
- 2) Find the marginal model

$$p(\mathbf{x}|\theta) = \int_{\Phi} p(\mathbf{x}|\theta, \lambda) \pi(\lambda|\theta) d\lambda \quad (597)$$

- 3) Get the reference prior  $\pi(\theta)$  from the marginal model  $p(\mathbf{x}|\theta)$

Then  $\pi(\theta, \lambda) \propto \pi(\lambda|\theta)\pi(\theta)$ . This is fine if  $\pi(\lambda|\theta)$  and  $\pi(\theta)$  are proper functions; seldom the case. Otherwise one has to define the appropriate sequence of compact sets observing, among other things, that this has to be done for the full parameter space and usually the limits depend on the parameters. Suppose that we have the sequence  $\Theta_i \times \Lambda_i \xrightarrow{i \rightarrow \infty} \Theta \times \Lambda$ . Then:

- 1) Obtain  $\pi_i(\lambda|\theta)$ :

$$\begin{aligned} \pi_i^*(\lambda|\theta) \mathbf{1}_{\Lambda_i}(\lambda) &\longrightarrow \pi_i^*(\lambda|\theta, \mathbf{z}_k) = \frac{p(\mathbf{z}_k|\theta, \lambda) \pi_i^*(\lambda|\theta)}{\int_{\Lambda_i} p(\mathbf{z}_k|\theta, \lambda) \pi_i^*(\lambda|\theta) d\lambda} \longrightarrow \\ &\longrightarrow \pi_i(\lambda|\theta) = \lim_{k \rightarrow \infty} \frac{f_k^*(\lambda|\Lambda_i, \theta, \dots)}{f_k^*(\lambda_0|\Lambda_i, \theta, \dots)} \end{aligned} \quad (598)$$

- 2) Get the marginal density  $p_i(\mathbf{x}|\theta)$ :

$$p_i(\mathbf{x}|\theta) = \int_{\Lambda_i} p(\mathbf{x}|\theta, \lambda) \pi_i(\lambda|\theta) d\lambda \quad (599)$$

- 3) Determine  $\pi_i(\theta)$ :

$$\pi_i^*(\theta) \mathbf{1}_{\Theta_i}(\theta) \longrightarrow \pi_i^*(\theta|\mathbf{z}_k) = \frac{p_i(\mathbf{z}_k|\theta) \pi_i^*(\theta)}{\int_{\Theta_i} p_i(\mathbf{z}_k|\theta) \pi_i^*(\theta) d\theta} \longrightarrow \pi_i(\theta) = \lim_{k \rightarrow \infty} \frac{f_k^*(\theta|\Theta_i, \Lambda_i, \dots)}{f_k^*(\theta_0|\Theta_i, \Lambda_i, \dots)} \quad (600)$$

- 4) The reference prior for the ordered parameterization  $\{\theta, \lambda\}$  will be:

$$\pi(\theta, \lambda) = \lim_{i \rightarrow \infty} \frac{\pi_i(\lambda|\theta) \pi_i(\theta)}{\pi_i(\lambda_0|\theta_0) \pi_i(\theta_0)} \quad (601)$$

In the case of two parameters, if  $\Lambda$  is independent of  $\theta$  the Fisher's matrix usually exists and, if  $\mathbf{I}(\theta, \lambda)$  and  $\mathbf{S}(\theta, \lambda) = \mathbf{I}^{-1}(\theta, \lambda)$  are such that:

$$\mathbf{I}_{22}(\theta, \lambda) = a_1^2(\theta) b_1^2(\lambda) \quad \text{and} \quad \mathbf{S}_{11}(\theta, \lambda) = a_0^{-2}(\theta) b_0^{-2}(\lambda) \quad (602)$$

then [?]  $\pi(\theta, \lambda) = \pi(\lambda|\theta)\pi(\theta) = a_0(\theta) b_1(\lambda)$  is a permissible prior even if the conditional reference priors are not proper. The reference priors are usually *probability matching priors*.

**Example 3.20:** A simple example is the Multinomial distribution  $\mathbf{X} \sim Mn(\mathbf{x}|\theta)$  with  $\dim \mathbf{X} = k + 1$  and probability

$$p(\mathbf{x}|\theta) \propto \theta_1^{x_1} \theta_2^{x_2} \dots \theta_k^{x_k} (1 - \delta_k)^{x_{k+1}} \quad ; \quad \delta_k = \sum_{j=1}^k \theta_j \quad (603)$$

Consider the ordered parameterization  $\{\theta_1, \theta_2, \dots, \theta_k\}$ . Then

$$\pi(\theta_1, \theta_2, \dots, \theta_k) = \pi(\theta_k|\theta_{k-1}, \theta_{k-2} \dots \theta_2, \theta_1) \pi(\theta_{k-1}|\theta_{k-2} \dots \theta_2, \theta_1) \dots \pi(\theta_2|\theta_1) \pi(\theta_1) \quad (604)$$

In this case, all the conditional densities are proper

$$\pi(\theta_m|\theta_{m-1}, \dots, \theta_1) \propto \theta_m^{-1/2} (1 - \delta_m)^{-1/2} \quad (605)$$

and therefore

$$\pi(\theta_1, \theta_2, \dots, \theta_k) \propto \prod_{i=1}^k \theta_i^{-1/2} (1 - \delta_i)^{-1/2} \quad (606)$$

The posterior density will be then

$$p(\theta|\mathbf{x}) \propto \left[ \prod_{i=1}^k \theta_i^{x_i-1/2} (1 - \delta_i)^{-1/2} \right] (1 - \delta_k)^{x_{k+1}} \quad (607)$$

**Example 3.21:** Consider again the case of two independent Poisson distributed random quantities  $X_1$  and  $X_2$  with joint density density

$$P(n_1, n_2 | \mu_1, \mu_2) = P(n_1 | \mu_1) P(n_2 | \mu_2) = e^{-(\mu_1 + \mu_2)} \frac{\mu_1^{n_1} \mu_2^{n_2}}{\Gamma(n_1 + 1) \Gamma(n_2 + 1)} \quad (608)$$

We are interested in the parameter  $\theta = \mu_1/\mu_2$  so setting  $\mu = \mu_2$  we have the ordered parameterization  $\{\theta, \mu\}$  and:

$$P(n_1, n_2 | \theta, \mu) = e^{-\mu(1 + \theta)} \frac{\theta^{n_1} \mu^n}{\Gamma(n_1 + 1) \Gamma(n_2 + 1)} \quad (609)$$

where  $n = n_1 + n_2$ . Since  $E[X_1] = \mu_1 = \theta\mu$  and  $E[X_2] = \mu_2 = \mu$  the Fisher's matrix and its inverse will be

$$\mathbf{I} = \begin{pmatrix} \mu/\theta & 1 \\ 1 & (1 + \theta)/\mu \end{pmatrix}; \quad \det(\mathbf{I}) = \theta^{-1} \quad \text{and} \quad \mathbf{S} = \mathbf{I}^{-1} = \begin{pmatrix} \theta(1 + \theta)/\mu & -\theta \\ -\theta & \mu \end{pmatrix} \quad (610)$$

Therefore

$$S_{11} = \theta(1 + \theta)/\mu \quad \text{and} \quad F_{22} = (1 + \theta)/\mu \quad (611)$$

and, in consequence:

$$\pi(\theta) f_1(\mu) \propto S_{11}^{-1/2} = \frac{\sqrt{\mu}}{\sqrt{\theta(1 + \theta)}} \quad \pi(\mu|\theta) f_2(\theta) \propto F_{22}^{1/2} = \frac{\sqrt{1 + \theta}}{\sqrt{\mu}} \quad (612)$$

Thus, we have for the ordered parameterization  $\{\theta, \mu\}$  the reference prior:

$$\pi(\theta, \mu) = \pi(\mu|\theta) \pi(\theta) \propto \frac{1}{\sqrt{\mu\theta(1 + \theta)}} \quad (613)$$

and the posterior density will be:

$$p(\theta, \mu | n_1, n_2) \propto \exp\{-\mu(1 + \theta)\} \theta^{n_1-1/2} (1 + \theta)^{-1/2} \mu^{n-1/2} \quad (614)$$

and, integrating the nuisance parameter  $\mu \in [0, \infty)$  we get finally

$$p(\theta | n_1, n_2) = N \frac{\theta^{n_1-1/2}}{(1 + \theta)^{n+1}} \quad (615)$$

with  $\theta = \mu_1/\mu_2$ ,  $n = n_1 + n_2$  and  $N^{-1} = B(n_1 + 1/2, n_2 + 1/2)$ . The distribution function will be:

$$P(\theta | n_1, n_2) = \int_0^\theta p(\theta' | n_1, n_2) d\theta' = I(\theta/(1 + \theta); n_1 + 1/2, n_2 + 1/2) \quad (616)$$

with  $I(x; a, b)$  the Incomplete Beta Function and the moments, when they exist;

$$E[\theta^m] = \frac{\Gamma(n_1 + 1/2 + m) \Gamma(n_2 + 1/2 - m)}{\Gamma(n_1 + 1/2) \Gamma(n_2 + 1/2)} \quad (617)$$

It is interesting to look at the problem from a different point of view. Consider again the ordered parameterization  $\{\theta, \lambda\}$  with  $\theta = \mu_1/\mu_2$  but now, the nuisance parameter is  $\lambda = \mu_1 + \mu_2$ . The likelihood will be:

$$P(n_1, n_2|\theta, \lambda) = \frac{1}{\Gamma(n_1 + 1)\Gamma(n_2 + 1)} e^{-\lambda} \lambda^n \frac{\theta^{n_1}}{(1 + \theta)^n} \tag{618}$$

The domains are  $\Theta = (0, \infty)$  and  $\Lambda = (0, \infty)$ , independent. Thus, no need to specify the prior for  $\lambda$  since

$$p(\theta|n_1, n_2) \propto \pi(\theta) \frac{\theta^{n_1}}{(1 + \theta)^n} \int_{\Lambda} e^{-\lambda} \lambda^n \pi(\lambda) d\lambda \propto \frac{\theta^{n_1}}{(1 + \theta)^n} \pi(\theta) \tag{619}$$

In this case we have that

$$I(\theta) \propto \frac{1}{\theta(1 + \theta)^2} \quad \rightarrow \quad \pi(\theta) = \frac{1}{\theta^{1/2}(1 + \theta)} \tag{620}$$

and, in consequence,

$$p(\theta|n_1, n_2) = N \frac{\theta^{n_1-1/2}}{(1 + \theta)^{n+1}} \tag{621}$$

**Problem 3.8:** Show that the reference prior for the Pareto distribution  $Pa(x|\theta, x_0)$  (see example 3.9) is  $\pi(\theta, x_0) \propto (\theta x_0)^{-1}$  and that for an iid sample  $\mathbf{x} = \{x_1, \dots, x_n\}$ , if  $x_m = \min\{x_i\}_{i=1}^n$  and  $a = \sum_{i=1}^n \ln(x_i/x_m)$  the posterior

$$p(\theta, x_0|\mathbf{x}) = \frac{na^{n-1}}{x_m \Gamma(n-1)} e^{-a\theta} \theta^{n-1} \left(\frac{x_0}{x_m}\right)^{n\theta-1} \mathbf{1}_{(0,\infty)}(\theta) \mathbf{1}_{(0,x_m)}(x_0) \tag{622}$$

is proper for a sample size  $n > 1$ . Obtain the marginal densities

$$p(\theta|\mathbf{x}) = \frac{a^{n-1}}{\Gamma(n-1)} e^{-a\theta} \theta^{n-2} \mathbf{1}_{(0,\infty)}(\theta) \quad \text{and} \tag{623}$$

$$p(x_0|\mathbf{x}) = \frac{n(n-1)}{a} x_0^{-1} \left[1 + \frac{n}{a} \ln\left(\frac{x_m}{x_0}\right)\right]^{-n} \mathbf{1}_{(0,x_m)}(x_0) \tag{624}$$

and show that for large  $n$  (see section 10.2)  $E[\theta] \simeq na^{-1}$  and  $E[x_0] \simeq x_m$ .

**Problem 3.9:** Show that for the shifted Pareto distribution (Lomax distribution):

$$p(x|\theta, x_0) = \frac{\theta}{x_0} \left(\frac{x_0}{x + x_0}\right)^{\theta+1} \mathbf{1}_{(0,\infty)}(x) \quad ; \quad \theta, x_0 \in R^+ \tag{625}$$

the reference prior for the ordered parameterization  $\{\theta, x_0\}$  is  $\pi_r(\theta, x_0) \propto (x_0\theta(\theta + 1))^{-1}$  and for  $\{x_0, \theta\}$  is  $\pi_r(x_0, \theta) \propto (x_0\theta)^{-1}$ . Show that the first one is a first order probability matching prior while the second is not. In fact, show that for  $\{x_0, \theta\}$ ,  $\pi_{pm}(x_0, \theta) \propto (x_0\theta^{3/2}\sqrt{\theta + 2})^{-1}$  is a matching prior and that for both orderings the Jeffrey's prior is  $\pi_J(\theta, x_0) \propto (x_0(\theta + 1)\sqrt{\theta(\theta + 2)})^{-1}$ .

**Problem 3.10:** Show that for the Weibull distribution

$$p(x|\alpha, \beta) = \alpha\beta x^{\beta-1} \exp\{-\alpha x^\beta\} \mathbf{1}_{(0,\infty)}(x) \tag{626}$$

with  $\alpha, \beta > 0$ , the reference prior functions are

$$\pi_r(\beta, \alpha) = (\alpha\beta)^{-1} \quad \text{and} \quad \pi_r(\alpha, \beta) = \left(\alpha\beta\sqrt{\zeta(2) + (\psi(2) - \ln \alpha)^2}\right)^{-1} \tag{627}$$

for the ordered parameterizations  $\{\beta, \alpha\}$  and  $\{\alpha, \beta\}$  respectively being  $\zeta(2) = \pi^2/6$  the Riemann Zeta Function and  $\psi(2) = 1 - \gamma$  the Digamma Function.

### 3.7 Hierarchical Structures

In many circumstances, even though the experimental observations respond to the same phenomena it is not always possible to consider the full set of observations as an exchangeable sequence but rather exchangeability within subgroups of observations. As stated earlier, this may be the case when the results come from different experiments or when, within the same experiment, data taking conditions (acceptances, efficiencies,...) change from run to run. A similar situation holds, for instance, for the results of responses under a drug performed at different hospitals when the underlying conditions of the population vary between zones, countries,... In general, we shall have different groups of observations

$$\begin{aligned} \mathbf{x}_1 &= \{x_{11}, x_{21}, \dots, x_{n_1 1}\} \\ &\vdots \\ \mathbf{x}_j &= \{x_{1j}, x_{2j}, \dots, x_{n_j j}\} \\ &\vdots \\ \mathbf{x}_J &= \{x_{1J}, x_{2J}, \dots, x_{n_J J}\} \end{aligned}$$

from  $J$  experiments  $e_1(n_1), e_2(n_2), \dots, e_J(n_J)$ . Within each sample  $\mathbf{x}_j$ , we can consider that exchangeability holds and also for the sets of observations  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_J\}$ . In this case, it is appropriate to consider *hierarchical structures*.

Let's suppose that for each experiment  $e(j)$  the observations are drawn from the model

$$p(\mathbf{x}_j | \theta_j) \quad ; \quad j = 1, 2, \dots, J \quad (628)$$

Since the experiments are independent we assume that the parameters of the sequence  $\{\theta_1, \theta_2, \dots, \theta_J\}$  are exchangeable and that, although different, they can be assumed to have a common origin since they respond to the same phenomena. Thus, we can set

$$p(\theta_1, \theta_2, \dots, \theta_J | \phi) = \prod_{i=1}^J p(\theta_i | \phi) \quad (629)$$

with  $\phi$  the *hyperparameters* for which we take a prior  $\pi(\phi)$ . Then we have the structure (see figure ??)

$$p(\mathbf{x}_1, \dots, \mathbf{x}_J, \theta_1, \dots, \theta_J, \phi) = \pi(\phi) \prod_{i=1}^J p(\mathbf{x}_i | \theta_i) \pi(\theta_i | \phi) \quad (630)$$

This structure can be repeated sequentially if we consider appropriate to assign a prior  $\pi(\phi | \tau)$  to the hyperparameters  $\phi$  so that

$$p(\mathbf{x}, \boldsymbol{\theta}, \phi, \tau) = p(\mathbf{x} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta} | \phi) \pi(\phi | \tau) \pi(\tau) \quad (631)$$

Now, consider the model  $p(\mathbf{x}, \theta, \phi)$ . We may be interested in  $\theta$ , in the hyperparameters  $\phi$  or in both. In general we shall need the conditional densities:

$$\bullet \quad p(\phi | \mathbf{x}) \propto p(\phi) \int p(\mathbf{x} | \boldsymbol{\theta}) p(\boldsymbol{\theta} | \phi) d\boldsymbol{\theta} \quad (632)$$

$$\bullet \quad p(\boldsymbol{\theta} | \mathbf{x}, \phi) = \frac{p(\boldsymbol{\theta}, \mathbf{x}, \phi)}{p(\mathbf{x}, \phi)} \quad \text{and} \quad (633)$$

$$\bullet \quad p(\boldsymbol{\theta} | \mathbf{x}) = \frac{p(\mathbf{x} | \boldsymbol{\theta})}{p(\mathbf{x})} p(\boldsymbol{\theta}) = \frac{p(\mathbf{x} | \boldsymbol{\theta})}{p(\mathbf{x})} \int p(\boldsymbol{\theta} | \phi) p(\phi) d\phi \quad (634)$$



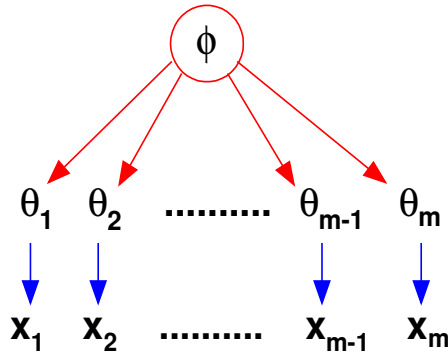


Fig. 8: Structure of the hierarchical model.

that can be expressed as

$$p(\boldsymbol{\theta}|\mathbf{x}) = \int \frac{p(\mathbf{x}, \boldsymbol{\theta}, \phi)}{p(\mathbf{x})} d\phi = \int p(\boldsymbol{\theta}|\mathbf{x}, \phi) p(\phi|\mathbf{x}) d\phi \quad (635)$$

and, since

$$p(\boldsymbol{\theta}|\mathbf{x}) = p(\mathbf{x}|\boldsymbol{\theta}) \int p(\boldsymbol{\theta}|\phi) \frac{p(\phi|\mathbf{x})}{p(\mathbf{x}|\phi)} d\phi \quad (636)$$

we can finally write

$$\frac{p(\boldsymbol{\theta}, \phi)}{p(\mathbf{x})} = p(\boldsymbol{\theta}|\phi) \frac{p(\phi)}{p(\mathbf{x})} = p(\boldsymbol{\theta}|\phi) \frac{p(\phi|\mathbf{x})}{p(\mathbf{x}|\phi)} \quad (637)$$

In general, this conditional densities have complicated expressions and we shall use Monte Carlo methods to proceed (see Gibbs Sampling, example 4.15, in section 3).

It is important to note that if the prior distributions are not proper we can have improper marginal and posterior densities that obviously have no meaning in the inferential process. Usually, conditional densities are better behaved but, in any case, we have to check that this is so. In general, the better behaved is the likelihood the wildest behavior we can accept for the prior functions. We can also used prior distributions that are a mixture of proper distributions:

$$p(\boldsymbol{\theta}|\phi) = \sum_i w_i p_i(\boldsymbol{\theta}|\phi) \quad (638)$$

with  $w_i \geq 0$  and  $\sum w_i = 1$  so that the combination is convex and we assure that it is proper density or, extending this to a continuous mixture:

$$p(\boldsymbol{\theta}|\phi) = \int w(\boldsymbol{\sigma}) p(\boldsymbol{\theta}|\phi, \boldsymbol{\sigma}) d\boldsymbol{\sigma} \quad (639)$$

### 3.8 Priors for discrete parameters

So far we have discussed parameters with continuous support but in some cases it is either finite or countable. If the parameter of interest can take only a finite set of  $n$  possible values, the reasonable option for an *uninformative prior* is a Discrete Uniform Probability  $P(X = x_i) = 1/n$ . In fact,

maximizing the expected information provided by the experiment with the normalization constraint (i.e. the probability distribution for which the *prior* knowledge is minimal) drives to  $P(X = x_i) = 1/n$  in accordance with the *Principle of Insufficient Reason*.

Even though finite discrete parameter spaces are either the most usual case we shall have to deal with or, at least, a sufficiently good approximation for the real situation, it may happen that a non-informative prior is not the most appropriate (see example 3.22). On the other hand, if the parameter takes values on a countable set the problem is more involved. A possible way out is to devise a hierarchical structure in which we assign the discrete parameter  $\theta$  a prior  $\pi(\theta|\lambda)$  with  $\lambda$  a set of continuous hyperparameters. Then, since

$$p(\mathbf{x}, \lambda) = \sum_{\theta \in \Theta} p(\mathbf{x}|\theta) \pi(\theta|\lambda) \pi(\lambda) = p(\mathbf{x}|\lambda) \pi(\lambda) \quad (640)$$

we get the prior  $\pi(\lambda)$  by any of the previous procedures for continuous parameters with the model  $p(\mathbf{x}|\lambda)$  and obtain

$$\pi(\theta) \propto \int_{\Lambda} \pi(\theta|\lambda) \pi(\lambda) d\lambda \quad (641)$$

Different procedures are presented and discussed in [?].

**Example 3.22:** The absolute value of the electric charge ( $Z$ ) of a particle is to be determined from the number of photons observed by a Cherenkov Counter. We know from test beam studies and Monte Carlo simulations that the number of observed photons  $n_\gamma$  produced by a particle of charge  $Z$  is well described by a Poisson distribution with parameter  $\mu = n_0 Z^2$ ; that is

$$P(n_\gamma|n_0, Z) = e^{-n_0 Z^2} \frac{(n_0 Z^2)^{n_\gamma}}{\Gamma(n_\gamma + 1)} \quad (642)$$

so  $E[n_\gamma|Z = 1] = n_0$ . First, by physics considerations  $Z$  has a finite support  $\Omega_Z = \{1, 2, \dots, n\}$ . Second, we know *a priori* that not all incoming nuclei are equally likely so a *non-informative* prior may not be the best choice. In any case, a discrete uniform prior will give the posterior:

$$P(Z = k|n_\gamma, n_0, n) = \frac{e^{-n_0 k^2} k^{2n_\gamma}}{\sum_{k=1}^n e^{-n_0 k^2} k^{2n_\gamma}} \quad (643)$$

### 3.9 Constrains on parameters and priors

Consider a parametric model  $p(\mathbf{x}|\theta)$  and the prior  $\pi_0(\theta)$ . Now we have some information on the parameters that we want to include in the prior. Typically we shall have say  $k$  constraints of the form

$$\int_{\Theta} g_i(\theta) \pi(\theta) d\theta = a_i \quad ; \quad i = 1, \dots, k \quad (644)$$

Then, we have to find the prior  $\pi(\theta)$  for which  $\pi_0(\theta)$  is the best approximation, in the Kullback-Leibler sense, including the constraints with the corresponding Lagrange multipliers  $\lambda_i$ ; that is, the extremal of

$$\mathcal{F} = \int_{\Theta} \pi(\theta) \log \frac{\pi(\theta)}{\pi_0(\theta)} d\theta + \sum_{i=1}^k \lambda_i \left( \int_{\Theta} g_i(\theta) \pi(\theta) d\theta - a_i \right) \quad (645)$$

Again, it is left as an exercise to show that from Calculus of Variations we have the well known solution

$$\pi(\theta) \propto \pi_0(\theta) \exp\left\{ \sum_{i=1}^k \lambda_i g_i(\theta) \right\} \quad \text{where} \quad \lambda_i \mid \int_{\Theta} g_i(\theta) \pi(\theta) d\theta = a_i \quad (646)$$

Quite frequently we are forced to include constraints on the support of the parameters: some are non-negative (masses, energies, momentum, life-times,...), some are bounded in  $(0, 1)$  ( $\beta = v/c$ , efficiencies, acceptances,...),... At least from a formal point of view, to account for constraints on the support is a trivial problem. Consider the model  $p(\mathbf{x}|\theta)$  with  $\theta \in \Theta_0$  and a reference prior  $\pi_0(\theta)$ . Then, our inferences on  $\theta$  shall be based on the posterior

$$p(\theta|\mathbf{x}) = \frac{p(\mathbf{x}|\theta) \pi_0(\theta)}{\int_{\Theta_0} p(\mathbf{x}|\theta) \pi_0(\theta) d\theta} \quad (647)$$

Now, if we require that  $\theta \in \Theta \subset \Theta_0$  we define

$$g_1(\theta) = \mathbf{1}_{\Theta}(\theta) \quad \longrightarrow \quad \int_{\Theta_0} g_1(\theta) \pi(\theta) d\theta = \int_{\Theta} \pi(\theta) d\theta = 1 - \epsilon \quad (648)$$

$$g_2(\theta) = \mathbf{1}_{\Theta^c}(\theta) \quad \longrightarrow \quad \int_{\Theta_0} g_2(\theta) \pi(\theta) d\theta = \int_{\Theta^c} \pi(\theta) d\theta = \epsilon \quad (649)$$

and in the limit  $\epsilon \rightarrow 0$  we have the *restricted reference prior*

$$\pi(\theta) = \frac{\pi_0(\theta)}{\int_{\Theta} \pi_0(\theta) d\theta} \mathbf{1}_{\Theta}(\theta) \quad (650)$$

as we have obviously expected. Therefore

$$p(\theta|\mathbf{x}, \theta \in \Theta) = \frac{p(\mathbf{x}|\theta) \pi(\theta)}{\int_{\Theta} p(\mathbf{x}|\theta) \pi(\theta) d\theta} = \frac{p(\mathbf{x}|\theta) \pi_0(\theta)}{\int_{\Theta} p(\mathbf{x}|\theta) \pi_0(\theta) d\theta} \mathbf{1}_{\Theta}(\theta) \quad (651)$$

that is, the same initial expression but normalized in the domain of interest  $\Theta$ .

### 3.10 Decision Problems

Even though all the information we have on the parameters of relevance is contained in the posterior density it is interesting, as we saw in section 1, to explicit some particular values that characterize the probability distribution. This certainly entails a considerable and unnecessary reduction of the available information but in the end, quoting Lord Kelvin, “... when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind”. In statistics, to specify a particular value of the parameter is termed *Point Estimation* and can be formulated in the framework of *Decision Theory*.

In general, *Decision Theory* studies how to choose the *optimal action* among several possible alternatives based on what has been experimentally observed. Given a particular problem, we have to explicit the set  $\Omega_{\theta}$  of the possible “states of nature”, the set  $\Omega_X$  of the possible experimental outcomes and the set  $\Omega_A$  of the possible actions we can take. Imagine, for instance, that we do a test on an individual suspected to have some disease for which the medical treatment has some potentially dangerous collateral effects. Then, we have:

$$\begin{aligned} \Omega_{\theta} &= \{ \text{healthy, sic} \} \\ \Omega_X &= \{ \text{test positive, test negative} \} \\ \Omega_A &= \{ \text{apply treatment, do not apply treatment} \} \end{aligned}$$

Or, for instance, a detector that provides within some accuracy the momentum ( $p$ ) and the velocity ( $\beta$ ) of charged particles. If we want to assign an hypothesis for the mass of the particle we have that  $\Omega_{\theta} = \mathcal{R}^+$  is the set of all possible states of nature (all possible values of the mass),  $\Omega_X$  the set of experimental observations (the momentum and the velocity) and  $\Omega_A$  the set of all possible actions that we can take (assign one or other value for the mass). In this case, we shall take a decision based on the probability density  $p(m|p, \beta)$ .

Obviously, unless we are in a state of absolute certainty we can not take an action without potential losses. Based on the observed experimental outcomes, we can for instance assign the particle a mass  $m_1$  when the *true state of nature* is  $m_2 \neq m_1$  or consider that the individual is healthy when is actually sic. Thus, the first element of Decision Theory is the *Loss Function*:

$$l(a, \theta) : (\theta, a) \in \Omega_\theta \times \Omega_A \longrightarrow \mathcal{R}^+ + \{0\} \quad (652)$$

This is a non-negative function, defined for all  $\theta \in \Omega_\theta$  and the set of possible actions  $a \in \Omega_A$ , that quantifies the *loss* associated to take the action  $a$  (decide for  $a$ ) when the state of nature is  $\theta$ .

Obviously, we do not have a perfect knowledge of the *state of nature*; what we know comes from the observed data  $\mathbf{x}$  and is contained in the posterior distribution  $p(\theta|\mathbf{x})$ . Therefore, we define the *Risk Function* (*risk* associated to take the action  $a$ , or decide for  $a$  when we have observed the data  $\mathbf{x}$ ) as the expected value of the Loss Function:

$$R(a|\mathbf{x}) = E_\theta[l(a, \theta)] = \int_{\Omega_\theta} l(a, \theta) p(\theta|\mathbf{x}) d\theta \quad (653)$$

Sound enough, the Bayesian decision criteria consists on taking the action  $a(\mathbf{x})$  (*Bayesian action*) that minimizes the risk  $R(a|\mathbf{x})$  (*minimum risk*); that is, that minimizes the expected loss under the posterior density function<sup>22</sup>. Then, we shall encounter to kinds of problems:

- *inferential problems*, where  $\Omega_A = \mathcal{R}$  y  $a(\mathbf{x})$  is a statistic that we shall take as *estimator* of the parameter  $\theta$ ;
- *decision problems* (or *hypothesis testing*) where  $\Omega_A = \{ \text{accept, reject} \}$  or choose one among a set of hypothesis.

Obviously, the actions depend on the loss function (that we have to specify) and on the posterior density and, therefore, on the data through the model  $p(\mathbf{x}|\theta)$  and the prior function  $\pi(\theta)$ . It is then possible that, for a particular model, two different loss functions drive to the same decision or that the same loss function, depending on the prior, take to different actions.

### 3.10.1 Hypothesis Testing

Consider the case where we have to choose between two exclusive and exhaustive hypothesis  $H_1$  and  $H_2 (= H_1^c)$ . From the data sample and our prior beliefs we have the posterior probabilities

$$P(H_i | \text{data}) = \frac{P(\text{data}|H_i) P(H_i)}{P(\text{data})} \quad ; \quad i = 1, 2 \quad (654)$$

and the actions to be taken are then:

$a_1$ : action to take if we decide upon  $H_1$

$a_2$ : action to take if we decide upon  $H_2$

Then, we define the loss function  $l(a_i, H_j); i, j = 1, 2$  as:

<sup>22</sup>The problems studied by *Decision Theory* can be addressed from the point of view of *Game Theory*. In this case, instead of *Loss Functions* one works with *Utility Functions*  $u(\theta, \mathbf{a})$  that, in essence, are nothing else but  $u(\theta, \mathbf{a}) = K - l(\theta, \mathbf{a}) \geq 0$ ; it is just matter of personal optimism to work with "*utilities*" or "*losses*". J. Von Neumann and O. Morgenstern introduced in 1944 the idea of *expected utility* and the criteria to take as *optimal action* hat which maximizes the expected utility.

$$l(a_i|H_j) = \begin{cases} l_{11} = l_{22} = 0 & \text{if we make the correct choice; that is, if we} \\ & \text{take action } a_1 \text{ when the state of nature is } H_1 \\ & \text{or } a_2 \text{ when it is } H_2; \\ l_{12} > 0 & \text{if we take action } a_1 \text{ (decide upon } H_1\text{) when} \\ & \text{the state of nature is } H_2 \\ l_{21} > 0 & \text{if we take action } a_2 \text{ (decide upon } H_2\text{) when} \\ & \text{the state of nature is } H_1 \end{cases}$$

so the risk function will be:

$$R(a_i | \text{data}) = \sum_{j=1}^2 l(a_i|H_j) P(H_j | \text{data}) \quad (655)$$

that is:

$$R(a_1 | \text{data}) = l_{11} P(H_1 | \text{data}) + l_{12} P(H_2 | \text{data}) \quad (656)$$

$$R(a_2 | \text{data}) = l_{21} P(H_1 | \text{data}) + l_{22} P(H_2 | \text{data}) \quad (657)$$

and, according to the minimum Bayesian risk, we shall choose the hypothesis  $H_1$  (action  $a_1$ ) if

$$R(a_1 | \text{data}) < R(a_2 | \text{data}) \quad \longrightarrow \quad P(H_1 | \text{data}) (l_{11} - l_{21}) < P(H_2 | \text{data}) (l_{22} - l_{12}) \quad (658)$$

Since we have chosen  $l_{11} = l_{22} = 0$  in this particular case, we shall take action  $a_1$  (decide for hypothesis  $H_1$ ) if:

$$\frac{P(H_1 | \text{data})}{P(H_2 | \text{data})} > \frac{l_{12}}{l_{21}} \quad (659)$$

or action  $a_2$  (decide in favor of hypothesis  $H_2$ ) if:

$$R(a_2, \text{data}) < R(a_1, \text{data}) \quad \longrightarrow \quad \frac{P(H_2 | \text{data})}{P(H_1 | \text{data})} > \frac{l_{21}}{l_{12}} \quad (660)$$

that is, we take action  $a_i$  ( $i = 1, 2$ ) if:

$$\frac{P(H_i | \text{data})}{P(H_j | \text{data})} = \left[ \frac{P(\text{data}|H_i)}{P(\text{data}|H_j)} \right] \left[ \frac{P(H_i)}{P(H_j)} \right] > \frac{l_{ij}}{l_{ji}} \quad (661)$$

The ratio of likelihoods

$$B_{ij} = \frac{P(\text{data}|H_i)}{P(\text{data}|H_j)} \quad (662)$$

is called **Bayes Factor**  $B_{ij}$  and changes our prior beliefs on the two alternative hypothesis based on the evidence we have from the data; that is, quantifies how strongly data favors one model over the other. Thus, we shall decide in favor of hypothesis  $H_i$  against  $H_j$  ( $i, j = 1, 2$ ) if

$$\frac{P(H_i | \text{data})}{P(H_j | \text{data})} > \frac{l_{ij}}{l_{ji}} \quad \longrightarrow \quad B_{ij} > \frac{P(H_j)}{P(H_i)} \frac{l_{ij}}{l_{ji}} \quad (663)$$

If we consider the same loss if we decide upon the wrong hypothesis whatever it be, we have  $l_{12} = l_{21}$  (Zero-One Loss Function). In general, we shall be interested in testing:

- 1) **Two simple hypothesis**,  $H_1$  vs  $H_2$ , for which the models  $M_i = \{X \sim p_i(x|\theta_i)\}$ ;  $i = 1, 2$  are fully specified including the values of the parameters (that is,  $\Theta_i = \{\theta_i\}$ ). In this case, the Bayes Factor will be given by the ratio of likelihoods

$$B_{12} = \frac{p_1(\mathbf{x}|\theta_1)}{p_2(\mathbf{x}|\theta_2)} \quad \left( \text{usually } \frac{p(\mathbf{x}|\theta_1)}{p(\mathbf{x}|\theta_2)} \right) \quad (664)$$

The classical Bayes Factor is the ratio of the likelihoods for the two competing models evaluated at their respective maximums.

- 2) **A simple ( $H_1$ ) vs a composite hypothesis  $H_2$**  for which the parameters of the model  $M_2 = \{X \sim p_2(x|\theta_2)\}$  have support on  $\Theta_2$ . Then we have to average the likelihood under  $H_2$  and

$$B_{12} = \frac{p_1(\mathbf{x}|\theta_1)}{\int_{\Theta_2} p_2(\mathbf{x}|\theta)\pi_2(\theta)d\theta} \quad (665)$$

- 3) **Two composite hypothesis**: in which the models  $M_1$  and  $M_2$  have parameters that are not specified by the hypothesis so

$$B_{12} = \frac{\int_{\Theta_1} p_1(\mathbf{x}|\theta_1)\pi_1(\theta_1)d\theta_1}{\int_{\Theta_2} p_2(\mathbf{x}|\theta_2)\pi_2(\theta_2)d\theta_2} \quad (666)$$

and, since  $P(H_1 | \text{data}) + P(H_2 | \text{data}) = 1$ , we can express the posterior probability  $P(H_1 | \text{data})$  as

$$P(H_1 | \text{data}) = \frac{B_{12} P(H_1)}{P(H_2) + B_{12} P(H_1)} \quad (667)$$

Usually, we consider equal prior probabilities for the two hypothesis ( $P(H_1) = P(H_2) = 1/2$ ) but be aware that in some cases this may not be a realistic assumption.

Bayes Factors are independent of the prior beliefs on the hypothesis ( $P(H_i)$ ) but, when we have composite hypothesis, we average the likelihood with a prior and if it is an improper function they are not well defined. If we have prior knowledge about the parameters, we may take informative priors that are proper but this is not always the case. One possible way out is to consider sufficiently general proper priors (conjugated priors for instance) so the Bayes factors are well defined and then study what is the sensitivity for different reasonable values of the hyperparameters. A more practical and interesting approach to avoid the indeterminacy due to improper priors [?], [?] is to take a subset of the observed sample to render a proper posterior (with, for instance, reference priors) and use that as proper prior density to compute the Bayes Factor with the remaining sample. Thus, if the sample  $\mathbf{x} = \{x_1, \dots, x_n\}$  consists on iid observations, we may consider  $\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2\}$  and, with the reference prior  $\pi(\theta)$ , obtain the proper posterior

$$\pi(\theta|\mathbf{x}_1) = \frac{p(\mathbf{x}_1|\theta)\pi(\theta)}{\int_{\Theta} p(\mathbf{x}_1|\theta)\pi(\theta)d\theta} \quad (668)$$

The remaining subsample ( $\mathbf{x}_2$ ) is then used to compute the partial Bayes Factor <sup>23</sup>:

$$B_{12}(x_2|x_1) = \frac{\int_{\Theta_1} p_1(\mathbf{x}_2|\theta_1)\pi_1(\theta_1|\mathbf{x}_1)d\theta_1}{\int_{\Theta_2} p_2(\mathbf{x}_2|\theta_2)\pi_2(\theta_2|\mathbf{x}_1)d\theta_2} \quad \left( = \frac{BF(\mathbf{x}_1, \mathbf{x}_2)}{BF(\mathbf{x}_1)} \right) \quad (669)$$

for the hypothesis testing. Berger and Pericchi propose to use the minimal amount of data needed to specify a proper prior (usually  $\max\{\dim(\theta_i)\}$ ) so as to leave most of the sample for the model testing and dilute the dependence on a particular election of the training sample evaluating the Bayes Factors

<sup>23</sup>Essentially, the ratio of the predictive inferences for  $\mathbf{x}_2$  after  $\mathbf{x}_1$  has been observed.

with all possible minimal samples and choosing the truncated mean, the geometric mean or the median, less sensitive to outliers, as a characteristic value (see example 3.24). A thorough analysis of Bayes Factors, with its caveats and advantages, is given in [?].

A different alternative to quantify the evidence in favour of a particular model that avoids the need of the prior specification and is easy to evaluate is the Schwarz criteria [?] (or "*Bayes Information Criterion (BIC)*"). The rationale is the following. Consider a sample  $\mathbf{x} = \{x_1, \dots, x_n\}$  and two alternative hypothesis for the models  $M_i = \{p_i(x|\theta_i); \dim(\theta_i) = d_i\}; i = 1, 2$ . Under the appropriate conditions we can approximate the likelihood as

$$l(\theta|\mathbf{x}) \simeq l(\hat{\theta}|\mathbf{x}) \exp \left\{ -\frac{1}{2} \sum_{k=1}^d \sum_{m=1}^d (\theta_k - \hat{\theta}_k) \left[ n \mathbf{I}_{km}(\hat{\theta}) \right] (\theta_m - \hat{\theta}_m) \right\} \quad (670)$$

so taking a uniform prior for the parameters  $\theta$ , reasonable in the region where the likelihood is dominant, we can approximate

$$J(\mathbf{x}) = \int_{\Theta} p(\mathbf{x}|\theta) \pi(\theta) d\theta \simeq p(\mathbf{x}|\hat{\theta}) (2\pi/n)^{d/2} |\det[\mathbf{I}(\hat{\theta})]|^{-1/2} \quad (671)$$

and, ignoring terms that are bounded as  $n \rightarrow \infty$ , define the  $BIC(M_i)$  for the model  $M_i$  as

$$2 \ln J_i(\mathbf{x}) \simeq BIC(M_i) \equiv 2 \ln p_i(\mathbf{x}|\hat{\theta}_i) - d_i \ln n \quad (672)$$

so:

$$B_{12} \simeq \frac{p_1(\mathbf{x}|\hat{\theta}_1)}{p_2(\mathbf{x}|\hat{\theta}_2)} n^{(d_2-d_1)/2} \longrightarrow \Delta_{12} = 2 \ln B_{12} \simeq 2 \ln \left( \frac{p_1(\mathbf{x}|\hat{\theta}_1)}{p_2(\mathbf{x}|\hat{\theta}_2)} \right) - (d_1 - d_2) \ln n \quad (673)$$

and therefore, larger values of  $\Delta_{12} = BIC(M_1) - BIC(M_2)$  indicate a preference for the hypothesis  $H_1(M_1)$  against  $H_2(M_2)$  being commonly accepted that for values greater than 6 the evidence is "*strong*"<sup>24</sup> although, in some cases, it is worth to study the behaviour with a Monte Carlo sampling. Note that the last term penalises models with larger number of parameters and that this quantification is sound when the sample size  $n$  is much larger than the dimensions  $d_i$  of the parameters.

**Example 3.23:** Suppose that from the information provided by a detector we estimate the mass of an incoming particle and we want to decide upon the two exclusive and alternative hypothesis  $H_1$  (particle of type 1) and  $H_2 (= H_1^c)$  (particle of type 2). We know from calibration data and Monte Carlo simulations that the mass distributions for both hypothesis are, to a very good approximation, Normal with means  $m_1$  and  $m_2$  variances  $\sigma_1^2$  and  $\sigma_2^2$  respectively. Then for an observed value of the mass  $m_0$  we have:

$$B_{12} = \frac{p(m_0|H_1)}{p(m_0|H_2)} = \frac{N(m_0|m_1, \sigma_1)}{N(m_0|m_2, \sigma_2)} = \frac{\sigma_2}{\sigma_1} \exp \left\{ \frac{(m_0 - m_2)^2}{2\sigma_2^2} - \frac{(m_0 - m_1)^2}{2\sigma_1^2} \right\} \quad (674)$$

Taking ( $l_{12} = l_{21}; l_{11} = l_{22} = 0$ ), the Bayesian decision criteria in favor of the hypothesis  $H_1$  is:

$$B_{12} > \frac{P(H_2)}{P(H_1)} \longrightarrow \ln B_{12} > \ln \frac{P(H_2)}{P(H_1)} \quad (675)$$

Thus, we have a critical value  $m_c$  of the mass:

$$\sigma_1^2 (m_c - m_2)^2 - \sigma_2^2 (m_c - m_1)^2 = 2 \sigma_1^2 \sigma_2^2 \ln \left( \frac{P(H_2) \sigma_1}{P(H_1) \sigma_2} \right) \quad (676)$$

such that, if  $m_0 < m_c$  we decide in favor of  $H_1$  and for  $H_2$  otherwise. In the case that  $\sigma_1 = \sigma_2$  and  $P(H_1) = P(H_2)$ , then  $m_c = (m_1 + m_2)/2$ . This, however, may be a quite unrealistic assumption for if  $P(H_1) > P(H_2)$ , it may be more likely that the event is of type 1 being  $B_{12} < 1$ .

<sup>24</sup>If  $P(H_1) = P(H_2) = 1/2$ , then  $P(H_1 | \text{data}) = 0.95 \longrightarrow B_{12} = 19 \longrightarrow \Delta_{12} \simeq 6$ .

**Example 3.24:** Suppose we have an iid sample  $\mathbf{x} = \{x_1, \dots, x_n\}$  of size  $n$  with  $X \sim N(x|\mu, 1)$  and the two hypothesis  $H_1 = \{N(x|0, 1)\}$  and  $H_2 = \{N(x|\mu, 1); \mu \neq 0\}$ . Let us take  $\{x_i\}$  as the minimum sample and, with the usual constant prior, consider the proper posterior

$$\pi(\mu|x_i) = \frac{1}{\sqrt{2\pi}} \exp\{-(\mu - x_i)^2/2\} \tag{677}$$

that we use as a prior for the rest of the sample  $\mathbf{x}' = \{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n\}$ . Then

$$\frac{P(H_1|\mathbf{x}', x_i)}{P(H_2|\mathbf{x}', x_i)} = B_{12}(i) \frac{P(H_1)}{P(H_2)} \tag{678}$$

where

$$B_{12}(i) = \frac{p(\mathbf{x}'|0)}{\int_{-\infty}^{\infty} p(\mathbf{x}'|\mu)\pi(\mu|x_i)d\mu} = n^{1/2} \exp\{-(n\bar{x}^2 - x_i^2)/2\} \tag{679}$$

and  $\bar{x} = n^{-1} \sum_{k=1}^n x_k$ . To avoid the effect that a particular choice of the minimal sample ( $\{x_i\}$ ) may have, this is evaluated for all possible minimal samples and the median (or the geometric mean) of all the  $B_{12}(i)$  is taken. Since  $P(H_1|\mathbf{x}) + P(H_2|\mathbf{x}) = 1$ , if we assign equal prior probabilities to the two hypothesis ( $P(H_1) = P(H_2) = 1/2$ ) we have that

$$P(H_1|\mathbf{x}) = \frac{B_{12}}{1 + B_{12}} = \left(1 + n^{-1/2} \exp\{(n\bar{x}^2 - \text{med}\{x_i^2\})/2\}\right)^{-1} \tag{680}$$

is the posterior probability that quantifies the evidence in favor of the hypothesis  $H_1$ . It is left as an exercise to compare the Bayes Factor obtained from the geometric mean with what you would get if you were to take a proper prior  $\pi(\mu|\sigma) = N(\mu|0, \sigma)$ .

**Problem 3.11:** Suppose we have  $n$  observations (independent, under the same experimental conditions,...) of energies or decay time of particles above a certain known threshold and we want to test the evidence of an exponential fall against a power law. Consider then a sample  $\mathbf{x} = \{x_1, \dots, x_n\}$  of observations with  $\text{supp}(X) = (1, \infty)$  and the two models

$$M_1 : p_1(x|\theta) = \theta \exp\{-\theta(x - 1)\} \mathbf{1}_{(1, \infty)}(x) \quad \text{and} \quad M_2 : p_2(x|\alpha) = \alpha x^{-(\alpha+1)} \mathbf{1}_{(1, \infty)}(x) \tag{681}$$

that is, Exponential and Pareto with unknown parameters  $\theta$  and  $\alpha$ . Show that for the minimal sample  $\{x_i\}$  and reference priors, the Bayes Factor  $B_{12}(i)$  is given by

$$B_{12}(i) = \left(\frac{x_g \ln x_g}{\bar{x} - 1}\right)^n \left(\frac{x_i - 1}{x_i \ln x_i}\right) = \frac{p_1(\mathbf{x}|\hat{\theta})}{p_2(\mathbf{x}|\hat{\alpha})} \left(\frac{x_i - 1}{x_i \ln x_i}\right) \tag{682}$$

where  $(\bar{x}, x_g)$  are the arithmetic and geometric sample means and  $(\hat{\theta}, \hat{\alpha})$  the values that maximize the likelihoods and therefore

$$\text{med}\{B_{12}(i)\}_{i=1}^n = \left(\frac{x_g \ln x_g}{\bar{x} - 1}\right)^n \text{med}\left\{\frac{x_i - 1}{x_i \ln x_i}\right\}_{i=1}^n \tag{683}$$

**Problem 3.12:** Suppose we have two experiments  $e_i(n_i); i = 1, 2$  in which, out of  $n_i$  trials,  $x_i$  successes have been observed and we are interested in testing whether both treatments are different or not (*contingency tables*). If we assume Binomial models  $Bi(x_i|n_i, \theta_i)$  for both experiments and the two hypothesis  $H_1 : \{\theta_1 = \theta_2\}$  and  $H_2 : \{\theta_1 \neq \theta_2\}$ , the Bayes Factor will be

$$B_{12} = \frac{\int_{\Theta} Bi(x_1|n_1, \theta)Bi(x_2|n_2, \theta)\pi(\theta)d\theta}{\int_{\Theta_1} Bi(x_1|n_1, \theta_1)\pi(\theta_1)d\theta_1 \int_{\Theta_2} Bi(x_2|n_2, \theta)\pi(\theta_2)d\theta_2} \tag{684}$$

We may consider proper Beta prior densities  $Be(\theta|a, b)$ . In a specific pharmacological analysis, a sample of  $n_1 = 52$  individuals were administered a placebo and  $n_2 = 61$  were treated with an a priori beneficial drug. After the essay, positive effects were observed in  $x_1 = 22$  out of the 52 and  $x_2 = 41$  out of the 61 individuals. It is left as an exercise to obtain the posterior probability  $P(H_2|\text{data})$  with Jeffreys' ( $a = b = 1/2$ ) and Uniform ( $a = b = 1$ ) priors and to determine the BIC difference  $\Delta_{12}$ .



### 3.10.2 Point Estimation

When we have to face the problem to characterize the posterior density by a single number, the most usual *Loss Functions* are:

- **Quadratic Loss:** In the simple one-dimensional case, the Loss Function is

$$l(\theta, a) = (\theta - a)^2 \quad (685)$$

so, minimizing the *Risk*:

$$\min \int_{\Omega_\theta} (\theta - a)^2 p(\theta|\mathbf{x}) d\theta \quad \longrightarrow \quad \int_{\Omega_\theta} (\theta - a) p(\theta|\mathbf{x}) d\theta = 0 \quad (686)$$

and therefore  $a = E[\theta]$ ; that is, the posterior mean.

In the  $k$ -dimensional case, if  $\mathcal{A} = \Omega_\theta = \mathcal{R}^k$  we shall take as Loss Function

$$l(\boldsymbol{\theta}, \mathbf{a}) = (\mathbf{a} - \boldsymbol{\theta})^T \mathbf{H} (\mathbf{a} - \boldsymbol{\theta}) \quad (687)$$

where  $\mathbf{H}$  is a positive defined symmetric matrix. It is clear that:

$$\min \int_{\mathcal{R}^k} (\mathbf{a} - \boldsymbol{\theta})^T \mathbf{H} (\mathbf{a} - \boldsymbol{\theta}) p(\boldsymbol{\theta}|\mathbf{x}) d\boldsymbol{\theta} \quad \longrightarrow \quad \mathbf{H} \mathbf{a} = \mathbf{H} E[\boldsymbol{\theta}] \quad (688)$$

so, if  $\mathbf{H}^{-1}$  exists, then  $\mathbf{a} = E[\boldsymbol{\theta}]$ . Thus, we have that the Bayesian estimate under a quadratic loss function is the mean of  $p(\boldsymbol{\theta}|\mathbf{x})$  (... if exists!).

- **Linear Loss:** If  $\mathcal{A} = \Omega_\theta = \mathcal{R}$ , we shall take the loss function:

$$l(\theta, a) = c_1 (a - \theta) \mathbf{1}_{\theta \leq a} + c_2 (\theta - a) \mathbf{1}_{\theta > a} \quad (689)$$

Then, the estimator will be such that

$$\min \int_{\Omega_\theta} l(a, \theta) p(\theta|\mathbf{x}) d\theta = \min \left( c_1 \int_{-\infty}^a (a - \theta) p(\theta|\mathbf{x}) d\theta + c_2 \int_a^{\infty} (\theta - a) p(\theta|\mathbf{x}) d\theta \right) \quad (690)$$

After derivative with respect to  $a$  we have  $(c_1 + c_2) P(\theta \leq a) - c_2 = 0$  and therefore the estimator will be the value of  $a$  such that

$$P(\theta \leq a) = \frac{c_2}{c_1 + c_2} \quad (691)$$

In particular, if  $c_1 = c_2$  then  $P(\theta \leq a) = 1/2$  and we shall have the median of the distribution  $p(\theta|\mathbf{x})$ . In this case, the Loss Function can be expressed more simply as  $l(\theta, a) = |\theta - a|$ .

- **Zero-One Loss:** Si  $\mathcal{A} = \Omega_\theta = \mathcal{R}^k$ , we shall take the Loss Function

$$l(\boldsymbol{\theta}, \mathbf{a}) = 1 - \mathbf{1}_{\mathcal{B}_\epsilon(\mathbf{a})} \quad (692)$$

where  $\mathcal{B}_\epsilon(\mathbf{a}) \in \Omega_\theta$  is an open ball of radius  $\epsilon$  centered at  $\mathbf{a}$ . The corresponding point estimator will be:

$$\min \int_{\Omega_\theta} (1 - \mathbf{1}_{\mathcal{B}_\epsilon(\mathbf{a})}) p(\boldsymbol{\theta}|\mathbf{x}) d\boldsymbol{\theta} = \max \int_{\mathcal{B}_\epsilon(\mathbf{a})} p(\boldsymbol{\theta}|\mathbf{x}) d\boldsymbol{\theta} \quad (693)$$

It is clear than, in the limit  $\epsilon \rightarrow 0$ , the Bayesian estimator for the Zero-One Loss Function will be the mode of  $p(\boldsymbol{\theta}|\mathbf{x})$  if exists.

As explained in section 1, the mode, the median and the mean can be very different if the distribution is not symmetric. Which one should we take then? Quadratic losses, for which large deviations from the *true* value are penalized quadratically, are the most common option but, even if for unimodal symmetric the three statistics coincide, it may be misleading to take this value as a characteristic number for the information we got about the parameters or even be nonsense. In the hypothetical case that the posterior is essentially the same as the likelihood (that is the case for a sufficiently smooth prior), the Zero-One Loss points to the classical estimate of the *Maximum Likelihood Method*. Other considerations of interest in Classical Statistics (like bias, consistency, minimum variance,...) have no special relevance in Bayesian inference.

**Problem 3.13: The Uniform Distribution.** Show that for the posterior density (see example 3.4)

$$p(\theta|x_M, n) = n \frac{x_M^n}{\theta^{n+1}} \mathbf{1}_{[x_M, \infty)}(\theta) \quad (694)$$

the point estimates under quadratic, linear and 0-1 loss functions are

$$\theta_{QL} = x_M \frac{n}{n-1} \quad ; \quad \theta_{LL} = x_M 2^{1/n} \quad \text{and} \quad \theta_{01L} = x_M \quad (695)$$

and discuss which one you consider more reasonable.

### 3.11 Credible Regions

Let  $p(\theta|\mathbf{x})$ , with  $\theta \in \Omega \subseteq \mathcal{R}^n$  be a posterior density function. A credible region with probability content  $1 - \alpha$  is a region of  $V_\alpha \subseteq \Theta$  of the parametric space such that

$$P(\theta \in V_\alpha) = \int_{V_\alpha} p(\theta|\mathbf{x}) d\theta = 1 - \alpha \quad (696)$$

Obviously, for a given probability content credible regions are not unique and a sound criteria is to specify the one that the smallest possible volume. A region  $C$  of the parametric space  $\Omega$  is called *Highest Probability Region* (HPD) with probability content  $1 - \alpha$  if:

- 1)  $P(\theta \in C) = 1 - \alpha$ ;  $C \subseteq \Omega$ ;
- 2)  $p(\theta_1|\cdot) \geq p(\theta_2|\cdot)$  for all  $\theta_1 \in C$  and  $\theta_2 \notin C$  except, at most, for a subset of  $\Omega$  with zero probability measure.

It is left as an exercise to show that condition 2) implies that the HPD region so defined is of minimum volume so both definitions are equivalent. Further properties that are easy to demonstrate are:

- 1) If  $p(\theta|\cdot)$  is *not uniform*, the HPD region with probability content  $1 - \alpha$  is *unique*;
- 2) If  $p(\theta_1|\cdot) = p(\theta_2|\cdot)$ , then  $\theta_1$  and  $\theta_2$  are both either included or excluded of the HPD region;
- 3) If  $p(\theta_1|\cdot) \neq p(\theta_2|\cdot)$ , there is an HPD region for some value of  $1 - \alpha$  that contains one value of  $\theta$  and not the other;
- 4)  $C = \{\theta \in \Theta | p(\theta|\mathbf{x}) \geq k_\alpha\}$  where  $k_\alpha$  is the largest constant for which  $P(\theta \in C) \geq \alpha$ ;
- 5) If  $\phi = f(\theta)$  is a one-to-one transformation, then
  - a) any region with probability content  $1 - \alpha$  for  $\theta$  will have probability content  $1 - \alpha$  for  $\phi$  but...
  - b) an HPD region for  $\theta$  will not, in general, be an HPD region for  $\phi$  unless the transformation is linear.

In general, evaluation of credible regions is a bit messy task. A simple way through is to do a Monte Carlo sampling of the posterior density and use the 4<sup>th</sup> property. For a one-dimensional parameter,

the condition that the HPD region with probability content  $1 - \alpha$  has the minimum length allows to write a relation that may be useful to obtain those regions in an easier manner. Let  $[\theta_1, \theta_2]$  be an interval such that

$$\int_{\theta_1}^{\theta_2} p(\theta|\cdot) d\theta = 1 - \alpha \tag{697}$$

For this to be an HPD region we have to find the extremal of the function

$$\phi(\theta_1, \theta_2, \lambda) = (\theta_2 - \theta_1) + \lambda \left( \int_{\theta_1}^{\theta_2} p(\theta|\cdot) d\theta - (1 - \alpha) \right) \tag{698}$$

Taking derivatives we get:

$$\left( \frac{\partial \phi(\theta_1, \theta_2, \lambda)}{\partial \theta_i} \right)_{i=1,2} = 0 \quad \longrightarrow \quad p(\theta_1|\cdot) = p(\theta_2|\cdot) \tag{699}$$

$$\frac{\partial \phi(\theta_1, \theta_2, \lambda)}{\partial \lambda} = 0 \quad \longrightarrow \quad \int_{\theta_1}^{\theta_2} p(\theta) d\theta = 1 - \alpha \tag{700}$$

Thus, from the first two conditions we have that  $p(\theta_1|\cdot) = p(\theta_2|\cdot)$  and, from the third, we know that  $\theta_1 \neq \theta_2$ . In the special case that the distribution is unimodal and symmetric the only possible solution is  $\theta_2 = 2E[\theta] - \theta_1$ .

The HPD regions are useful to summarize the information on the parameters contained in the posterior density  $p(\theta|\mathbf{x})$  but it should be clear that there is no justification to reject a particular value  $\theta_0$  just because is not included in the HPD region (or, in fact, in whatever confidence region) and that in some circumstances (distributions with more than one mode for instance) it may be the union of disconnected regions.

### 3.12 Bayesian ( $\mathcal{B}$ ) vs Classical ( $\mathcal{F}$ ) Philosophy

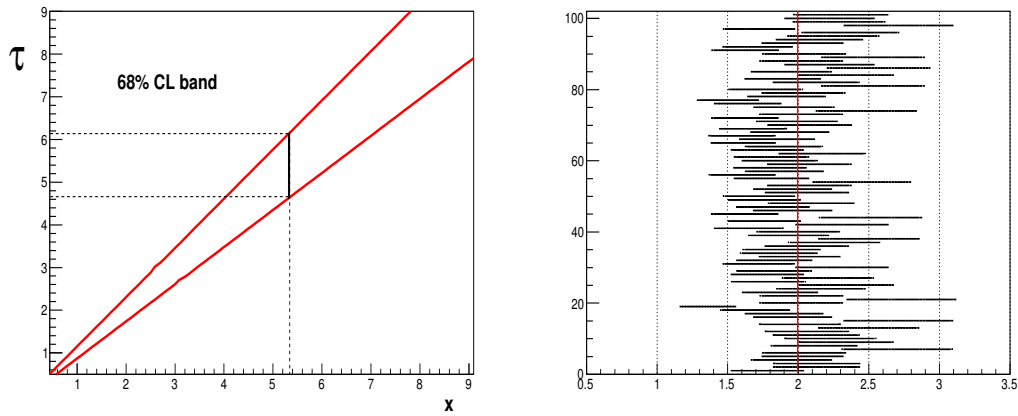
The Bayesian philosophy aims at the right questions in a very intuitive and, at least conceptually, simple manner. However the "classical" (frequentist) approach to statistics, that has been very useful in scientific reasoning over the last century, is at present more widespread in the Particle Physics community and most of the stirred up controversies are originated by misinterpretations. It is worth to take a look for instance at [?]. Let's see how a simple problem is attacked by the two schools. "We" are  $\mathcal{B}$ , "they" are  $\mathcal{F}$ .

Suppose we want to estimate the life-time of a particle. We both "assume" an exponential model  $X \sim Ex(x|1/\tau)$  and do an experiment  $e(n)$  that provides an iid sample  $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ . In this case there is a sufficient statistic  $\mathbf{t} = (n, \bar{x})$  with  $\bar{x}$  the sample mean so let's define the random quantity

$$X = \frac{1}{n} \sum_{i=1}^n X_i \quad \sim \quad p(x|n, \tau) = \left(\frac{n}{\tau}\right)^n \frac{1}{\Gamma(n)} \exp\{-nx\tau^{-1}\} x^{n-1} \mathbf{1}_{(0, \infty)}(x) \tag{701}$$

What can we say about the parameter of interest  $\tau$ ?

$\mathcal{F}$  will start by finding the estimator (statistic)  $\hat{\tau}$  that maximizes the likelihood (MLE). In this case it is clear that  $\hat{\tau} = \bar{x}$ , the sample mean. We may ask about the rationale behind because, apparently, there is no serious mathematical reasoning that justifies this procedure.  $\mathcal{F}$  will respond that, in a certain sense, even for us this should be a reasonable way because if we have a smooth prior function, the posterior is dominated by the likelihood and one possible point estimator is the mode of the posterior. Beside that, he will argue that maximizing the likelihood renders an estimator that often has "good" properties like unbiasedness, invariance under monotonous one-to-one transformations, consistency (convergence



**Fig. 9:** (1): 68% confidence level bands in the  $(\tau, X)$  plane. (2): 68% confidence intervals intervals obtained for 100 repetitions of the experiment.

in probability), smallest variance within the class of unbiased estimators (Cramèr-Rao bound), approximately well known distribution,... We may question some of them (unbiased estimators are not always the best option and invariance... well, if the transformation is not linear usually the MLE is biased), argue that the others hold in the asymptotic limit,... Anyway; for this particular case one has that:

$$E[\hat{\tau}] = \tau \quad \text{and} \quad V[\hat{\tau}] = \frac{\tau^2}{n} \tag{702}$$

and  $\mathcal{F}$  will claim that “if you repeat the experiment” many times under the same conditions, you will get a sequence of estimators  $\{\hat{\tau}_1, \hat{\tau}_2, \dots\}$  that eventually will cluster around the life-time  $\tau$ . Fine but we shall point out that, first, although desirable we usually do not repeat the experiments (and under the same conditions is even more rare) so we have just one observed sample  $(\mathbf{x} \rightarrow \bar{x} = \hat{\tau})$  from  $e(n)$ . Second, “if you repeat the experiment you will get” is a free and unnecessary hypothesis. You do not know what you will get, among other things, because the model we are considering may not be the way nature behaves. Besides that, it is quite unpleasant that inferences on the life-time depend upon what you think you will get if you do what you know you are not going to do. And third, that this is in any case a nice sampling property of the estimator  $\hat{\tau}$  but eventually we are interested in  $\tau$  so, What can we say about it?

For us, the answer is clear. Being  $\tau$  a scale parameter we write the posterior density function

$$p(\tau|n, \bar{x}) = \frac{(n\bar{x})^n}{\Gamma(n)} \exp\{-n\bar{x}\tau^{-1}\} \tau^{-(n+1)} \mathbf{1}_{(0,\infty)}(\tau) \tag{703}$$

for the *degree of belief* we have on the parameter and easily get for instance:

$$E[\tau^k] = (n\bar{x})^k \frac{\Gamma(n-k)}{\Gamma(n)} \quad \longrightarrow \quad E[\tau] = \bar{x} \frac{n}{n-1} \quad ; \quad V[\tau] = \bar{x}^2 \frac{n^2}{(n-1)^2(n-2)} \quad ; \quad \dots \tag{704}$$

Cleaner and simpler impossible.

To bound the life-time,  $\mathcal{F}$  proceeds with the determination of the *Confidence Intervals*. The classical procedure was introduced by J. Neyman in 1933 and rests on establishing, for an specified probability content, the domain of the random quantity (usually a statistic) as function of the possible values the parameters may take. Consider a one dimensional parameter  $\theta$  and the model  $X \sim p(x|\theta)$ . Given a desired probability content  $\beta \in [0, 1]$ , he determines the interval  $[x_1, x_2] \subset \Omega_X$  such that

$$P(X \in [x_1, x_2]) = \int_{x_1}^{x_2} p(x|\theta) dx = \beta \tag{705}$$

for a particular fixed value of  $\theta$ . Thus, for each possible value of  $\theta$  he has one interval  $[x_1 = f_1(\theta; \beta), x_2 = f_2(\theta; \beta)] \subset \Omega_X$  and the sequence of those intervals gives a band in the  $\Omega_\theta \times \Omega_X$  region of the real plane. As for the *Credible Regions*, these intervals are not uniquely determined so one usually adds the condition:

$$1) \quad \int_{-\infty}^{x_1} p(x|\theta) dx = \int_{x_2}^{\infty} p(x|\theta) dx = \frac{1 - \beta}{2} \quad \text{or} \quad (706)$$

$$2) \quad \int_{x_1}^{\theta} p(x|\theta) dx = \int_{\theta}^{x_2} p(x|\theta) dx = \frac{\beta}{2} \quad (707)$$

or, less often, (3) chooses the interval with smallest size. Now, for an invertible mapping  $x_i \rightarrow f_i(\theta)$  one can write

$$\beta = P(f_1(\theta) \leq X \leq f_2(\theta)) = P(f_2^{-1}(X) \leq \theta \leq f_1^{-1}(X)) \quad (708)$$

and get the random interval  $[f_2^{-1}(X), f_1^{-1}(X)]$  that contains the given value of  $\theta$  with probability  $\beta$ . Thus, for each possible value that  $X$  may take he will get an interval  $[f_2^{-1}(X), f_1^{-1}(X)]$  on the  $\theta$  axis and a particular experimental observation  $\{x\}$  will single out one of them. This is the *Confidence Interval* that the frequentist analyst will quote. Let's continue with the life-time example and take, for illustration,  $n = 50$  and  $\beta = 0.68$ . The bands  $[x_1 = f_1(\tau), x_2 = f_2(\tau)]$  in the  $(\tau, X)$  plane, in this case obtained with the third prescription, are shown in figure ?? (1). They are essentially straight lines so  $P[X \in (0.847\tau, 1.126\tau)] = 0.68$ . This is a correct statement, but doesn't say anything about  $\tau$  so he inverts that and gets  $0.89X < \tau < 1.18X$  in such a way that an observed value  $\{x\}$  singles out an interval in the vertical  $\tau$  axis. We, Bayesians, will argue this does not mean that  $\tau$  has a 0.68 chance to lie in this interval and the frequentist will certainly agree on that. In fact, this is not an admissible question for him because in the classical philosophy  $\tau$  is a number, unknown but a *fixed* number. If he repeats the experiment  $\tau$  will not change; it is the interval that will be different because  $x$  will change. They are *random intervals* and what the 68% means is just that if he repeats the experiment a large number  $N$  of times, he will end up with  $N$  intervals of which  $\sim 68\%$  will contain the true value  $\tau$  whatever it is. But the experiment is done only once so: Does the interval derived from this observation contain  $\tau$  or not? We don't know, we have no idea if it does contain  $\tau$ , if it does not and how far is the unknown true value. Figure ?? (2) shows the 68% confidence intervals obtained after 100 repetitions of the experiment for  $\tau = 2$  and 67 of them did contain the true value. But when the experiment is done once, he picks up one of those intervals and has a 68% chance that the one chosen contains the true value. WeB shall proceed in a different manner. After integration of the posterior density we get the HPD interval  $P[\tau \in (0.85x, 1.13x)] = 0.68$ ; almost the same but with a direct interpretation in terms of what we are interested in. Thus, both have an absolutely different philosophy:

$\mathcal{F}$  : "Given a particular value of the parameters of interest, How likely is the observed data?"

$\mathcal{B}$  : "Having observed this data, What can we say about the parameters of interest?"

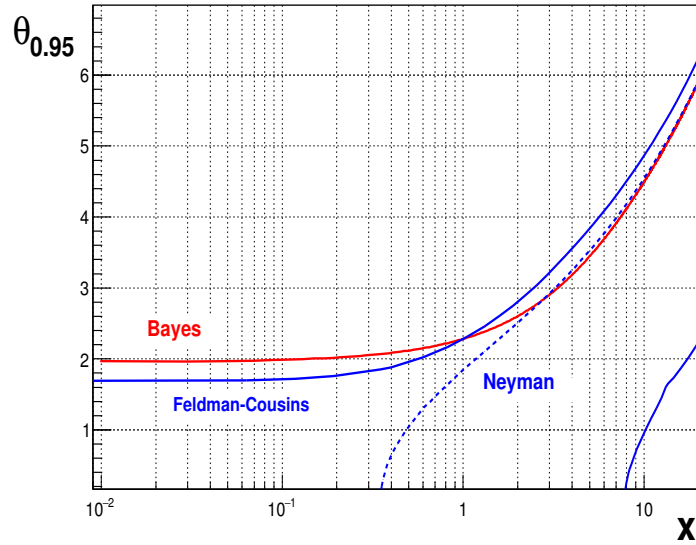
... and the probability if the causes, as Poincare said, is the most important from the point of view of scientific applications.

In many circumstances we are also interested in one-sided intervals. That is for instance the case when the data is consistent with the hypothesis  $H : \{\theta = \theta_0\}$  and we want to give an upper bound on  $\theta$  so that  $P(\theta \in (-\infty, \theta_\beta]) = \beta$ . The frequentist rationale is the same: obtain the interval  $[-\infty, x_2] \subset \Omega_X$  such that

$$P(X \leq x_2) = \int_{-\infty}^{x_2} p(x|\theta) dx = \beta \quad (709)$$

where  $x_2 = f_2(\theta)$ ; in this case without ambiguity. For the the random interval  $(-\infty, f_2^{-1}(X))$   $\mathcal{F}$  has that

$$P(\theta < f_2^{-1}(X)) = 1 - P(\theta \geq f_2^{-1}(X)) = 1 - \beta \quad (710)$$



**Fig. 10:** 95% upper bounds on the parameter  $\theta$  following the Bayesian approach (red), the Neyman approach (broken blue) and Feldman and Cousins (solid blue line).

so, for a probability content  $\alpha$  (say 0.95), one should set  $\beta = 1 - \alpha (=0.05)$ . Now, consider for instance the example of the anisotropy is cosmic rays discussed in the last section 13.3. For a dipole moment (details are unimportant now) we have a statistic

$$X \sim p(x|\theta, 1/2) = \frac{\exp\{-\theta^2/2\}}{\sqrt{2\pi\theta}} \exp\{-x/2\} \sinh(\theta\sqrt{x}) \mathbf{1}_{(0,\infty)}(x) \quad (711)$$

where the parameter  $\theta$  is the dipole coefficient multiplied by a factor that is irrelevant for the example. It is argued in section 13.3 that the reasonable prior for this model is  $\pi(\theta) = \text{constant}$  so we have the posterior

$$p(\theta|x, 1/2) = \frac{\sqrt{2}}{\sqrt{\pi x} M(1/2, 3/2, x/2)} \exp\{-\theta^2/2\} \theta^{-1} \sinh(\theta\sqrt{x}) \mathbf{1}_{(0,\infty)}(\theta) \quad (712)$$

with  $M(a, b, z)$  the Kummer's Confluent Hypergeometric Function. In fact,  $\theta$  has a compact support but since the observed values of  $X$  are consistent with  $H_0 : \{\theta = 0\}$  and the sample size is very large [AMS13]<sup>25</sup>,  $p(\theta|x, 1/2)$  is concentrated in a small interval  $(0, \epsilon)$  and it is easier for the evaluations to extend the domain to  $\mathcal{R}^+$  without any effect on the results. Then we, Bayesians, shall derive the one-sided upper credible region  $[0, \theta_{0.95}(x)]$  with  $\alpha = 95\%$  probability content as simply as:

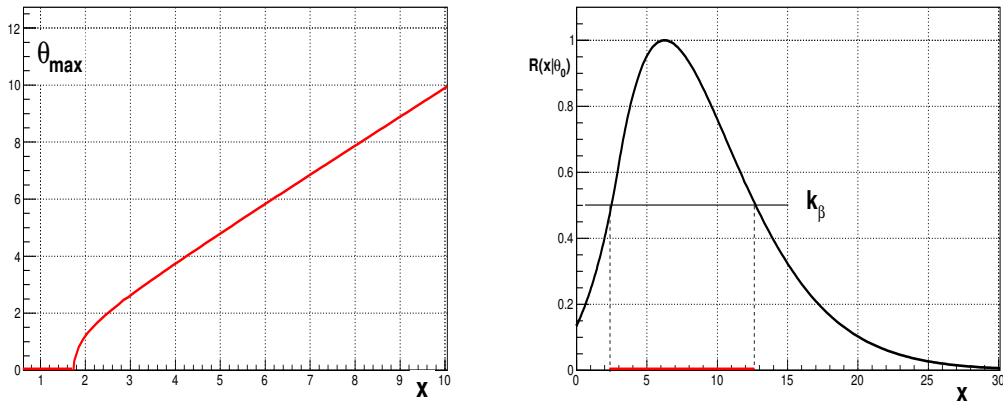
$$\int_0^{\theta_{0.95}} p(\theta|x, 1/2) d\theta = \alpha = 0.95 \quad (713)$$

This upper bound shown as function of  $x$  in figure ?? under "Bayes" (red line). Neyman's construction is also straight forward. From

$$\int_0^{x_2} p(x|\theta, 1/2) dx = 1 - \alpha = 0.05 \quad (714)$$

(essentially a  $\chi^2$  probability for  $\nu = 3$ ),  $\mathcal{F}$  will get the upper bound shown in the same figure under "Neyman" (blue broken line). As you can see, they get closer as  $x$  grows but, first, there is no solution

<sup>25</sup>[AMS13]: Aguilar M. et al. (2013); Phys. Rev. Lett. 110, 141102.



**Fig. 11:** (1) Dependence of  $\theta_m$  with  $x$ . (2) Probability density ratio  $R(x|\theta)$  for  $\theta = 2$ .

for  $x \leq x_c = 0.352$ . In fact,  $E[X] = \theta^2 + 3$  so if the dipole moment is  $\delta = 0$  ( $\theta = 0$ ),  $E[X] = 3$  and observed values below  $x_c$  will be an unlikely fluctuation downwards (assuming of course that the model is correct) but certainly a possible experimental outcome. In fact, you can see that for values of  $x$  less than 2, even though there is a solution Neyman's upper bound is underestimated. To avoid this "little" problem, a different prescription has to be taken.

The most interesting solution is the one proposed by Feldman and Cousins [?] in which the region  $\Delta_X \subset \Omega_X$  that is considered for the specified probability content is determined by the ratio of probability densities. Thus, for a given value  $\theta_0$ , the interval  $\Delta_X$  is such that

$$\int_{\Delta_X} p(x|\theta_0) dx = \beta \quad \text{with} \quad R(x|\theta_0) = \frac{p(x|\theta_0)}{p(x|\theta_b)} > k_\beta \quad ; \forall x \in \Delta_X \quad (715)$$

and where  $\theta_b$  is the best estimation of  $\theta$  for a given  $\{x\}$ ; usually the one that maximizes the likelihood ( $\theta_m$ ). In our case, it is given by:

$$\theta_m = \begin{cases} 0 & \text{if } x \leq \sqrt{3} \\ \theta_m + \theta_m^{-1} - \sqrt{x} \coth(\theta_m \sqrt{x}) = 0 & \text{if } x > \sqrt{3} \end{cases} \quad (716)$$

and the dependence with  $x$  is shown in figure ?? (1) ( $\theta_m \simeq x$  for  $x \gg$ ). As illustration, function  $R(x|\theta)$  is shown in figure ?? (2) for the particular value  $\theta_0 = 2$ . Following this procedure<sup>26</sup>, the 0.95 probability content band is shown in figure ?? under "Feldman-Cousins" (blue line). Note that for large values of  $x$ , the confidence region becomes an interval. It is true that if we observe a large value of  $X$ , the hypothesis  $H_0 : \{\delta = 0\}$  will not be favoured by the data and a different analysis will be more relevant although, by a simple modification of the ordering rule, we still can get an upper bound if desired or use the standard Neyman's procedure.

The Feldman and Cousins prescription allows to consider constrains on the parameters in a simpler way than Neyman's procedure and, as opposed to it, will always provide a region with the specified probability content. However, on the one hand, they are frequentist intervals and as such have to be interpreted. On the other hand, for discrete random quantities with image in  $\{x_1, \dots, x_k, \dots\}$  it may not be possible to satisfy exactly the probability content equation since for the Distribution Function one has that  $F(x_{k+1}) = F(x_k) + P(X = x_{k+1})$ . And last, it is not straight forward to deal with nuisance parameters. Therefore, the best advice: "Be Bayesian!".

<sup>26</sup>In most cases, a Monte Carlo simulation will simplify life.

### 3.13 Some worked examples

#### 3.13.1 Regression

Consider the exchangeable sequence  $\mathbf{z} = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$  of  $n$  samplings from the two-dimensional model  $N(x_i, y_i | \cdot) = N(x_i | \mu_{x_i}, \sigma_{x_i}^2) N(y_i | \mu_{y_i}, \sigma_{y_i}^2)$ . Then

$$p(\mathbf{z} | \cdot) \propto \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \left[ \frac{(y_i - \mu_{y_i})^2}{\sigma_{y_i}^2} + \frac{(x_i - \mu_{x_i})^2}{\sigma_{x_i}^2} \right] \right\} \quad (717)$$

We shall assume that the precisions  $\sigma_{x_i}$  and  $\sigma_{y_i}$  are known and that there is a functional relation  $\mu_y = f(\mu_x; \theta)$  with unknown parameters  $\theta$ . Then, in terms of the new parameters of interest:

$$p(\mathbf{y} | \cdot) \propto \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \left[ \frac{(y_i - f(\mu_{x_i}; \theta))^2}{\sigma_{y_i}^2} + \frac{(x_i - \mu_{x_i})^2}{\sigma_{x_i}^2} \right] \right\} \quad (718)$$

Consider a linear relation  $f(\mu_x; a, b) = a + b\mu_x$  with  $a, b$  the unknown parameters so:

$$p(\mathbf{z} | \cdot) \propto \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \left[ \frac{(y_i - a - b\mu_{x_i})^2}{\sigma_{y_i}^2} + \frac{(x_i - \mu_{x_i})^2}{\sigma_{x_i}^2} \right] \right\} \quad (719)$$

and assume, in first place, that  $\mu_{x_i} = x_i$  without uncertainty. Then,

$$p(\mathbf{y} | a, b) \propto \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \left[ \frac{(y_i - a - bx_i)^2}{\sigma_{y_i}^2} \right] \right\} \quad (720)$$

There is a set of sufficient statistics for  $(a, b)$ :

$$\mathbf{t} = \{t_1, t_2, t_3, t_4, t_5\} = \left\{ \sum_{i=1}^n \frac{1}{\sigma_i^2}, \sum_{i=1}^n \frac{x_i^2}{\sigma_i^2}, \sum_{i=1}^n \frac{x_i}{\sigma_i^2}, \sum_{i=1}^n \frac{y_i}{\sigma_i^2}, \sum_{i=1}^n \frac{y_i x_i}{\sigma_i^2} \right\} \quad (721)$$

and, after a simple algebra, it is easy to write

$$p(\mathbf{y} | a, b) \propto \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \frac{(a-a_0)^2}{\sigma_a^2} + \frac{(b-b_0)^2}{\sigma_b^2} - 2\rho \frac{(a-a_0)(b-b_0)}{\sigma_a \sigma_b} \right] \right\} \quad (722)$$

where the new statistics  $\{a_0, b_0, \sigma_a, \sigma_b, \rho\}$  are defined as:

$$a_0 = \frac{t_2 t_4 - t_3 t_5}{t_1 t_2 - t_3^2}, \quad b_0 = \frac{t_1 t_5 - t_3 t_4}{t_1 t_2 - t_3^2} \quad (723)$$

$$\sigma_a^2 = \frac{t_2}{t_1 t_2 - t_3^2}, \quad \sigma_b^2 = \frac{t_1}{t_1 t_2 - t_3^2}, \quad \rho = -\frac{t_3}{\sqrt{t_1 t_2}} \quad (724)$$

Both  $(a, b)$  are position parameters so we shall take a uniform prior and in consequence

$$p(a, b | \cdot) = \frac{1}{2\pi\sigma_a\sigma_b\sqrt{1-\rho^2}} e^{\left\{ -\frac{1}{2(1-\rho^2)} \left[ \frac{(a-a_0)^2}{\sigma_a^2} + \frac{(b-b_0)^2}{\sigma_b^2} - 2\rho \frac{(a-a_0)(b-b_0)}{\sigma_a \sigma_b} \right] \right\}} \quad (725)$$

This was obviously expected.

When  $\mu_{x_i}$  are  $n$  unknown parameters, if we take  $\pi(\mu_{x_i}) = \mathbf{1}_{(0,\infty)}(\mu_{x_i})$  and marginalize for  $(a, b)$  we have

$$p(a, b | \cdot) \propto \pi(a, b) \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \frac{(y_i - a - bx_i)^2}{\sigma_{y_i}^2 + b^2 \sigma_{x_i}^2} \right\} \left\{ \prod_{i=1}^n (\sigma_{y_i}^2 + b^2 \sigma_{x_i}^2) \right\}^{-1/2} \quad (726)$$



In general, the expressions one gets for non-linear regression problems are complicated and setting up priors is a non-trivial task but fairly vague priors easy to deal with are usually a reasonable choice. In this case, for instance, one may consider uniform priors or normal densities  $N(\cdot|0, \sigma \gg)$  for both parameters  $(a, b)$  and and sample the proper posterior with a Monte Carlo algorithm (Gibbs sampling will be appropriate).

The same reasoning applies if we want to consider other models or more involved relations with several explanatory variables like  $\theta_i = \sum_{j=1}^k \alpha_j x_{ij}^{b_j}$ . In counting experiments, for example,  $y_i \in \mathcal{N}$  so we may be interested in a Poisson model  $Po(y_i|\mu_i)$  where  $\mu_i$  is parameterized as a simple log-linear form  $\ln(\mu_i) = \alpha_0 + \alpha_1 x_i$  (so  $\mu_i > 0$  for whatever  $\alpha_0, \alpha_1 \in \mathcal{R}$ ). Suppose for instance that we have the sample  $\{(y_i, x_i)\}_{i=1}^n$ . Then:

$$p(\mathbf{y}|\alpha_1, \alpha_2, \mathbf{x}) \propto \prod_{i=1}^n \exp\{-\mu_i\} \mu_i^{y_i} = \exp\left\{\alpha_1 s_1 + \alpha_2 s_2 - e^{\alpha_1} \sum_{i=1}^n e^{\alpha_2 x_i}\right\} \quad (727)$$

where  $s_1 = \sum_{i=1}^n y_i$  and  $s_2 = \sum_{i=1}^n y_i x_i$ . In this case, the Normal distribution  $N(\alpha_i|a_i, \sigma_i)$  with  $\sigma_i \gg$  is a reasonable smooth and easy to handle proper prior density for both parameters. Thus, we get the posterior conditional densities

$$p(\alpha_i|\alpha_j, \mathbf{y}, \mathbf{x}) \propto \exp\left\{-\frac{\alpha_i^2}{2\sigma_i^2} + \alpha_i \left(\frac{a_i}{\sigma_i^2} + s_i\right) - e^{\alpha_1} \sum_{i=1}^n e^{\alpha_2 x_i}\right\} \quad ; \quad i = 1, 2 \quad (728)$$

that are perfectly suited for the Gibbs sampling to be discussed in section 3.

**Example 3.25: Proton Flux in Primary Cosmic Rays.** For energies between  $\sim 20$  and  $\sim 200$  GeV, the flux of protons of the primary cosmic radiation is reasonably well described by a power law  $\phi(r) = cr^\gamma$  where  $r$  is the *rigidity*<sup>27</sup> and  $\gamma = d\ln\phi/ds$ , with  $s = \ln r$ , is the *spectral index*. At lower energies, this dependence is significantly modified by the geomagnetic cut-off and the solar wind but at higher energies, where these effects are negligible, the observations are not consistent with a single power law (figure ?? (1)). One may characterize this behaviour with a simple phenomenological model where the spectral index is no longer constant but has a dependence  $\gamma(s) = \alpha + \beta \tanh[a(s - s_0)]$  such that  $\lim_{s \rightarrow -\infty} \gamma(s) = \gamma_1$  ( $r \rightarrow 0$ ) and  $\lim_{s \rightarrow \infty} \gamma(s) = \gamma_2$  ( $r \rightarrow +\infty$ ). After integration, the flux can be expressed in terms of 5 parameters  $\theta = \{\phi_0, \gamma_1, \delta = \gamma_2 - \gamma_1, r_0, \sigma\}$  as:

$$\phi(r; \theta) = \phi_0 r^{\gamma_1} \left[1 + \left(\frac{r}{r_0}\right)^\sigma\right]^{\delta/\sigma} \quad (729)$$

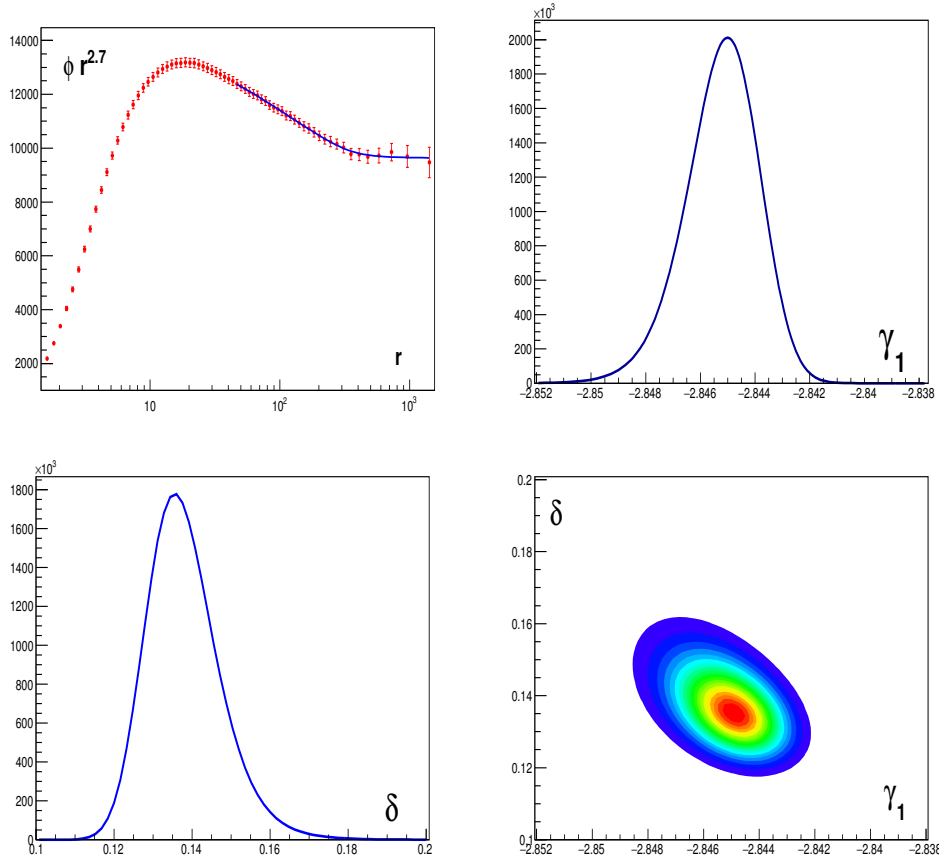
For this example, I have used the data above 45 GeV published by the AMS experiment<sup>28</sup> and considered only the quoted *statistical errors*. Last, for a better description of the flux the previous expression has been modified to account for the effect of the solar wind with the force-field approximation in consistency with [AMS15]. This is just a technical detail, irrelevant for the purpose of the example. Then, assuming a Normal model for the observations we can write the posterior density

$$p(\theta|\text{data}) = \pi(\theta) \prod_{i=1}^n \exp\left\{-\frac{1}{2\sigma_i^2} (\phi_i - \phi(r_i; \theta))^2\right\} \quad (730)$$

I have taken Normal priors with large variances ( $\sigma_i \gg$ ) for the parameters  $\gamma_1$  and  $\delta$  and restricted the support to  $\mathcal{R}^+$  for  $\{\phi_0, r_0, \sigma\}$ . The posterior densities for the parameters  $\gamma_1$  and  $\delta$  are shown in figure ?? (2,3) together with the projection ?? (4) that gives an idea of correlation between them. For a visual inspection, the phenomenological form of the flux is shown in figure ?? (1) (blue line) overimposed to the data when the parameters are set to their expected posterior values.

<sup>27</sup>The *rigidity* ( $r$ ) is defined as the momentum ( $p$ ) divided by the electric charge ( $Z$ ) so  $r = p$  for protons.

<sup>28</sup>[AMS15]: Aguilar M. et al. (2015); PRL 114, 171103 and references therein.



**Fig. 12:** (1): Observed flux multiplied by  $r^{2.7}$  in  $m^{-2}sr^{-1}sec^{-1}GV^{1.7}$  as given in [AMS15]; (2): Posterior density of the parameter  $\gamma_1$  (arbitrary vertical scale); (3): Posterior density of the parameter  $\delta = \gamma_2 - \gamma_1$  (arbitrary vertical scale); (4): Projection of the posterior density  $p(\gamma_1, \delta)$ .

### 3.13.2 Characterization of a Possible Source of Events

Suppose that we observe a particular region  $\Omega$  of the sky during a time  $t$  and denote by  $\lambda$  the rate at which events from this region are produced. We take a Poisson model to describe the number of produced events:  $k \sim Po(k|\lambda t)$ . Now, denote by  $\epsilon$  the probability to detect one event (detection area, efficiency of the detector,...). The number of observed events  $n$  from the region  $\Omega$  after an exposure time  $t$  and detection probability  $\epsilon$  will follow:

$$n \sim \sum_{k=n}^{\infty} Bi(k|n, \epsilon) Po(k|\lambda t) = Po(n|\lambda t \epsilon) \quad (731)$$

The approach to the problem will be the same for other counting process like, for instance, events collected from a detector for a given integrated luminosity. We suspect that the events observed in a particular region  $\Omega_o$  of the sky are background events together with those from an emitting source. To determine the significance of the potential source we analyze a nearby region,  $\Omega_b$ , to infer about the expected background. If after a time  $t_b$  we observe  $n_b$  events from this region with detection probability  $e_b$  then, defining  $\beta = e_b t_b$  we have that

$$n_b \sim Po(n_b|\lambda_b \beta) = \exp\{-\beta \lambda_b\} \frac{(\beta \lambda_b)^{n_b}}{\Gamma(n_b + 1)} \quad (732)$$

At  $\Omega_o$  we observe  $n_o$  events during a time  $t_o$  with a detection probability  $\epsilon_o$ . Since  $n_o = n_1 + n_2$  with  $n_1 \sim Po(n_1|\lambda_s\alpha)$  signal events ( $\alpha = \epsilon_o t_o$ ) and  $n_2 \sim Po(n_2|\lambda_b\alpha)$  background events (assume reasonably that  $e_s = e_b = e_o$  in the same region), we have that

$$n_o \sim \sum_{n_1=0}^{n_o} Po(n_1|\lambda_s\alpha)Po(n_o - n_1|\lambda_b\alpha) = Po(n_o|(\lambda_s + \lambda_b)\alpha) \quad (733)$$

Now, we can do several things. We can assume for instance that the overall rate from the region  $\Omega_o$  is  $\lambda$ , write  $n_o \sim Po(n_o|\alpha\lambda)$  and study the fraction  $\lambda/\lambda_b$  of the rates from the information provided by the observations in the two different regions. Then, reparameterizing the model in terms of  $\theta = \lambda/\lambda_b$  and  $\phi = \lambda_b$  we have

$$p(n_o, n_b|\cdot) = Po(n_o|\alpha\lambda)Po(n_b|\beta\lambda_b) \sim e^{-\beta\phi(1+\gamma\theta)}\theta^{n_o}\phi^{n_o+n_b} \quad (734)$$

where  $\gamma = \alpha/\beta = (\epsilon_s t_s)/(\epsilon_b t_b)$ . For the ordering  $\{\theta, \phi\}$  we have that the Fisher's matrix and its inverse are

$$\mathbf{I}(\theta, \phi) = \begin{pmatrix} \frac{\gamma\beta\phi}{\theta} & \gamma\beta \\ \gamma\beta & \frac{\beta(1+\gamma\theta)}{\phi} \end{pmatrix} \quad \text{and} \quad \mathbf{I}^{-1}(\mu_1, \mu_2) = \begin{pmatrix} \frac{\theta(1+\gamma\theta)}{\phi\gamma\beta} & -\frac{\theta}{\beta} \\ -\frac{\theta}{\beta} & \frac{\phi}{\beta} \end{pmatrix} \quad (735)$$

Then

$$\pi(\theta, \phi) = \pi(\phi|\theta)\pi(\theta) \propto \frac{\phi^{-1/2}}{\sqrt{\theta(1+\gamma\theta)}} \quad (736)$$

and integrating the nuisance parameter  $\phi$  we get finally:

$$p(\theta|n_o, n_b, \gamma) = \frac{\gamma^{n_o+1/2}}{B(n_o+1/2, n_b+1/2)} \frac{\theta^{n_o-1/2}}{(1+\gamma\theta)^{n_o+n_b+1}} \quad (737)$$

From this:

$$E[\theta^m] = \frac{1}{\gamma^m} \frac{\Gamma(n_o+1/2+m)\Gamma(n_b+1/2-m)}{\Gamma(n_o+1/2)\Gamma(n_b+1/2)} \rightarrow E[\theta] = \frac{1}{\gamma} \frac{n_o+1/2}{n_b-1/2} \quad (738)$$

and

$$P(\theta \leq \theta_0) = \int_0^{\theta_0} p(\theta|\cdot) d\theta = 1 - IB(n_b+1/2, n_o+1/2; (1+\gamma\theta_0)^{-1}) \quad (739)$$

with  $IB(x, y; z)$  the Incomplete Beta Function. Had we interest in  $\theta = \lambda_s/\lambda_b$ , the corresponding reference prior will be

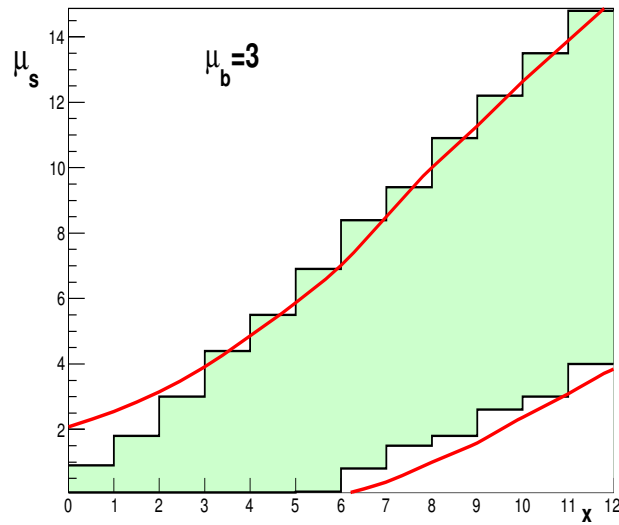
$$\pi(\theta, \phi) \propto \frac{\phi^{-1/2}}{\sqrt{(1+\theta)(\delta+\theta)}} \quad \text{with} \quad \delta = \frac{1+\gamma}{\gamma} \quad (740)$$

A different analysis can be performed to make inferences on  $\lambda_s$ . In this case, we may consider as an informative prior for the nuisance parameter the posterior what we had from the study of the background in the region  $\Omega_b$ ; that is:

$$p(\lambda_b|n_b, \beta) \propto \exp\{-\beta\lambda_b\} \lambda_b^{n_b-1/2} \quad (741)$$

and therefore:

$$p(\lambda_s|\cdot) \propto \pi(\lambda_s) \int_0^\infty p(n_o|\alpha(\lambda_s + \lambda_b)) p(\lambda_b|n_b, \beta) d\lambda_b \propto \pi(\lambda_s) e^{-\alpha\lambda_s} \lambda_s^{n_o} \sum_{k=0}^{n_o} a_k \lambda_s^{-k} \quad (742)$$



**Fig. 13:** 90% Confidence Belt derived with Feldman and Cousins (filled band) and the Bayesian HPD region (red lines) for a background parameter  $\mu_b = 3$ .

where

$$a_k = \binom{n_o}{k} \frac{\Gamma(k + n_b + 1/2)}{[(\alpha + \beta)]^k} \tag{743}$$

A reasonable choice for the prior will be a conjugated prior  $\pi(\lambda_s) = Ga(\lambda_s|a, b)$  that simplifies the calculations and provides enough freedom analyze the effect of different shapes on the inferences. The same reasoning is valid if the knowledge on  $\lambda_b$  is represented by a different  $p(\lambda_b|\cdot)$  from, say, a Monte Carlo simulation. Usual distributions in this case are the Gamma and the Normal with non-negative support. Last, it is clear that if the rate of background events is known with high accuracy then, with  $\mu_i = \alpha\lambda_i$  and  $\pi(\mu_s) \propto (\mu_s + \mu_b)^{-1/2}$  we have

$$p(\mu_s|\cdot) = \frac{1}{\Gamma(x + 1/2, \mu_b)} \exp\{-(\mu_s + \mu_b)\} (\mu_s + \mu_b)^{x-1/2} \mathbf{1}_{(0,\infty)}(\mu_s) \tag{744}$$

As an example, we show in figure ?? the 90% HPD region obtained from the previous expression (red lines) as function of  $x$  for  $\mu_b = 3$  (conditions as given in the example of [?]) and the Confidence Belt derived with the Feldman and Cousins approach (filled band). In this case,  $\mu_{s,m} = \max\{0, x - \mu_b\}$  and therefore, for a given  $\mu_s$ :

$$\sum_{x_1}^{x_2} Po(x|\mu_s + \mu_b) = \beta \quad \text{with} \quad R(x|\mu_s) = e^{(\mu_{s,m} - \mu_s)} \left( \frac{\mu_s + \mu_b}{\mu_{s,m} + \mu_b} \right)^x > k\beta \tag{745}$$

for all  $x \in [x_1, x_2]$ .

**Problem 3.14:** In the search for a new particle, assume that the number of observed events follows a Poisson distribution with  $\mu_b = 0.7$  known with enough precision from extensive Monte Carlo simulations. Consider the hypothesis  $H_0 : \{\mu_s = 0\}$  and  $H_1 : \{\mu_s \neq 0\}$ . It is left as an exercise to obtain the Bayes Factor  $BF_{01}$  with the proper prior  $\pi(\mu_s|\mu_b) = \mu_b(\mu_s + \mu_b)^{-2}$  proposed in [?],  $P(H_1|n)$  and the BIC difference  $\Delta_{01}$  as function of  $n = 1, \dots, 7$  and decide when, based on this results, will you consider that there is evidence for a signal.

### 3.13.3 Anisotropies of Cosmic Rays

The angular distribution of cosmic rays in galactic coordinates is analyzed searching for possible anisotropies. A well-behaved real function  $f(\theta, \phi) \in L_2(\Omega)$ , with  $(\theta, \phi) \in \Omega = [0, \pi] \times [0, 2\pi]$ , can be expressed in the real harmonics basis as:

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l a_{lm} Y_{lm}(\theta, \phi) \quad \text{where} \quad a_{lm} = \int_{\Omega} f(\theta, \phi) Y_{lm}(\theta, \phi) d\mu ; \quad (746)$$

$a_{lm} \in \mathbb{R}$  and  $d\mu = \sin \theta d\theta d\phi$ . The convention adopted for the spherical harmonic functions is such that (orthonormal basis):

$$\int_{\Omega} Y_{lm}(\theta, \phi) Y_{l'm'}(\theta, \phi) d\mu = \delta_{ll'} \delta_{mm'} \quad \text{and} \quad \int_{\Omega} Y_{lm}(\theta, \phi) d\mu = \sqrt{4\pi} \delta_{l0} \quad (747)$$

In consequence, a probability density function  $p(\theta, \phi)$  with support in  $\Omega$  can be expanded as

$$p(\theta, \phi) = c_{00} Y_{00}(\theta, \phi) + \sum_{l=1}^{\infty} \sum_{m=-l}^l c_{lm} Y_{lm}(\theta, \phi) \quad (748)$$

The normalization imposes that  $c_{00} = 1/\sqrt{4\pi}$  so we can write

$$p(\theta, \phi | \mathbf{a}) = \frac{1}{4\pi} (1 + a_{lm} Y_{lm}(\theta, \phi)) \quad (749)$$

where  $l \geq 1$ ,

$$a_{lm} = 4\pi c_{lm} = 4\pi \int_{\Omega} p(\theta, \phi) Y_{lm}(\theta, \phi) d\mu = 4\pi E_{p;\mu}[Y_{lm}(\theta, \phi)] \quad (750)$$

and summation over repeated indices understood. Obviously, for any  $(\theta, \phi) \in \Omega$  we have that  $p(\theta, \phi | \mathbf{a}) \geq 0$  so the set of parameters  $\mathbf{a}$  are constrained on a compact support.

Even though we shall study the general case, we are particularly interested in the expansion up to  $l = 1$  (dipole terms) so, to simplify the notation, we redefine the indices  $(l, m) = \{(1, -1), (1, 0), (1, 1)\}$  as  $i = \{1, 2, 3\}$  and, accordingly, the coefficients  $\mathbf{a} = (a_{1-1}, a_{10}, a_{11})$  as  $\mathbf{a} = (a_1, a_2, a_3)$ . Thus:

$$p(\theta, \phi | \mathbf{a}) = \frac{1}{4\pi} (1 + a_1 Y_1 + a_2 Y_2 + a_3 Y_3) \quad (751)$$

In this case, the condition  $p(\theta, \phi | \mathbf{a}) \geq 0$  implies that the coefficients are bounded by the sphere  $a_1^2 + a_2^2 + a_3^2 \leq 4\pi/3$  and therefore, the coefficient of anisotropy

$$\delta \stackrel{\text{def.}}{=} \sqrt{\frac{3}{4\pi}} (a_1^2 + a_2^2 + a_3^2)^{1/2} \leq 1 \quad (752)$$

There are no sufficient statistics for this model but the Central Limit Theorem applies and, given the large amount of data, the experimental observations can be cast in the statistic  $\mathbf{a} = (a_1, a_2, a_3)$  such that <sup>29</sup>

$$p(\mathbf{a} | \mu) = \prod_{i=1}^3 N(a_i | \mu_i, \sigma_i^2) \quad (753)$$

with  $V(a_i) = 4\pi/n$  known and with negligible correlations ( $\rho_{ij} \simeq 0$ ).

<sup>29</sup>Essentially,  $a_{lm} = \frac{4\pi}{n} \sum_{i=1}^n Y_{lm}(\theta_i, \phi_i)$  for a sample of size  $n$ .

Consider then a  $k$ -dimensional random quantity  $\mathbf{Z} = \{Z_1, \dots, Z_k\}$  and the distribution

$$p(\mathbf{z}|\mu, \sigma) = \prod_{j=1}^k N(z_j|\mu_j, \sigma_j^2) \quad (754)$$

The interest is centered on the euclidean norm  $\|\mu\|$ , with  $\dim\{\mu\} = k$ , and its square; in particular, in

$$\delta = \sqrt{\frac{3}{4\pi}}\|\mu\| \quad \text{for } k = 3 \quad \text{and} \quad C_k = \frac{\|\mu\|^2}{k} \quad (755)$$

First, let us define  $X_j = Z_j/\sigma_j$  and  $\rho_j = \mu_j/\sigma_j$  so  $X_j \sim N(x_j|\rho_j, 1)$  and make a transformation of the parameters  $\rho_j$  to spherical coordinates:

$$\begin{aligned} \rho_1 &= \rho \cos \phi_1 \\ \rho_2 &= \rho \sin \phi_1 \cos \phi_2 \\ \rho_3 &= \rho \sin \phi_1 \sin \phi_2 \cos \phi_3 \\ &\vdots \\ \rho_{k-1} &= \rho \sin \phi_1 \sin \phi_2 \dots \sin \phi_{k-2} \cos \phi_{k-1} \\ \rho_k &= \rho \sin \phi_1 \sin \phi_2 \dots \sin \phi_{k-2} \sin \phi_{k-1} \end{aligned} \quad (756)$$

The Fisher's matrix is the Riemann metric tensor so the square root of the determinant is the  $k$ -dimensional volume element:

$$dV^k = \rho^{k-1} d\rho dS^{k-1} \quad (757)$$

with

$$dS^{k-1} = \sin^{k-2} \phi_1 \sin^{k-3} \phi_2 \dots \sin \phi_{k-2} d\phi_1 d\phi_2 \dots d\phi_{k-1} = \prod_{j=1}^{k-1} \sin^{(k-1)-j} \phi_j d\phi_j \quad (758)$$

the  $k - 1$  dimensional spherical surface element,  $\phi_{k-1} \in [0, 2\pi)$  and  $\phi_{1, \dots, k-2} \in [0, \pi]$ . The interest we have is on the parameter  $\rho$  so we should consider the ordered parameterization  $\{\rho; \phi\}$  with  $\phi = \{\phi_1, \phi_2, \dots, \phi_{k-1}\}$  nuisance parameters. Being  $\rho$  and  $\phi_i$  independent for all  $i$ , we shall consider the surface element (that is, the determinant of the submatrix obtained for the angular part) as prior density (proper) for the nuisance parameters. As we have commented in section 1, this is just the Lebesgue measure on the  $k - 1$  dimensional sphere (the Haar invariant measure under rotations) and therefore the natural choice for the prior; in other words, a uniform distribution on the  $k - 1$  dimensional sphere. Thus, we start integrating the the angular parameters. Under the assumption that the variances  $\sigma_i^2$  are all the same and considering that

$$\int_0^\pi e^{\pm\beta \cos \theta} \sin^{2\nu} \theta d\theta = \sqrt{\pi} \left(\frac{2}{\beta}\right)^\nu \Gamma\left(\nu + \frac{1}{2}\right) I_\nu(\beta) \quad \text{for } \text{Re}(\nu) > -\frac{1}{2} \quad (759)$$

one gets  $p(\phi | \text{data}) \propto p(\phi_m | \phi) \pi(\phi)$  where

$$p(\phi_m | \phi, \nu) = b e^{-b(\phi + \phi_m)} \left(\frac{\phi_m}{\phi}\right)^{\nu/2} I_\nu(2b\sqrt{\phi_m}\sqrt{\phi}) \quad (760)$$

is properly normalized,

$$\nu = k/2 - 1; \quad \phi = \|\mu\|^2; \quad \phi_m = \|\mathbf{a}\|^2; \quad b = \frac{1}{2\sigma^2} = \frac{n}{8\pi} \quad (761)$$

and  $\dim\{\mu\} = \dim\{\mathbf{a}\} = k$ . This is nothing else but a non-central  $\chi^2$  distribution.

From the series expansion of the Bessel functions it is easy to prove that this process is just a compound Poisson-Gamma process

$$p(\phi_m|\phi, \nu) = \sum_{k=0}^{\infty} Po(k|b\phi) Ga(\phi_m|b, \nu + k + 1) \tag{762}$$

and therefore the sampling distribution is a Gamma-weighted Poisson distribution with the parameter of interest that of the Poisson. From the Mellin Transform:

$$\mathcal{M}(s)_{<-\nu, \infty>} = \frac{b e^{-b\phi}}{\Gamma(\nu + 1)} \frac{\Gamma(s + \nu)}{b^s} M(s + \nu, \nu + 1, b\phi) \tag{763}$$

with  $M(a, b, z)$  the Kummer's function one can easily get the moments ( $E[\phi_m^n] = M(n + 1)$ ); in particular

$$E[\phi_m] = \phi + b^{-1}(\nu + 1) \quad \text{and} \quad V[\phi_m] = 2\phi b^{-1} + b^{-2}(\nu + 1) \tag{764}$$

Now that we have the model  $p(\phi_m|\phi)$ , let's go for the prior function  $\pi(\phi)$  or  $\pi(\delta)$ . One may guess already what shall we get. The first element of the Fisher's matrix (diagonal) corresponds to the norm and is constant so it would not be surprising to get the Lebesgue measure for the norm  $d\lambda(\delta) = \pi(\delta)d\delta = c d\delta$ . As a second argument, for large sample sizes ( $n \gg$ ) we have  $b \gg$  so  $\phi_m \sim N(\phi_m|\phi, \sigma^2 = 2\phi/b)$  and, to first order, Jeffreys' prior is  $\pi(\phi) \sim \phi^{-1/2}$ . From the reference analysis, if we take for instance

$$\pi^*(\phi) = \phi^{(\nu-1)/2} \tag{765}$$

we end up, after some algebra, with

$$\pi(\phi) \propto \pi(\phi_0) \lim_{k \rightarrow \infty} \frac{f_k(\phi)}{f_k(\phi_0)} \propto \left(\frac{\phi_0}{\phi}\right)^{1/2} \lim_{b \rightarrow \infty} e^{-3b(\phi - \phi_0)/2} + [I(\phi, b) - I(\phi_0, b)] \tag{766}$$

where

$$I(\phi, b) = \int_0^{\infty} p(\phi_m|\phi) \log \frac{I_{\nu}(2b\sqrt{\phi\phi_m})}{I_{\nu/2}(b\phi_m/2)} d\phi_m \tag{767}$$

and  $\phi_0$  any interior point of  $\Lambda(\phi) = [0, \infty)$ . From the asymptotic behavior of the Bessel functions one gets

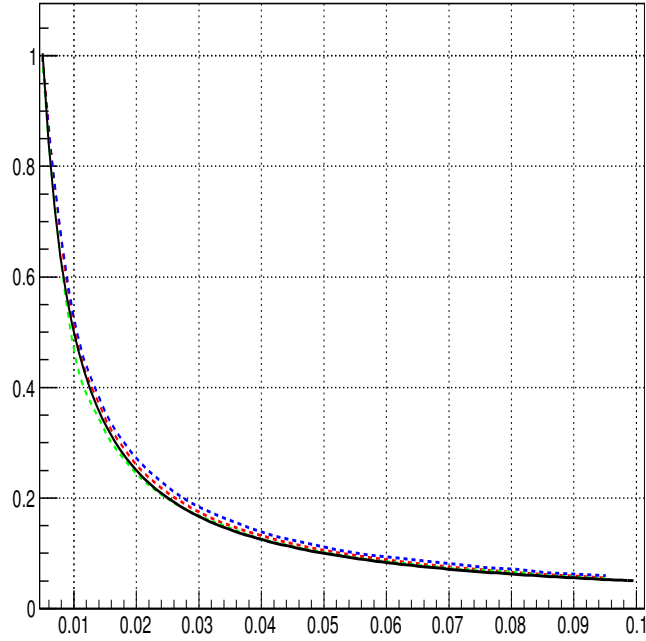
$$\pi(\phi) \propto \phi^{-1/2} \tag{768}$$

and therefore,  $\pi(\delta) = c$ . It is left as an exercise to get the same result with other priors like  $\pi^*(\phi) = c$  or  $\pi^*(\phi) = \phi^{-1/2}$ .

For this problem, it is easier to derive the prior from the reference analysis. Nevertheless, the Fisher's information that can be expressed as:

$$F(\phi; \nu) = b^2 \left\{ -1 + b \frac{e^{-b\phi}}{\phi^{\nu/2+1}} \int_0^{\infty} e^{-bz} z^{\nu/2+1} \frac{I_{\nu+1}^2(2b\sqrt{z\phi})}{I_{\nu}(2b\sqrt{z\phi})} dz \right\} \tag{769}$$

and, for large  $b$  (large sample size),  $F(\lambda; \nu) \rightarrow \phi^{-1}$  regardless the number of degrees of freedom  $\nu$ . Thus, Jeffrey's prior is consistent with the result from reference analysis. In fact, from the asymptotic behavior of the Bessel Function in the corresponding expressions of the pdf, one can already see that  $F(\phi; \nu) \sim \phi^{-1}$ . A cross check from a numeric integration is shown in figure ?? where, for  $k = 3, 5, 7$  ( $\nu = 1/2, 3/2, 5/2$ ),  $F(\phi; \nu)$  is depicted as function of  $\phi$  compared to  $1/\phi$  in black for a sufficiently



**Fig. 14:** Fisher's information (numeric integration) as function of  $\phi$  for  $k = 3, 5, 7$  (discontinuous lines) and  $f(\phi) = \phi^{-1}$  (continuous line). All are scaled so that  $F(\phi = 0.005, \nu) = 1$ .

large value of  $b$ . Therefore we shall use  $\pi(\phi) = \phi^{-1/2}$  for the cases of interest (dipole, quadrupole, ... any-pole).

The posterior densities are

- For  $\phi = \|\mu\|^2$ :  $p(\phi|\phi_m, \nu) = N e^{-b\phi} \phi^{-(\nu+1)/2} I_\nu(2b\sqrt{\phi_m}\sqrt{\phi})$  with

$$N = \frac{\Gamma(\nu + 1) b^{1/2-\nu} \phi_m^{-\nu/2}}{\sqrt{\pi} M(1/2, \nu + 1, b\phi_m)} \quad (770)$$

The Mellin Transform is

$$\mathcal{M}_\phi(s)_{<1/2, \infty>} = \frac{\Gamma(s - 1/2)}{b^{s-1} \sqrt{\pi}} \frac{M(s - 1/2, \nu + 1, b\phi_m)}{M(1/2, \nu + 1, b\phi_m)} \quad (771)$$

and therefore the moments

$$E[\phi^n] = M(n + 1) = \frac{\Gamma(n + 1/2) M(n + 1/2, \nu + 1, b\phi_m)}{\sqrt{\pi} b^n M(1/2, \nu + 1, b\phi_m)} \quad (772)$$

In the limit  $|b\phi_m| \rightarrow \infty$ ,  $E[\phi^n] = \phi_m^n$ .

- For  $\rho = \|\mu\|$ :  $p(\rho|\phi_m, \nu) = 2N e^{-b\rho^2} \rho^{-\nu} I_\nu(2b\sqrt{\phi_m}\rho)$  and

$$\mathcal{M}_\rho(s) = \mathcal{M}_\phi(s/2 + 1/2) \longrightarrow E[\rho^n] = \frac{\Gamma(n/2 + 1/2) M(n/2 + 1/2, \nu + 1, b\phi_m)}{\sqrt{\pi} b^{n/2} M(1/2, \nu + 1, b\phi_m)} \quad (773)$$



In the particular case that  $k = 3$  (dipole;  $\nu = 1/2$ ), we have for  $\delta = \sqrt{3/4\pi}\rho$  that the first two moments are:

$$E[\delta] = \frac{\text{erf}(z)}{a\delta_m M(1, 3/2, -z^2)} \quad E[\delta^2] = \frac{1}{a M(1, 3/2, -z^2)} \quad (774)$$

with  $z = 2\delta_m\sqrt{b\pi/3}$  and, when  $\delta_m \rightarrow 0$  we get

$$E[\delta] = \sqrt{\frac{2}{\pi a}} \simeq \frac{1.38}{\sqrt{n}} \quad E[\delta^2] = \frac{1}{a} \quad \sigma_\delta \simeq \frac{1.04}{\sqrt{n}} \quad (775)$$

and a one sided 95% upper credible region (see section 12 for more details) of  $\delta_{0.95} = \frac{3.38}{\sqrt{n}}$ .

So far, the analysis has been done assuming that the variances  $\sigma_j^2$  are of the same size (equal in fact) and the correlations are small. This is a very reasonable assumption but may not always be the case. The easiest way to proceed then is to perform a transformation of the parameters of interest ( $\mu$ ) to polar coordinates  $\mu(\rho, \Omega)$  and do a Monte Carlo sampling from the posterior:

$$p(\rho, \Omega | \mathbf{z}, \Sigma^{-1}) \propto \left[ \prod_{j=1}^n N(z_j | \mu_j(\rho, \Omega), \Sigma^{-1}) \right] \pi(\rho) d\rho dS^{n-1} \quad (776)$$

with a constant prior for  $\delta$  or  $\pi(\phi) \propto \phi^{-1/2}$  for  $\phi$ .

“Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin”

*J. Von Neumann*

#### 4 Monte Carlo Methods

The Monte Carlo Method is a very useful and versatile numerical technique that allows to solve a large variety of problems difficult to tackle by other procedures. Even though the central idea is to simulate experiments on a computer and make inferences from the “observed” sample, it is applicable to problems that do not have an explicit random nature; it is enough if they have an adequate probabilistic approach. In fact, a frequent use of Monte Carlo techniques is the evaluation of definite integrals that at first sight have no statistical nature but can be interpreted as expected values under some distribution.

Detractors of the method used to argue that one uses Monte Carlo Methods because a manifest incapability to solve the problems by other *more academic* means. Well, consider a “simple” process in particle physics:  $ee \rightarrow ee\mu\mu$ . Just four particles in the final state; the differential cross section in terms of eight variables that are not independent due to kinematic constraints. To see what we expect for a particular experiment, it has to be integrated within the acceptance region with dead zones between subdetectors, different materials and resolutions that distort the momenta and energy, detection efficiencies,... Yes. Admittedly we are not able to get nice expressions. Nobody in fact and Monte Carlo comes to our help. Last, it may be a truism but worth to mention that Monte Carlo is not a magic black box and will not give the answer to our problem out of nothing. It will simply present the available information in a different and more suitable manner after more or less complicated calculations are performed but all the needed information has to be put in to start with in some way or another.

In this lecture we shall present and justify essentially all the procedures that are commonly used in particle physics and statistics leaving aside subjects like Markov Chains that deserve a whole lecture by themselves and for which only the relevant properties will be stated without demonstration. A general introduction to Monte Carlo techniques can be found in [?].

## 4.1 Pseudo-Random Sequences

Sequences of random numbers  $\{x_1, x_2, \dots, x_n\}$  are the basis of Monte Carlo simulations and, in principle, their production is equivalent to perform an experiment  $e(n)$  sampling  $n$  times the random quantity  $X \sim p(x|\theta)$ . Several procedures have been developed for this purpose (real experiments, dedicated machines, digits of transcendental numbers,...) but, besides the lack of precise knowledge behind the generated sequences and the need of periodic checks, the complexity of the calculations we are interested in demands large sequences and fast generation procedures. We are then forced to devise simple and efficient arithmetical algorithms to be implemented in a computer. Obviously neither the sequences produced are random nor we can produce truly random sequences by arithmetical algorithms but we really do not need them. It is enough for them to *simulate* the relevant properties of truly random sequences and be such that if I give you one of these sequences and no additional information, you won't be able to tell after a bunch of tests [2] whether it is a truly random sequence or not (at least for the needs of the problem at hand). That's why they are called *pseudo-random* although, in what follows we shall call them *random*. The most popular (and essentially the best) algorithms are based on congruential relations (used for this purpose as far as in the 1940s) together with binary and/or shuffling operations with some free parameters that have to be fixed before the sequence is generated. They are fast, easy to implement on any computer and, with the adequate initial setting of the parameters, produce very long sequences with sufficiently good properties. And the easiest and fastest pseudo-random distribution to be generated on a computer is the *Discrete Uniform* <sup>30</sup>.

Thus, let's assume that we have a *good Discrete Uniform random number generator* <sup>31</sup> although, as Marsaglia said, "A Random Number Generator is like sex: When it is good it is wonderful; when it is bad... it is still pretty good". Each call in a computer algorithm will produce an output ( $x$ ) that we shall represent as  $x \leftarrow Un(0, 1)$  and simulates a sampling of the random quantity  $X \sim Un(x|0, 1)$ . Certainly, we are not very much interested in the Uniform Distribution so the task is to obtain a sampling of densities  $p(\mathbf{x}|\theta)$  other than Uniform from a *Pseudo-Uniform Random Number Generator* for which there are several procedures.

**Example 4.1: Estimate the value of  $\pi$ .** As a simple first example, let's see how we may estimate the value of  $\pi$ . Consider a circle of radius  $r$  inscribed in a square with sides of length  $2r$ . Imagine now that we throw random *points* evenly distributed inside the square and count how many have fallen inside the circle. It is clear that since the area of the square is  $4r^2$  and the area enclosed by the circle is  $\pi r^2$ , the probability that a throw falls inside the circle is  $\theta = \pi/4$ .

If we repeat the experiment  $N$  times, the number  $n$  of throws falling inside the circle follows a Binomial law  $Bi(n|N, p)$  and therefore, having observed  $n$  out of  $N$  trials we have that

$$p(\theta|n, N) \propto \theta^{n-1/2}(1 - \theta)^{N-n-1/2} \quad (777)$$

Let's take  $\pi^* = 4E[\theta]$  as point estimator and  $\sigma^* = 4\sigma_\theta$  as a measure of the precision. The results obtained for samplings of different size are shown in the following table:

It is interesting to see that the precision decreases with the sampling size as  $1/\sqrt{N}$ . This dependence is a general feature of Monte Carlo estimations *regardless* the number of dimensions of the problem.

A similar problem is that of Buffon's needle: *A needle of length  $l$  is thrown at random on a horizontal plane with stripes of width  $d > l$ . What is the probability that the needle intersects one of the lines between the stripes?* It is left as an exercise to shown that, as given already by Buffon in 1777,  $P_{cut} = 2l/\pi d$ . Laplace pointed out, in what may be the first use of the Monte Carlo method, that doing the experiment one may estimate the value of  $\pi$  "... *although with large error*".

<sup>30</sup>See [?] and [?] for a detailed review on random and quasi-random number generators.

<sup>31</sup>For the examples in this lecture I have used *RANMAR* [?] that can be found, for instance, at the CERN Computing Library.

**Table 1:** Estimation of  $\pi$  from a Binomial random process.

Throws ( $N$ )	Accepted ( $n$ )	$\pi^*$	$\sigma^*$
100	83	3.3069	0.1500
1000	770	3.0789	0.0532
10000	7789	3.1156	0.0166
100000	78408	3.1363	0.0052
1000000	785241	3.1410	0.0016

## 4.2 Basic Algorithms

### 4.2.1 Inverse Transform

This is, at least formally, the easiest procedure. Suppose we want a sampling of the continuous one-dimensional random quantity  $X \sim p(x)$  <sup>32</sup> so

$$P[X \in (-\infty, x]] = \int_{-\infty}^x p(x') dx' = \int_{-\infty}^x dF(x') = F(x) \quad (778)$$

Now, we define the new random quantity  $U = F(X)$  with support in  $[0, 1]$ . How is it distributed? Well,

$$F_U(u) \equiv P[U \leq u] = P[F(X) \leq u] = P[X \leq F^{-1}(u)] = \int_{-\infty}^{F^{-1}(u)} dF(x') = u \quad (779)$$

and therefore  $U \sim Un(u|0, 1)$ . The algorithm is then clear; at step  $i$ :

- $i_1) u_i \leftarrow Un(u|0, 1)$
- $i_2) x_i = F^{-1}(u_i)$

After repeating the sequence  $n$  times we end up with a sampling  $\{x_1, x_2, \dots, x_n\}$  of  $X \sim p(x)$ .

**Example 4.2:** Let's see how we generate a sampling of the Laplace distribution  $X \sim La(x|\alpha, \beta)$  with  $\alpha \in \mathcal{R}$ ,  $\beta \in (0, \infty)$  and density

$$p(x|\alpha, \beta) = \frac{1}{2\beta} e^{-|x-\alpha|/\beta} \mathbf{1}_{(-\infty, \infty)}(x) \quad (780)$$

The distribution function is

$$F(x) = \int_{-\infty}^x p(x'|\alpha, \beta) dx' = \begin{cases} \frac{1}{2} \exp\left(\frac{x-\alpha}{\beta}\right) & \text{if } x < \alpha \\ 1 - \frac{1}{2} \exp\left(-\frac{x-\alpha}{\beta}\right) & \text{if } x \geq \alpha \end{cases} \quad (781)$$

Then, if  $u \leftarrow Un(0, 1)$ :

$$x = \begin{cases} \alpha + \beta \ln(2u) & \text{if } u < 1/2 \\ \alpha - \beta \ln(2(1-u)) & \text{if } u \geq 1/2 \end{cases} \quad (782)$$

The generalization of the *Inverse Transform* method to  $n$ -dimensional random quantities is trivial. We just have to consider the marginal and conditional distributions

$$F(x_1, x_2, \dots, x_n) = F_n(x_n|x_{n-1}, \dots, x_1) \cdots F_2(x_2|x_1) \cdots F_1(x_1) \quad (783)$$

<sup>32</sup>Remember that if  $\text{supp}(X) = \Omega \subseteq \mathcal{R}$ , it is assumed that the density is  $p(x)\mathbf{1}_\Omega(x)$ .

or, for absolute continuous quantities, the probability densities

$$p(x_1, x_2, \dots, x_n) = p_n(x_n|x_{n-1}, \dots, x_1) \cdots p_2(x_2|x_1) \cdots p_1(x_1) \tag{784}$$

and then proceed sequentially; that is:

$$\begin{aligned} i_{2,1}) \quad & u_1 \leftarrow Un(u|0, 1) \text{ and } x_1 = F_1^{-1}(u_1); \\ i_{2,2}) \quad & u_2 \leftarrow Un(u|0, 1) \text{ and } x_2 = F_2^{-1}(u_2|x_1); \\ i_{2,1}) \quad & u_3 \leftarrow Un(u|0, 1) \text{ and } x_3 = F_3^{-1}(u_3|x_1, x_2); \\ & \vdots \\ i_{2,n}) \quad & u_n \leftarrow Un(u|0, 1) \text{ and } x_n = F_n^{-1}(u_n|x_{n-1}, \dots, x_1) \end{aligned}$$

If the random quantities are independent there is a unique decomposition

$$p(x_1, x_2, \dots, x_n) = \prod_{i=1}^n p_i(x_i) \quad \text{and} \quad F(x_1, x_2, \dots, x_n) = \prod_{i=1}^n F_i(x_i) \tag{785}$$

but, if this is not the case, note that there are  $n!$  ways to do the decomposition and some may be easier to handle than others (see example 4.3).

**Example 4.3:** Consider the probability density

$$p(x, y) = 2 e^{-x/y} \mathbf{1}_{(0,\infty)}(x) \mathbf{1}_{(0,1]}(y) \tag{786}$$

We can express the Distribution Function as  $F(x, y) = F(x|y)F(y)$  where:

$$p(y) = \int_0^\infty p(x, y) dx = 2y \quad \longrightarrow \quad F(y) = y^2 \tag{787}$$

$$p(x|y) = \frac{p(x, y)}{p(y)} = \frac{1}{y} e^{-x/y} \quad \longrightarrow \quad F(x|y) = 1 - e^{-x/y} \tag{788}$$

Both  $F(y)$  and  $F(x|y)$  are easy to invert so:

$$\begin{aligned} i_1) \quad & u \leftarrow Un(0, 1) \text{ and get } y = u^{1/2} \\ i_2) \quad & w \leftarrow Un(0, 1) \text{ and get } x = -y \ln w \end{aligned}$$

Repeating the algorithm  $n$  times, we get the sequence  $\{(x_1, y_1), (x_2, y_2), \dots\}$  that simulates a sampling from  $p(x, y)$ .

Obviously, we can also write  $F(x, y) = F(y|x)F(x)$  and proceed in an analogous manner. However

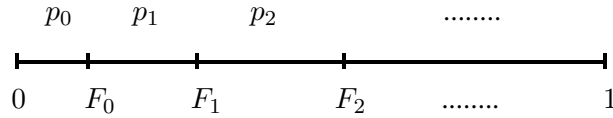
$$p(x) = \int_0^1 p(x, y) dy = 2x \int_x^\infty e^{-u} u^{-2} du \tag{789}$$

is not so easy to sample.

Last, let's see how to use the *Inverse Transform* procedure for discrete random quantities. If  $X$  can take the values in  $\Omega_X = \{x_0, x_1, x_2, \dots\}$  with probabilities  $P(X = x_k) = p_k$ , the Distribution Function will be:

$$\begin{aligned} F_0 &= P(X \leq x_0) = p_0 \\ F_1 &= P(X \leq x_1) = p_0 + p_1 \\ F_2 &= P(X \leq x_2) = p_0 + p_1 + p_2 \\ &\dots \end{aligned} \tag{790}$$

Graphically, we can represent the sequence  $\{0, F_0, F_1, F_2, \dots, 1\}$  as:



Then, it is clear that a random quantity  $u_i$  drawn from  $U(x|0, 1)$  will determine a point in the interval  $[0, 1]$  and will belong to the subinterval  $[F_{k-1}, F_k]$  with probability  $p_k = F_k - F_{k-1}$  so we can set up the following algorithm:

- $i_1)$  Get  $u_i \sim Un(u|0, 1)$ ;
- $i_2)$  Find the value  $x_k$  such that  $F_{k-1} < u_i \leq F_k$

The sequence  $\{x_0, x_1, x_2, \dots\}$  so generated will be a sampling of the probability law  $P(X = x_k) = p_k$ . Even though discrete random quantities can be sampled in this way, some times there are specific properties based on which faster algorithms can be developed. That is the case for instance for the Poisson Distribution as the following example shows.

**Example 4.4: Poisson Distribution**  $Po(k|\mu)$ . From the recurrence relation

$$p_k = e^{-\mu} \frac{\mu^k}{\Gamma(k+1)} = \frac{\mu}{k} p_{k-1} \tag{791}$$

- $i_1)$   $u_i \leftarrow Un(0, 1)$
- $i_2)$  Find the value  $k = 0, 1, \dots$  such that  $F_{k-1} < u_i \leq F_k$  and deliver  $x_k = k$

For the Poisson Distribution, there is a faster procedure. Consider a sequence of  $n$  independent random quantities  $\{X_1, X_2, \dots, X_n\}$ , each distributed as  $X_i \sim Un(x|0, 1)$ , and introduce a new random quantity

$$W_n = \prod_{k=1}^n X_k \tag{792}$$

with  $\text{supp}\{W_n\} = [0, 1]$ . Then

$$W_n \sim p(w_n|n) = \frac{(-\log w_n)^{n-1}}{\Gamma(n)} \longrightarrow P(W_n \leq a) = \frac{1}{\Gamma(n)} \int_{-\log a}^{\infty} e^{-t} t^{n-1} dt \tag{793}$$

and if we take  $a = e^{-\mu}$  we have, in terms of the Incomplete Gamma Function  $P(a, x)$ :

$$P(W_n \leq e^{-\mu}) = 1 - P(n, \mu) = e^{-\mu} \sum_{k=0}^{n-1} \frac{\mu^k}{\Gamma(k+1)} = Po(X \leq n-1|\mu) \tag{794}$$

Therefore,

- $i_0)$  Set  $w_p = 1$ ;
- $i_1)$   $u_i \leftarrow Un(0, 1)$  and set  $w_p = w_p u$ ;
- $i_2)$  Repeat step  $i_1$  while  $w_p \leq e^{-\mu}$ , say  $k$  times, and deliver  $x_k = k - 1$

**Example 4.5: Binomial Distribution**  $Bn(k|N, \theta)$ . From the recurrence relation

$$p_k = \binom{N}{k} \theta^k (1-\theta)^{n-k} = \frac{\theta}{1-\theta} \frac{n-k+1}{k} p_{k-1} \tag{795}$$

with  $p_0 = (1-\theta)^n$

- $i_1)$   $u_i \leftarrow Un(0, 1)$   
 $i_2)$  Find the value  $k = 0, 1, \dots, N$  such that  $F_{k-1} < u_i \leq F_k$  and deliver  $x_k = k$

**Example 4.6: Simulation of the response of a Photomultiplier tube.**

Photomultiplier tubes are widely used devices to detect electromagnetic radiation by means of the external photoelectric effect. A typical photomultiplier consists of a vacuum tube with an input window, a photocathode, a focusing and a series of amplifying electrodes (dynodes) and an electron collector (anode). Several materials are used for the input window (borosilicate glass, synthetic silica,...) which transmit radiation in different wavelength ranges and, due to absorptions (in particular in the UV range) and external reflexions, the transmittance of the window is never 100%. Most photocathodes are compound semiconductors consisting of alkali metals with a low work function. When the photons strike the photocathode the electrons in the valence band are excited and, if they get enough energy to overcome the vacuum level barrier, they are emitted into the vacuum tube as photoelectrons. The trajectory of the electrons inside the photomultiplier is determined basically by the applied voltage and the geometry of the focusing electrode and the first dynode. Usually, the photoelectron is driven towards the first dynode and originates an electron shower which is amplified in the following dynodes and collected at the anode. However, a fraction of the incoming photons pass through the photocathode and originates a smaller electron shower when it strikes the first dynode of the amplification chain.

To study the response of a photomultiplier tube, an experimental set-up has been made with a LED as photon source. We are interested in the response for isolated photons so we regulate the current and the frequency so as to have a low intensity source. Under this conditions, the number of photons that arrive at the window of the photomultiplier is well described by a Poisson law. When one of this photons strikes on the photocathode, an electron is ejected and driven towards the first dynode to start the electron shower. We shall assume that the number of electrons so produced follows also a Poisson law  $n_{gen} \sim Po(n|\mu)$ . The parameter  $\mu$  accounts for the efficiency of this first process and depends on the characteristics of the photocathode, the applied voltage and the geometry of the focusing electrodes (essentially that of the first dynode). It has been estimated to be  $\mu = 0.25$ . Thus, we start our simulation with

$$1) \ n_{gen} \leftarrow Po(n|\mu)$$

electrons leaving the photocathode. They are driven towards the the first dynode to start the electron shower but there is a chance that they miss the first and start the shower at the second. Again, the analysis of the experimental data suggest that this happens with probability  $p_{d2} \simeq 0.2$ . Thus, we have to decide how many of the  $n_{gen}$  electrons start the shower at the second dynode. A Binomial model is appropriate in this case:

$$2) \ n_{d2} \leftarrow Bi(n_{d2}|n_{gen}, p_{d2}) \text{ and therefore } n_{d1} = n_{gen} - n_{d2}.$$

Obviously, we shall do this second step if  $n_{gen} > 0$ .

Now we come to the amplification stage. Our photomultiplier has 12 dynodes so let's see the response of each of them. For each electron that strikes upon dynode  $k$  ( $k = 1, \dots, 12$ ),  $n_k$  electrons will be produced and directed towards the next element of the chain (dynode  $k + 1$ ), the number of them again well described by a Poisson law  $Po(n_k|\mu_k)$ . If we denote by  $V$  the total voltage applied between the photocathode and the anode and by  $R_k$  the resistance previous to dynode  $k$  we have that the current intensity through the chain will be

$$I = \frac{V}{\sum_{i=1}^{13} R_i}$$

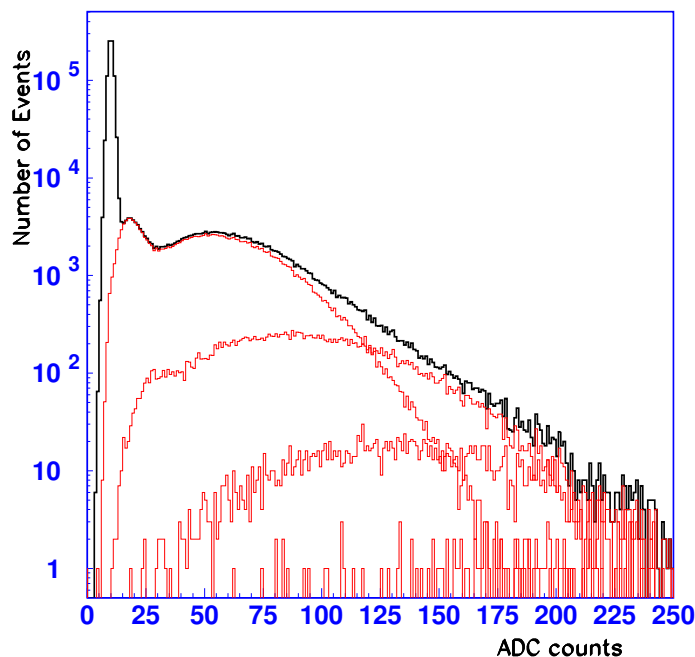
where we have considered also the additional resistance between the last dynode and the anode that collects the electron shower. Therefore, the parameters  $\mu_k$  are determined by the relation

$$\mu_k = a (I R_k)^b$$

where  $a$  and  $b$  are characteristic parameters of the photomultiplier. In our case we have that  $N = 12$ ,  $a = 0.16459$ ,  $b = 0.75$ , a total applied voltage of 800 V and a resistance chain of  $\{2.4, 2.4, 2.4, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.2, 2.4\}$  Ohms. It is easy to see that if the response of dynode  $k$  to one electron is modeled as  $Po(n_k|\mu_k)$ , the response to  $n_i$  incoming electrons is described by  $Po(n_k|n_i\mu_k)$ . Thus, we simulate the amplification stage as:

- 3.1) If  $n_{d1} > 0$  :            do from  $k=1$  to 12:

$$\mu = \mu_k n_{d1} \longrightarrow n_{d1} \leftarrow Po(n|\mu)$$



**Fig. 15:** Result of the simulation of the response of a photomultiplier tube. The histogram contains  $10^6$  events and shows the final ADC distribution detailing the contribution of the pedestal and the response to 1, 2 and 3 incoming photons.

3.2) If  $n_{d2} > 0$  :      do from  $k=2$  to 12:

$$\mu = \mu_k n_{d2} \longrightarrow n_{d2} \leftarrow Po(n|\mu)$$

Once this is done, we have to convert the number of electrons at the anode in ADC counts. The electron charge is  $Q_e = 1.602176 \cdot 10^{-19}$  C and in our set-up we have  $f_{ADC} = 2.1 \cdot 10^{14}$  ADC counts per Coulomb so

$$ADC_{pm} = (n_{d1} + n_{d2}) (Q_e f_{ADC})$$

Last, we have to consider the noise ( *pedestal* ). In our case, the number of pedestal ADC counts is well described by a mixture model with two Normal densities

$$p_{ped}(x|\cdot) = \alpha N_1(x|10., 1.) + (1 - \alpha) N_1(x|10., 1.5) \quad (796)$$

with  $\alpha = 0.8$ . Thus, with probability  $\alpha$  we obtain  $ADC_{ped} \leftarrow N_1(x|10, 1.5)$ , and with probability  $1 - \alpha$ ,  $ADC_{ped} \leftarrow N_1(x|10, 1)$  so the total number of ADC counts will be

$$ADC_{tot} = ADC_{ped} + ADC_{pm}$$

Obviously, if in step 1) we get  $n_{gen} = 0$ , then  $ADC_{tot} = ADC_{ped}$ . Figure ?? shows the result of the simulation for a sampling size of  $10^6$  together with the main contributions (1,2 or 3 initial photoelectrons) and the pedestal. From these results, the parameters of the device can be adjusted (voltage, resistance chain,...) to optimize the response for our specific requirements.

The *Inverse Transform* method is conceptually simple and easy to implement for discrete distributions and many continuous distributions of interest. Furthermore, is *efficient* in the sense that for each generated value  $u_i$  as  $Un(x|0, 1)$  we get a value  $x_i$  from  $F(x)$ . However, with the exception of easy

distributions the inverse function  $F^{-1}(x)$  has no simple expression in terms of elementary functions and may be difficult or time consuming to invert. This is, for instance, the case if you attempt to invert the Error Function for the Normal Distribution. Thus, apart from simple cases, the *Inverse Transform* method is used in combination with other procedures to be described next.

---

**NOTE 6: Bootstrap.** Given the iid sample  $\{x_1, x_2, \dots, x_n\}$  of the random quantity  $X \sim p(x|\theta)$  we know (Glivenko-Cantelli theorem; see lecture 1 (7.6)) that:

$$F_n(x) = 1/n \sum_{k=1}^n \mathbf{1}_{(-\infty, x]}(x_k) \xrightarrow{\text{unif.}} F(x|\theta) \quad (797)$$

Essentially the idea behind the *bootstrap* is to sample from the empirical Distribution Function  $F_n(x)$  that, as we have seen for discrete random quantities, is equivalent to draw samplings  $\{x'_1, x'_2, \dots, x'_n\}$  of size  $n$  from the original sample with replacement. Obviously, increasing the number of resamplings does not provide more information than what is contained in the original data but, used with good sense, each *bootstrap* will lead to a posterior and can also be useful to give insight about the form of the underlying model  $p(x|\theta)$  and the distribution of some statistics. We refer to [?] for further details.

---

#### 4.2.2 Acceptance-Rejection (Hit-Miss; J. Von Neumann, 1951)

The *Acceptance-Rejection* algorithm is easy to implement and allows to sample a large variety of  $n$ -dimensional probability densities with a less detailed knowledge of the function. But nothing is for free; these advantages are in detriment of the generation efficiency.

Let's start with the one dimensional case where  $X \sim p(x|\theta)$  is a continuous random quantity with  $\text{supp}(X) = [a, b]$  and  $p_m = \max_x p(x|\theta)$ . Consider now two independent random quantities  $X_1 \sim Un(x_1|\alpha, \beta)$  and  $X_2 \sim Un(x_2|0, \delta)$  where  $[a, b] \subseteq \text{supp}(X_1) = [\alpha, \beta]$  and  $[0, p_m] \subseteq \text{supp}(X_2) = [0, \delta]$ . The covering does not necessarily have to be a rectangle in  $\mathcal{R}^2$  (nor a hypercube  $\mathcal{R}^{n+1}$ ) and, in fact, in some cases it may be interesting to consider other coverings to improve the efficiency but the generalization is obvious. Then

$$p(x_1, x_2|\cdot) = \frac{1}{\beta - \alpha} \frac{1}{\delta} \quad (798)$$

Now, let's find the distribution of  $X_1$  conditioned to  $X_2 \leq p(x_1|\theta)$ :

$$\begin{aligned} P(X_1 \leq x | X_2 \leq p(x|\theta)) &= \frac{P(X_1 \leq x, X_2 \leq p(x|\theta))}{P(X_2 \leq p(x|\theta))} = \frac{\int_{\alpha}^x dx_1 \int_0^{p(x_1|\theta)} p(x_1, x_2|\cdot) dx_2}{\int_{\alpha}^{\beta} dx_1 \int_0^{p(x_1|\theta)} p(x_1, x_2|\cdot) dx_2} = \\ &= \frac{\int_{\alpha}^x p(x_1|\theta) \mathbf{1}_{[a, b]}(x_1) dx_1}{\int_{\alpha}^{\beta} p(x_1|\theta) \mathbf{1}_{[a, b]}(x_1) dx_1} = \int_a^x p(x_1|\theta) dx_1 = F(x|\theta) \end{aligned} \quad (799)$$

so we set up the following algorithm:

- $i_1)$   $u_i \leftarrow Un(u|\alpha \leq a, \beta \geq b)$  and  $w_i \leftarrow Un(w|0, \delta \geq p_m)$ ;
- $i_2)$  If  $w_i \leq p(u_i|\theta)$  we *accept*  $x_i$ ; otherwise we *reject*  $x_i$  and start again from  $i_1$

Repeating the algorithm  $n$  times we get the sampling  $\{x_1, x_2, \dots, x_n\}$  from  $p(x|\theta)$ .



Besides its simplicity, the Acceptance-Rejection scheme does not require to have normalized densities for it is enough to know an upper bound and in some cases, for instance when the support of the random quantity  $X$  is determined by functional relations, it is easier to deal with a simpler covering of the support. However, the price to pay is a low *generation efficiency*:

$$\epsilon \stackrel{def.}{=} \frac{\text{accepted trials}}{\text{total trials}} = \frac{\text{area under } p(x|\theta)}{\text{area of the covering}} \leq 1 \tag{800}$$

Note that the efficiency so defined refers only to the fraction of accepted trials and, obviously, the more adjusted the covering is the better but for the *Inverse Transform*  $\epsilon = 1$  and it does not necessarily imply that it is more efficient attending to other considerations. It is interesting to observe that if we do not know the normalization factor of the density function, it can be estimated as

$$\int_X p(x|\theta) dx \simeq (\text{area of the covering}) \epsilon \tag{801}$$

Let's see some examples before we proceed.

**Example 4.7:** Consider  $X \sim Be(x|\alpha, \beta)$ . In this case, what follows is just for pedagogical purposes since other procedures to be discussed later are more efficient. Anyway, the density is

$$p(x|\alpha, \beta) \propto x^{\alpha-1} (1-x)^{\beta-1} \quad ; \quad x \in [0, 1] \tag{802}$$

Suppose that  $\alpha, \beta > 1$  so the mode  $x_o = (\alpha - 1)(\alpha + \beta - 2)^{-1}$  exists and is unique. Then

$$p_m \equiv \max_x \{p(x|\alpha, \beta)\} = p(x_o|\alpha, \beta) = \frac{(\alpha - 1)^{\alpha-1} (\beta - 1)^{\beta-1}}{(\alpha + \beta - 2)^{\alpha+\beta-2}} \tag{803}$$

so let's take then the domain  $[\alpha = 0, \beta = 1] \times [0, p_m]$  and

- $i_1$ ) Get  $x_i \sim Un(x|0, 1)$  and  $y_i \sim Un(y|0, p_m)$ ;
- $i_2$ ) If
  - a)  $y_i \leq p(x_i|\alpha, \beta)$  we deliver ( *accept* )  $x_i$
  - r)  $y_i > p(x_i|\alpha, \beta)$  we *reject*  $x_i$  and start again from  $i_1$

Repeating the procedure  $n$  times, we get a sampling  $\{x_1, x_2, \dots, x_n\}$  from  $Be(x|\alpha, \beta)$ . In this case we know the normalization so the area under  $p(x|\alpha, \beta)$  is  $Be(x|\alpha, \beta)$  and the *generation efficiency* will be:

$$\epsilon = B(\alpha, \beta) \frac{(\alpha + \beta - 2)^{\alpha+\beta-2}}{(\alpha - 1)^{\alpha-1} (\beta - 1)^{\beta-1}} \tag{804}$$

**Example 4.8:** Let's generate a sampling of the spatial distributions of a bounded electron in a Hydrogen atom. In particular, as an example, those for the principal quantum number  $n = 3$ . The wave-function is  $\psi_{nlm}(r, \theta, \phi) = R_{nl}(r)Y_{lm}(\theta, \phi)$  with:

$$R_{30} \propto (1 - 2r/2 - 2r^2/27)e^{-r/3} \quad ; \quad R_{31} \propto r(1 - r/6)e^{-r/3} \quad \text{and} \quad R_{32} \propto r^2 e^{-r/3} \tag{805}$$

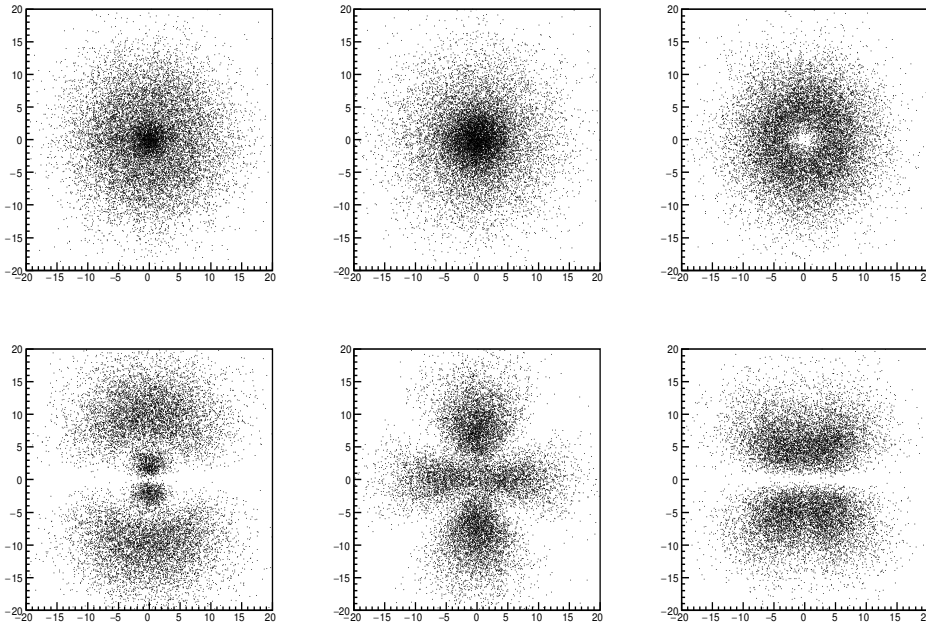
the radial functions of the 3s, 3p and 3d levels and

$$|Y_{10}|^2 \propto \cos^2\theta \quad ; \quad |Y_{1\pm 1}|^2 \propto \sin^2\theta \quad ; \quad |Y_{20}|^2 \propto (3 \cos^2\theta - 1)^2 \quad ; \quad |Y_{2\pm 1}|^2 \propto \cos^2\theta \sin^2\theta \quad ; \quad |Y_{2\pm 2}|^2 \propto \sin^4\theta \tag{806}$$

the angular dependence from the spherical harmonics. Since  $d\mu = r^2 \sin\theta dr d\theta d\phi$ , the probability density will be

$$p(r, \theta, \phi|n, l, m) = R_{nl}^2(r) |Y_{lm}|^2 r^2 \sin\theta = p_r(r|n, l) p_\theta(\theta|l, m) p_\phi(\phi) \tag{807}$$

so we can sample independently  $r, \theta$  and  $\phi$ . It is left as an exercise to explicit a sampling procedure. Note however that, for the marginal radial density, the mode is at  $r = 13, 12$  and  $9$  for  $l = 0, 1$  and  $2$  and decreases exponentially so even if the support is  $r \in [0, \infty)$  it will be a reasonable approximation to take a covering  $r \in [0, r_{max})$  such that



**Fig. 16:** Spatial probability distributions of an electron in a hydrogen atom corresponding to the quantum states  $(n, l, m) = (3, 1, 0)$ ,  $(3, 2, 0)$  and  $(3, 2, \pm 1)$  (columns 1,2,3) and projections  $(x, y)$  and  $(x, z) = (y, z)$  (rows 1 and 2) (see example 4.8).

$P(r \geq r_{max})$  is small enough. After  $n = 4000$  samplings for the quantum numbers  $(n, l, m) = (3, 1, 0)$ ,  $(3, 2, 0)$  and  $(3, 2, \pm 1)$ , the projections on the planes  $\pi_{xy}$ ,  $\pi_{xz}$  and  $\pi_{yz}$  are shown in figure ??.

The generalization of the *Acceptance-Rejection* method to sample a  $n$ -dimensional density  $\mathbf{X} \sim p(\mathbf{x}|\theta)$  with  $\dim(\mathbf{x}) = n$  is straight forward. Covering with an  $n+1$  dimensional hypercube:

$i_1)$  Get a sampling  $\{x_i^1, x_i^2, \dots, x_i^n; y_i\}$  where

$$\{x_i^k\} \leftarrow Un(x_i^k | \alpha_k, \beta_k)_{k=1}^n; \quad y_i \leftarrow Un(y|0, k) \quad \text{and} \quad k \geq \max_{\mathbf{x}} p(\mathbf{x}|\theta)$$

$i_2)$  Accept the  $n$ -tuple  $\mathbf{x}_i = (x_i^1, x_i^2, \dots, x_i^n)$  if  $y_i \leq p(\mathbf{x}_i|\theta)$  or reject it otherwise.

#### 4.2.2.1 Incorrect estimation of $\max_{\mathbf{x}}\{p(\mathbf{x}|\cdot)\}$

Usually, we know the support  $[\alpha, \beta]$  of the random quantity but the pdf is complicated enough to know the maximum. Then, we start the generation with our best guess for  $\max_{\mathbf{x}} p(\mathbf{x}|\cdot)$ , say  $k_1$ , and after having generated  $N_1$  events (*generated*, not *accepted*) in  $[\alpha, \beta] \times [0, k_1]$ ,... wham!, we generate a value  $x_m$  such that  $p(x_m) > k_1$ . Certainly, our estimation of the maximum was not correct. A possible solution is to forget about what has been generated and start again with the new maximum  $k_2 = p(x_m) > k_1$  but, obviously, this is not desirable among other things because we have no guarantee that this is not going to happen again. We better keep what has been done and proceed in the following manner:

- 1) We have generated  $N_1$  pairs  $(x_1, x_2)$  in  $[\alpha, \beta] \times [0, k_1]$  and, in particular,  $X_2$  uniformly in  $[0, k_1]$ . How many additional pairs  $N_a$  do we have to generate? Since the density of pairs is constant in both domains  $[\alpha, \beta] \times [0, k_1]$  and  $[\alpha, \beta] \times [0, k_2]$  we have that

$$\frac{N_1}{(\beta - \alpha) k_1} = \frac{N_1 + N_a}{(\beta - \alpha) k_2} \quad \longrightarrow \quad N_a = N_1 \left( \frac{k_2}{k_1} - 1 \right) \quad (808)$$

2) How do we generate them? Obviously in the domain  $[\alpha, \beta] \times [k_1, k_2]$  but from the truncated density

$$p_e(x|\cdot) = (p(x|\cdot) - k_1) \mathbf{1}_{p(x|\cdot) > k_1}(x) \quad (809)$$

3) Once the  $N_a$  additional events have been generated (out of which some have been hopefully accepted) we continue with the usual procedure but on the domain  $[\alpha, \beta] \times [0, k_2]$ .

The whole process is repeated as many times as needed.

**NOTE 7: Weighted events.** The *Acceptance-Rejection* algorithm just explained is equivalent to:

- $i_1$ ) Sample  $x_i$  from  $Un(x|\alpha, \beta)$  and  $u_i$  from  $Un(u|0, 1)$ ;
- $i_2$ ) Assign to each generated event  $x_i$  a *weight*:  $w_i = p(x_i|\cdot)/p_m$ ;  $0 \leq w_i \leq 1$  and accept the event if  $u_i \leq w_i$  or reject it otherwise.

It is clear that:

- Events with a higher weight will have a higher chance to be accepted;
- After applying the *acceptance-rejection* criteria at step  $i_2$ , all events will have a weight either 1 if it has been accepted or 0 if it was rejected.
- The generation efficiency will be

$$\epsilon = \frac{\text{accepted trials}}{\text{total trials}(N)} = \frac{1}{N} \sum_{i=1}^N w_i = \bar{w} \quad (810)$$

In some cases it is interesting to keep all the events, accepted or not.

**Example 4.9:** Let's obtain a sampling  $\{\mathbf{x}_1, \mathbf{x}_2, \dots\}$ ,  $\dim(\mathbf{x}) = n$ , of points inside a n-dimensional sphere centered at  $\mathbf{x}^c$  and radius  $r$ . For a direct use of the *Acceptance-Rejection* algorithm we enclose the sphere in a n-dimensional hypercube

$$C_n = \prod_{i=1}^n [x_i^c - r, x_i^c + r] \quad (811)$$

and:

- 1)  $x_i \leftarrow Un(x|x_i^c - r, x_i^c + r)$  for  $i = 1, \dots, n$
- 2) Accept  $\mathbf{x}_i$  if  $\rho_i = \|\mathbf{x}_i - \mathbf{x}^c\| \leq r$  and reject otherwise.

The generation efficiency will be

$$\epsilon(n) = \frac{\text{volume of the sphere}}{\text{volume of the covering}} = \frac{2 \pi^{n/2}}{n \Gamma(n/2)} \frac{1}{2^n} \quad (812)$$

Note that the sequence  $\{\mathbf{x}_i/\rho_i\}_{i=1}^n$  will be a sampling of points uniformly distributed on the sphere of radius  $r = 1$ . This we can get also as:

- 1)  $z_i \leftarrow N(z|0, 1)$  for  $i = 1, \dots, n$
- 2)  $\rho = \|\mathbf{z}_i\|$  and  $\mathbf{x}_i = \mathbf{z}_i/\rho$

Except for simple densities, the efficiency of the *Acceptance-Rejection* algorithm is not very high and decreases quickly with the number of dimensions. For instance, we have seen in the previous example that covering the n-dimensional sphere with a hypercube has a generation efficiency

$$\epsilon(n) = \frac{2 \pi^{n/2}}{n \Gamma(n/2)} \frac{1}{2^n} \quad (813)$$

and  $\lim_{n \rightarrow \infty} \epsilon(n) = 0$ . Certainly, some times we can refine the covering since there is no need other than simplicity for a hypercube (see *Stratified Sampling*) but, in general, the origin of the problem will remain: when we generate points uniformly in whatever domain, we are sampling with constant density regions that have a very low probability content or even zero when they have null intersection with the support of the random quantity  $\mathbf{X}$ . This happens, for instance, when we want to sample from a differential cross-section that has very sharp peaks (sometimes of several orders of magnitude as in the case of bremsstrahlung). Then, the problem of having a low efficiency is not just the time expend in the generation but the accuracy and convergence of the evaluations. We need a more clever way to generate sequences and the *Importance Sampling* method comes to our help.

### 4.2.3 Importance Sampling

The *Importance Sampling* generalizes the *Acceptance-Rejection* method sampling the density function with higher frequency in regions of the domain where the probability of acceptance is larger (more *important*). Let's see the one-dimensional case since the extension to n-dimensions is straight forward.

Suppose that we want a sampling of  $X \sim p(x)$  with support  $\Omega_X \in [a, b]$  and  $F(x)$  the corresponding distribution function. We can always express  $p(x)$  as:

$$p(x) = c g(x) h(x) \quad (814)$$

where:

- 1)  $h(x)$  is a probability density function, i.e., non-negative and normalized in  $\Omega_X$ ;
- 2)  $g(x) \geq 0$  ;  $\forall x \in \Omega_X$  and has a finite maximum  $g_m = \max\{g(x); x \in \Omega_X\}$ ;
- 3)  $c > 0$  a constant normalization factor.

Now, consider a sampling  $\{x_1, x_2, \dots, x_n\}$  drawn from the density  $h(x)$ . If we apply the *Acceptance-Rejection* criteria with  $g(x)$ , how are the accepted values distributed? It is clear that, if  $g_m = \max(g(x))$  and  $Y \sim Un(y|0, g_m)$

$$P(X \leq x | Y \leq g(x)) = \frac{\int_a^x h(x) dx \int_0^{g(x)} dy}{\int_a^b h(x) dx \int_0^{g(x)} dy} = \frac{\int_a^x h(x) g(x) dx}{\int_a^b h(x) g(x) dx} = F(x) \quad (815)$$

and therefore, from a sampling of  $h(x)$  we get a sampling of  $p(x)$  applying the *Acceptance-Rejection* with the function  $g(x)$ . There are infinite options for  $h(x)$ . First, the simpler the better for then the Distribution Function can be easily inverted and the *Inverse Transform* applied efficiently. The Uniform Density  $h(x) = Un(x|a, b)$  is the simplest one but then  $g(x) = p(x)$  and this is just the *Acceptance-Rejection* over  $p(x)$ . The second consideration is that  $h(x)$  be a fairly good approximation to  $p(x)$  so that  $g(x) = p(x)/h(x)$  is as smooth as possible and the *Acceptance-Rejection* efficient. Thus, if  $h(x) > 0$   $\forall x \in [a, b]$ :

$$p(x) dx = \frac{p(x)}{h(x)} h(x) dx = g(x) dH(x) \quad (816)$$

#### 4.2.3.1 Stratified Sampling

The *Stratified Sampling* is a particular case of the *Importance Sampling* where the density  $p(x); x \in \Omega_X$  is approximated by a simple function over  $\Omega_X$ . Thus, in the one-dimensional case, if  $\Omega = [a, b]$  and we take the partition ( *stratification* )

$$\Omega = \cup_{i=1}^n \Omega_i = \cup_{i=1}^n [a_{i-1}, a_i] \quad ; \quad a_0 = a, \quad a_n = b \quad (817)$$

with measure  $\lambda(\Omega_i) = (a_i - a_{i-1})$ , we have

$$h(x) = \sum_{i=1}^n \frac{\mathbf{1}_{[a_{i-1}, a_i)}(x)}{\lambda(\Omega_i)} \quad \longrightarrow \quad \int_{a_0}^{a_n} h(x) dx = 1 \quad (818)$$

Denoting by  $p_m(i) = \max_x \{p(x) | x \in \Omega_i\}$ , we have that for the *Acceptance-Rejection* algorithm the volume of each sampling domain is  $V_i = \lambda(\Omega_i) p_m(i)$ . In consequence, for a partition of size  $n$ , if  $Z \in \{1, 2, \dots, n\}$  and define

$$P(Z = k) = \frac{V_k}{\sum_{i=1}^n V_i} \quad ; \quad F(k) = P(Z \leq k) = \sum_{j=1}^k P(Z = j) \quad (819)$$

we get a sampling of  $p(x)$  from the following algorithm:

- $i_1)$   $u_i \leftarrow Un(u|0, 1)$  and select the partition  $k = \text{Int}[\min\{F_i | F_i > n \cdot u_i\}]$ ;
- $i_2)$   $x_i \leftarrow Un(x|a_{k-1}, a_k)$ ,  $y_i \leftarrow Un(y|0, p_m(k))$  and accept  $x_i$  if  $y_i \leq p(x_i)$  (reject otherwise).

#### 4.2.4 Decomposition of the probability density

Some times it is possible to express in a simple manner the density function as a linear combination of densities; that is

$$p(x) = \sum_{j=1}^k a_j p_j(x) \quad ; \quad a_j > 0 \quad \forall j = 1, 2, \dots, k \quad (820)$$

that are easier to sample. Since normalization imposes that

$$\int_{-\infty}^{\infty} p(x) dx = \sum_{j=1}^k a_j \int_{-\infty}^{\infty} p_j(x) dx = \sum_{j=1}^k a_m = 1 \quad (821)$$

we can sample from  $p(x)$  selecting, at each step  $i$ , one of the  $k$  densities  $p_i(x)$  with probability  $p_i = a_i$  from which we shall obtain  $x_i$  and therefore sampling with higher frequency from those densities that have a higher relative weight. Thus:

- $i_1)$  Select which density  $p_i(x)$  are we going to sample at step  $i_2$  with probability  $p_i = a_i$ ;
- $i_2)$  Get  $x_i$  from  $p_i(x)$  selected at  $i_1$ .

It may happen that some densities  $p_j(x)$  can not be easily integrated so we do not know a priori the relative weights. If this is the case, we can sample from  $f_j(x) \propto p_j(x)$  and estimate with the generated events from  $f_i(x)$  the corresponding normalizations  $I_i$  with, for instance, from the sample mean

$$I_i = \frac{1}{n} \sum_{k=1}^n f_i(x_k) \quad (822)$$

Then, since  $p_i(x) = f_i(x)/I_i$  we have that

$$p(x|\cdot) = \sum_{i=1}^K a_i f_i(x|\cdot) = \sum_{i=1}^K a_i I_i \frac{f_i(x)}{I_i} = \sum_{i=1}^K a_i I_i p_i(x) \quad (823)$$

so each generated event from  $f_i(x)$  has a weight  $w_i = a_i I_i$ .

**Example 4.10:** Suppose we want to sample from the density

$$p(x) = \frac{3}{8}(1 + x^2) \quad ; \quad x \in [-1, 1] \quad (824)$$

Then, we can take:

$$\left. \begin{array}{l} p_1(x) \propto 1 \\ p_2(x) \propto x^2 \end{array} \right\} \longrightarrow \text{normalization} \longrightarrow \begin{cases} p_1(x) = 1/2 \\ p_2(x) = 3x^2/2 \end{cases} \quad (825)$$

so:

$$p(x) = \frac{3}{4}p_1(x) + \frac{1}{4}p_2(x) \quad (826)$$

Then:

$i_1)$  Get  $u_i$  and  $w_i$  as  $Un(u|0, 1)$ ;

$i_2)$  Get  $x_i$  as:

if  $u_i \leq 3/4$  then  $x_i = 2w_i - 1$

if  $u_i > 3/4$  then  $x_i = (2w_i - 1)^{1/3}$

In this case, 75% of the times we sample from the trivial density  $Un(x|-1, 1)$ .

### 4.3 Everything at work

#### 4.3.1 The Compton Scattering

When problems start to get complicated, we have to combine several of the aforementioned methods; in this case *Importance Sampling*, *Acceptance-Rejection* and *Decomposition* of the probability density.

Compton Scattering is one of the main processes that occur in the interaction of photons with matter. When a photon interacts with one of the atomic electrons with an energy greater than the binding energy of the electron, it suffers an inelastic scattering resulting in a photon of less energy and different direction than the incoming one and an ejected free electron from the atom. If we make the simplifying assumptions that the atomic electron initially at rest and neglect the binding energy we have that if the incoming photon has an energy  $E_\gamma$  its energy after the interaction ( $E'_\gamma$ ) is:

$$\epsilon = \frac{E'_\gamma}{E_\gamma} = \frac{1}{1 + a(1 - \cos\theta)} \quad (827)$$

where  $\theta \in [0, \pi]$  is the angle between the momentum of the outgoing photon and the incoming one and  $a = E_\gamma/m_e$ . It is clear that if the dispersed photon goes in the forward direction (that is, with  $\theta = 0$ ), it will have the maximum possible energy ( $\epsilon = 1$ ) and when it goes backwards (that is,  $\theta = \pi$ ) the smallest possible energy ( $\epsilon = (1 + 2a)^{-1}$ ). Being a two body final state, given the energy (or the angle) of the outgoing photon the rest of the kinematic quantities are determined uniquely:

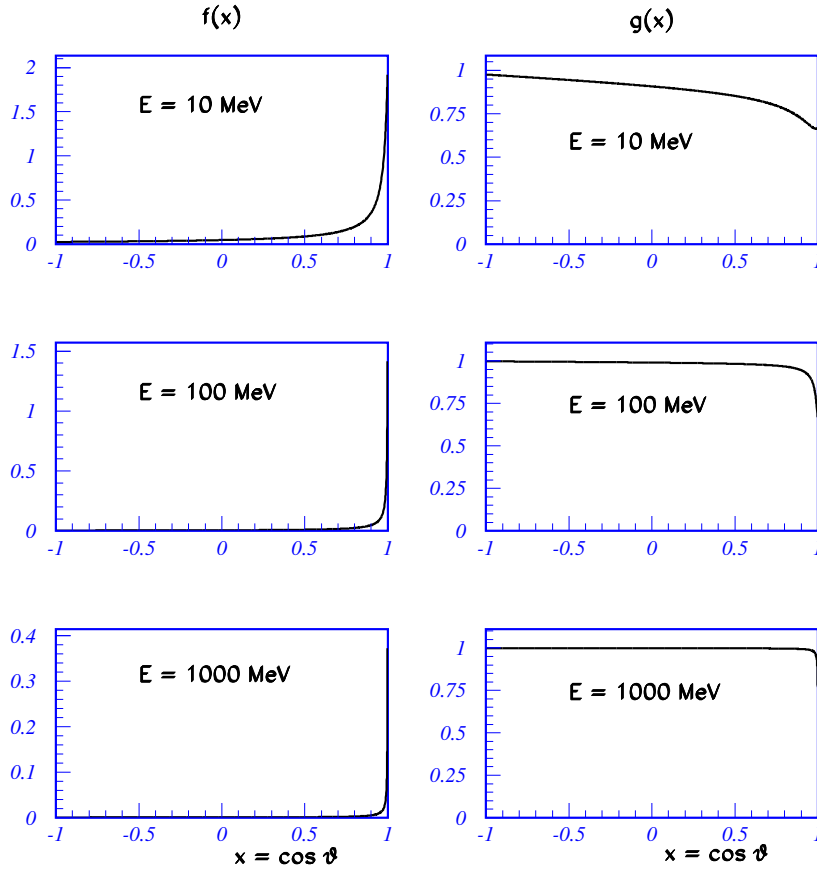
$$E'_e = E_\gamma \left(1 + \frac{1}{a} - \epsilon\right) \quad \text{and} \quad \tan\theta_e = \frac{\cot\theta/2}{1 + a} \quad (828)$$

The cross-section for the Compton Scattering can be calculated perturbatively in Relativistic Quantum Mechanics resulting in the Klein-Nishina expression:

$$\frac{d\sigma_0}{dx} = \frac{3\sigma_T}{8} f(x) \quad (829)$$

where  $x = \cos(\theta)$ ,  $\sigma_T = 0.665 \text{ barn} = 0.665 \cdot 10^{-24} \text{ cm}^2$  is the *Thomson cross-section* and

$$f(x) = \frac{1}{[1 + a(1 - x)]^2} \left(1 + x^2 + \frac{a^2(1 - x)^2}{1 + a(1 - x)}\right) \quad (830)$$



**Fig. 17:** Functions  $f(x)$  (left) and  $g(x)$  (right) for different values of the incoming photon energy.

has all the angular dependence. Due to the azimuthal symmetry, there is no explicit dependence with  $\phi \in [0, 2\pi]$  and has been integrated out. Last, integrating this expression for  $x \in [-1, 1]$  we have the total cross-section of the process:

$$\sigma_0(E_\gamma) = \frac{\sigma_T}{4} \left[ \left( \frac{1+a}{a^2} \right) \left( \frac{2(1+a)}{1+2a} - \frac{\ln(1+2a)}{a} \right) + \frac{\ln(1+2a)}{2a} - \frac{1+3a}{(1+2a)^2} \right] \quad (831)$$

For a material with  $Z$  electrons, the atomic cross-section can be approximated by  $\sigma = Z \sigma_0 \text{ cm}^2/\text{atom}$ .

Let's see how to simulate this process sampling the angular distribution  $p(x) \sim f(x)$ . Figure ?? (left) shows this function for incoming photon energies of 10, 100 and 1000 MeV. It is clear that it is peaked at  $x$  values close to 1 and gets sharper with the incoming energy; that is, when the angle between the incoming and outgoing photon momentum becomes smaller. In consequence, for high energy photons the *Acceptance-Rejection* algorithm becomes very inefficient. Let's then define the functions

$$f_n(x) = \frac{1}{[1 + a(1-x)]^n} \quad (832)$$

and express  $f(x)$  as  $f(x) = (f_1(x) + f_2(x) + f_3(x)) \cdot g(x)$  where

$$g(x) = 1 - (2-x^2) \frac{f_1(x)}{1 + f_1(x) + f_2(x)} \quad (833)$$

The functions  $f_n(x)$  are easy enough to use the *Inverse Transform* method and apply afterward the *Acceptance-Rejection* on  $g(x) > 0 \forall x \in [-1, 1]$ . The shape of this function is shown in figure ?? (right) for different values of the incoming photon energy and clearly is much more smooth than  $f(x)$  so the *Acceptance-Rejection* will be significantly more efficient. Normalizing properly the densities

$$p_i(x) = \frac{1}{w_i} f_i(x) \quad \text{such that} \quad \int_{-1}^1 p_i(x) dx = 1 \quad ; \quad i = 1, 2, 3 \quad (834)$$

we have that, with  $b = 1 + 2a$ :

$$w_1 = \frac{1}{a} \ln b \quad w_2 = \frac{2}{a^2 (b^2 - 1)} \quad w_3 = \frac{2b}{a^3 (b^2 - 1)^2} \quad (835)$$

and therefore

$$\begin{aligned} f(x) &= (f_1(x) + f_2(x) + f_3(x)) \cdot g(x) = (w_1 p_1(x) + w_2 p_2(x) + w_3 p_3(x)) \cdot g(x) = \\ &= w_t (\alpha_1 p_1(x) + \alpha_2 p_2(x) + \alpha_3 p_3(x)) \cdot g(x) \end{aligned} \quad (836)$$

where  $w_t = w_1 + w_2 + w_3$ ,

$$\alpha_i = \frac{w_i}{w_t} > 0 \quad ; \quad i = 1, 2, 3 \quad \text{and} \quad \sum_{i=1}^{i=3} \alpha_i = 1 \quad (837)$$

Thus, we set up the following algorithm:

- 1) Generate  $u \leftarrow Un(u|0, 1)$ ,
  - 1.1) if  $u \leq \alpha_1$  we sample  $x_g \sim p_1(x)$ ;
  - 1.2) if  $\alpha_1 < u \leq \alpha_1 + \alpha_2$  we sample  $x_g \sim p_2(x)$  and
  - 1.3) if  $\alpha_1 + \alpha_2 < u$  we sample  $x_g \sim p_3(x)$ ;
- 2) Generate  $w \leftarrow Un(w|0, g_M)$  where

$$g_M \equiv \max[g(x)] = g(x = -1) = 1 - \frac{b}{1 + b + b^2}$$

If  $w \leq g(x_g)$  we accept  $x_g$ ; otherwise we go back to step 1).

Let's see now to to sample from the densities  $p_i(x)$ . If  $u \leftarrow Un(u|0, 1)$  and

$$F_i(x) = \int_{-1}^x p_i(s) ds \quad i = 1, 2, 3 \quad (838)$$

then:

$$\begin{aligned} x \sim p_1(x) : F_1(x) &= 1 - \frac{\ln(1 + a(1 - x))}{\ln(b)} \quad \longrightarrow \quad x_g = \frac{1 + a - b^u}{a} \\ x \sim p_2(x) : F_2(x) &= \frac{b^2 - 1}{2(b - x)} - \frac{1}{2a} \quad \longrightarrow \quad x_g = b - \frac{a(b^2 - 1)}{1 + 2au} \\ x \sim p_3(x) : F_3(x) &= \frac{1}{4a(1 + a)} \left( \frac{(b + 1)^2}{(b - x)^2} - 1 \right) \quad \longrightarrow \quad x_g = b - \frac{b + 1}{[1 + 4a(1 + a)u]^{1/2}} \end{aligned} \quad (839)$$

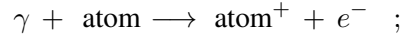


Once we have  $x_g$  we can deduce the remaining quantities of interest from the kinematic relations. In particular, the energy of the outgoing photon will be

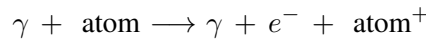
$$\epsilon_g = \frac{E'_g}{E} = \frac{1}{1 + a(1 - x_g)} \quad (840)$$

Last, we sample the azimuthal outgoing photon angle as  $\phi \leftarrow Un(u|0, 2\pi)$ .

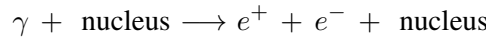
Even though in this example we are going to simulate only the Compton effect, there are other processes by which the photon interacts with matter. At low energies (essentially ionization energies:  $\leq E_\gamma \leq 100 \text{ KeV}$ ) the dominant interaction is the photoelectric effect



at intermediate energies ( $E_\gamma \sim 1 - 10 \text{ MeV}$ ) the Compton effect



and at high energies ( $E_\gamma \geq 100 \text{ MeV}$ ) the dominant one is pair production



To first approximation, the contribution of other processes is negligible. Then, at each step in the evolution of the photon along the material we have to decide first which interaction is going to occur next. The cross section is a measure of the interaction probability expressed in  $\text{cm}^2$  so, since the total interaction cross section will be in this case:

$$\sigma_t = \sigma_{\text{phot.}} + \sigma_{\text{Compt.}} + \sigma_{\text{pair}} \quad (841)$$

we decide upon the process  $i$  that is going to happen next with probability  $p_i = \sigma_i/\sigma_t$ ; that is,  $u \leftarrow Un(0, 1)$  and

- 1) if  $u \leq p_{\text{phot.}}$  we simulate the photoelectric interaction;
- 2) if  $p_{\text{phot.}} < u \leq (p_{\text{phot.}} + p_{\text{Compt.}})$ : we simulate the Compton effect and otherwise
- 3) we simulate the pair production

Once we have decided which interaction is going to happen next, we have to decide where. The probability that the photon interacts after traversing a distance  $x$  (cm) in the material is given by

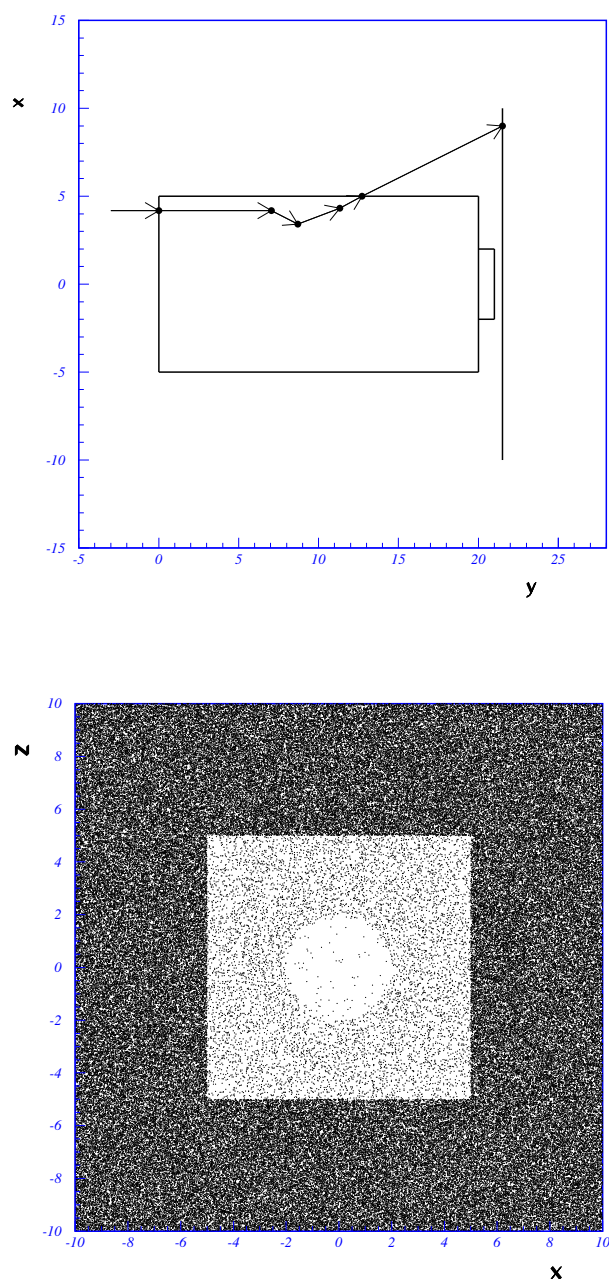
$$F_{int} = 1 - e^{-x/\lambda} \quad (842)$$

where  $\lambda$  is the *mean free path*. Being  $A$  the atomic mass number of the material,  $N_A$  the Avogadro's number,  $\rho$  the density of the material in  $\text{g}/\text{cm}^3$ , and  $\sigma$  the cross-section of the process under discussion, we have that

$$\lambda = \frac{A}{\rho N_A \sigma} \quad [\text{cm}] \quad (843)$$

Thus, if  $u \leftarrow Un(0, 1)$ , the next interaction is going to happen at  $x = -\lambda \ln u$  along the direction of the photon momentum.

As an example, we are going to simulate what happens when a beam of photons of energy  $E_\gamma = 1 \text{ MeV}$  (X rays) incide normally on the side of a rectangular block of carbon ( $Z = 6, A = 12.01, \rho = 2.26$ ) of  $10 \times 10 \text{ cm}^2$  surface y  $20 \text{ cm}$  depth. Behind the block, we have hidden an iron coin ( $Z = 26, A = 55.85, \rho = 7.87$ ) centered on the surface and in contact with it of  $2 \text{ cm}$  radius and  $1 \text{ cm}$  thickness. Last, at  $0.5 \text{ cm}$  from the coin there is a photographic film that collects the incident photons.



**Fig. 18:** The upper figure shows a sketch of the experimental set-up and the trajectory of one of the simulated photons until it is detected on the screen. The lower figure shows the density of photons collected at the screen for a initially generated sample  $10^5$  events.

The beam of photons is wider than the block of carbon so some of them will go right the way without interacting and will burn the film. We have assumed for simplicity that when the photon energy is below 0.01 MeV, the photoelectric effect is dominant and the ejected electron will be absorbed in the material. The photon will then be lost and we shall start with the next one. Last, an irrelevant technical issue: the angular variables of the photon after the interaction are referred to the direction of the incident photon so in each case we have to do the appropriate rotation.

Figure ?? (up) shows the sketch of the experimental set-up and the trajectory of one of the traced photons of the beam collected by the film. The *radiography* obtained after tracing 100,000 photons is shown in figure ?? (down). The black zone corresponds to photons that either go straight to the screen

or at some point leave the block before getting to the end. The mid zone are those photons that cross the carbon block and the central circle, with less interactions, those that cross the carbon block and afterward the iron coin.

### 4.3.2 An incoming flux of particles

Suppose we have a detector and we want to simulate a flux of isotropically distributed incoming particles. It is obvious that generating them one by one in the space and tracing them backwards is extremely inefficient. Consider a large cubic volume  $V$  that encloses the detector, both centered in the reference frame  $S_0$ . At time  $t_0$ , we have for particles uniformly distributed inside this volume that:

$$p(\mathbf{r}_0) d\mu_0 = \frac{1}{V} dx_0 dy_0 dz_0 \quad (844)$$

Assume now that the velocities are isotropically distributed; that is:

$$p(\mathbf{v}) d\mu_v = \frac{1}{4\pi} \sin \theta d\theta d\phi f(v) dv \quad (845)$$

with  $f(v)$  properly normalized. Under independence of positions and velocities at  $t_0$ , we have that:

$$p(\mathbf{r}_0, \mathbf{v}) d\mu_0 d\mu_v = \frac{1}{V} dx_0 dy_0 dz_0 \frac{1}{4\pi} \sin \theta d\theta d\phi f(v) dv \quad (846)$$

Given a square of surface  $S = (2l)^2$ , parallel to the  $(x, y)$  plane, centered at  $(0, 0, z_c)$  and well inside the volume  $V$ , we want to find the probability and distribution of particles that, coming from the top, cross the surface  $S$  in unit time.

For a particle having a coordinate  $z_0$  at  $t_0 = 0$ , we have that  $z(t) = z_0 + v_z t$ . The surface  $S$  is parallel to the plane  $(x, y)$  at  $z = z_c$  so particles will cross this plane at time  $t_c = (z_c - z_0)/v_z$  from above iff:

- 0)  $z_0 \geq z_c$ ; obvious for otherwise they are below the plane  $S$  at  $t_0 = 0$ ;
- 1)  $\theta \in [\pi/2, \pi)$ ; also obvious because if they are above  $S$  at  $t_0 = 0$  and cut the plane at some  $t > 0$ , the only way is that  $v_z = v \cos \theta < 0 \rightarrow \cos \theta < 0 \rightarrow \theta \in [\pi/2, \pi)$ .

But to cross the squared surface  $S$  of side  $2l$  we also need that

$$2) -l \leq x(t_c) = x_0 + v_x t_c \leq l \quad \text{and} \quad -l \leq y(t_c) = y_0 + v_y t_c \leq l$$

Last, we want particles crossing in unit time; that is  $t_c \in [0, 1]$  so  $0 \leq t_c = (z_c - z_0)/v_z \leq 1$  and therefore

$$3) z_0 \in [z_c, z_c - v \cos \theta]$$

Then, the desired subspace with conditions 1), 2), 3) is

$$\Omega_c = \{\theta \in [\pi/2, \pi); z_0 \in [z_c, z_c - v \cos \theta]; x_0 \in [-l - v_x t_c, l - v_x t_c]; y_0 \in [-l - v_y t_c, l - v_y t_c]\} \quad (847)$$

After integration:

$$\int_{z_c}^{z_c - v \cos \theta} dz_0 \int_{-l - v_x t_c}^{l - v_x t_c} dx_0 \int_{-l - v_y t_c}^{l - v_y t_c} dy_0 = -(2l)^2 v \cos \theta \quad (848)$$

Thus, we have that for the particles crossing the surface  $S = (2l)^2$  from above in unit time

$$p(\theta, \phi, v) d\theta d\phi dv = -\frac{(2l)^2}{V} \frac{1}{4\pi} \sin \theta \cos \theta d\theta d\phi f(v) v dv \quad (849)$$

with  $\theta \in [\pi/2, \pi)$  and  $\phi \in [0, 2\pi)$ . If we define the *average velocity*

$$E[v] = \int_{\Omega_v} v f(v) dv \quad (850)$$

the probability to have a cut per unit time is

$$P_{cut}(t_c \leq 1) = \int_{\Omega_c \times \Omega_v} p(\theta, \phi, v) d\theta d\phi dv = \frac{S E[v]}{4V} \quad (851)$$

and the pdf for the angular distribution of velocities (direction of crossing particles) is

$$p(\theta, \phi) d\theta d\phi = -\frac{1}{\pi} \sin \theta \cos \theta d\theta d\phi = \frac{1}{2\pi} d(\cos^2 \theta) d\phi \quad (852)$$

If we have a density of  $n$  particles per unit volume, the expected number of crossings per unit time due to the  $n_V = nV$  particles in the volume is

$$n_c = n_V P_{cut}(t_c \leq 1) = \frac{n E[v]}{4} S \quad (853)$$

so the *flux*, number of particles crossing the surface from one side per unit time and unit surface is

$$\Phi_c^0 = \frac{n_c}{S} = \frac{n E[v]}{4} \quad (854)$$

Note that the requirement that the particles cross the square surface  $S$  in a finite time ( $t_c \in [0, 1]$ ) modifies the angular distribution of the direction of particles. Instead of

$$p_1(\theta, \phi) \propto \sin \theta \quad ; \theta \in [0, \pi); \phi \in [0, 2\pi) \quad (855)$$

we have

$$p_2(\theta, \phi) \propto -\sin \theta \cos \theta \quad ; \theta \in [\pi/2, \pi); \phi \in [0, 2\pi) \quad (856)$$

The first one spans a solid angle of

$$\int_0^\pi d\theta \int_0^{2\pi} d\phi p_1(\theta, \phi) = 4\pi \quad (857)$$

while for the second one we have that

$$\int_0^{\pi/2} d\theta \int_0^{2\pi} d\phi p_2(\theta, \phi) = \pi \quad (858)$$

that is; one fourth the solid angle spanned by the sphere. Therefore, the flux expressed as number of particles crossing from one side of the square surface  $S$  per unit time and solid angle is

$$\Phi_c = \frac{\Phi_c^0}{\pi} = \frac{n_c}{\pi S} = \frac{n E[v]}{4\pi} \quad (859)$$

Thus, if we generate a total of  $n_T = 6n_c$  particles on the of the surface of a cube, each face of area  $S$ , with the angular distribution

$$p(\theta, \phi) d\theta d\phi = \frac{1}{2\pi} d(\cos^2 \theta) d\phi \quad (860)$$

for each surface with  $\theta \in [\pi/2, \pi)$  and  $\phi \in [0, 2\pi)$  defined with  $\vec{k}$  normal to the surface, the equivalent *generated flux* per unit time, unit surface and solid angle is  $\Phi_T = n_T/6\pi S$  and corresponds to a density of  $n = 2n_T/3SE[v]$  particles per unit volume.

---

**NOTE 8: Sampling some continuous distributions of interest.**

These are some procedures to sample from continuous distributions of interest. There are several algorithms for each case with efficiency depending on the parameters but those outlined here have in general high efficiency. In all cases, it is assumed that  $u \leftarrow Un(0, 1)$ .

- **Beta:**  $Be(x|\alpha, \beta); \alpha, \beta \in (0, \infty)$ :

$$p(x|\cdot) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} \mathbf{1}_{(0,1)}(x) \quad \longrightarrow \quad x = \frac{x_1}{x_1 + x_2} \quad (861)$$

where  $x_1 \leftarrow Ga(x|1/2, \alpha)$  and  $x_2 \leftarrow Ga(x|1/2, \beta)$

- **Cauchy:**  $Ca(x|\alpha, \beta); \alpha \in \mathcal{R}; \beta \in (0, \infty)$ :

$$p(x|\cdot) = \frac{\beta/\pi}{(1 + \beta^2(x - \alpha)^2)} \mathbf{1}_{(-\infty, \infty)}(x) \quad \longrightarrow \quad x = \alpha + \beta^{-1} \tan(\pi(u - 1/2)) \quad (862)$$

- **Chi-squared:** For  $\chi^2(x|\nu)$  see  $Ga(x|1/2, \nu/2)$ .

- **Dirichlet**  $Di(\mathbf{x}|\alpha); \dim(\mathbf{x}, \alpha) = n, \alpha_j \in (0, \infty), x_j \in (0, 1)$  and  $\sum_{j=1}^n x_j = 1$

$$p(\mathbf{x}|\alpha) = \frac{\Gamma(\alpha_1 + \dots + \alpha_n)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_n)} \prod_{j=1}^n x_j^{\alpha_j-1} \mathbf{1}_{(0,1)}(x_j) \quad \longrightarrow \quad \{x_j = z_j/z_0\}_{j=1}^n \quad (863)$$

where  $z_j \leftarrow Ga(z|1, \alpha_j)$  and  $z_0 = \sum_{j=1}^n z_j$ .

**Generalized Dirichlet**  $GDi(\mathbf{x}|\alpha, \beta); \dim(\beta) = n, \beta_j \in (0, \infty), \sum_{j=1}^{n-1} x_j < 1$

$$p(x_1, \dots, x_{n-1}|\alpha, \beta) = \prod_{i=1}^{n-1} \frac{\Gamma(\alpha_i + \beta_i)}{\Gamma(\alpha_i)\Gamma(\beta_i)} x_i^{\alpha_i-1} \left(1 - \sum_{k=1}^i x_k\right)^{\beta_i} \quad (864)$$

with

$$\gamma_i = \begin{cases} \beta_i - \alpha_{i+1} - \beta_{i+1} & \text{for } i = 1, 2, \dots, n-2 \\ \beta_{n-1} - 1 & \text{for } i = n-1 \end{cases} \quad (865)$$

When  $\beta_i = \alpha_{i+1} + \beta_{i+1}$  reduces to the usual Dirichlet. If  $z_k \leftarrow Be(z|\alpha_k, \beta_k)$  then  $x_k = z_k(1 - \sum_{j=1}^{k-1} x_j)$  for  $k = 1, \dots, n-1$  and  $x_n = 1 - \sum_{i=1}^{n-1} x_i$ .

- **Exponential:**  $Ex(x|\alpha); \alpha \in (0, \infty)$ :

$$p(x|\cdot) = \alpha \exp\{-\alpha x\} \mathbf{1}_{[0, \infty)}(x) \quad \longrightarrow \quad x = -\alpha^{-1} \ln u \quad (866)$$

- **Gamma Distribution**  $Ga(x|\alpha, \beta); \alpha, \beta \in (0, \infty)$ .

The probability density is

$$p(x|\alpha, \beta) = \frac{\alpha^\beta}{\Gamma(\beta)} e^{-\alpha x} x^{\beta-1} \mathbf{1}_{(0, \infty)}(x) \quad (867)$$

Note that  $Z = \alpha X \sim Ga(z|1, \beta)$  so let's see how to simulate a sampling of  $Ga(z|1, \beta)$  and, if  $\alpha \neq 1$ , take  $x = z/\alpha$ . Depending on the value of the parameter  $\beta$  we have that

- ▷  $\beta = 1$ : This is the Exponential distribution  $Ex(x|1)$  already discussed;
- ▷  $\beta = m \in \mathcal{N}$ : As we know, the sum  $X_s = X_1 + \dots + X_n$  of  $n$  independent random quantities  $X_i \sim Ga(x_i|\alpha, \beta_i); i = 1, \dots, n$  is a random quantity distributed as  $Ga(|x_s|\alpha, \beta_1 + \dots + \beta_n)$ . Thus, if we have  $m$  independent samplings  $x_i \leftarrow Ga(x|1, 1) = Ex(x|1)$ , that is

$$x_1 = -\ln u_1, \quad \dots, \quad x_m = -\ln u_m \tag{868}$$

with  $u_i \leftarrow Un(0, 1)$ , then

$$x_s = x_1 + x_2 \dots + x_m = -\ln \prod_{i=1}^m u_i \tag{869}$$

will be a sampling of  $Ga(x_s|1, \beta = m)$ .

- ▷  $\beta > 1 \in \mathcal{R}$ : Defining  $m = [\beta]$  we have that  $\beta = m + \delta$  with  $\delta \in [0, 1)$ . Then, if  $u_i \leftarrow Un(0, 1); i = 1, \dots, m$  and  $w \leftarrow Ga(w|1, \delta)$ ,

$$z = -\ln \prod_{i=1}^m u_i + w \tag{870}$$

will be a sampling from  $Ga(x|1, \beta)$ . The problem is reduced to get a sampling  $w \leftarrow Ga(w|1, \delta)$  with  $\delta \in (0, 1)$ .

- ▷  $0 < \beta < 1$ : In this case, for small values of  $x$  the density is dominated by  $p(x) \sim x^{\beta-1}$  and for large values by  $p(x) \sim e^{-x}$ . Let's then take the approximant

$$g(x) = x^{\beta-1} \mathbf{1}_{(0,1)}(x) + e^{-x} \mathbf{1}_{[1,\infty)}(x) \tag{871}$$

Defining

$$p_1(x) = \beta x^{\beta-1} \mathbf{1}_{(0,1)}(x) \quad \longrightarrow \quad F_1(x) = x^\beta \tag{872}$$

$$p_2(x) = e^{-(x-1)} \mathbf{1}_{[1,\infty)}(x) \quad \longrightarrow \quad F_2(x) = 1 - e^{-(x-1)} \tag{873}$$

$w_1 = e/(e + \beta)$  and  $w_2 = \beta/(e + \beta)$  we have that

$$g(x) = w_1 p_1(x) + w_2 p_2(x) \tag{874}$$

and therefore:

- 1)  $u_i \leftarrow Un(0, 1); i = 1, 2, 3$
- 2) If  $u_1 \leq w_1$ , set  $x = u_2^{1/\beta}$  and accept  $x$  if  $u_3 \leq e^{-x}$ ; otherwise go to 1);  
If  $u_1 > w_1$ , set  $x = 1 - \ln u_2$  and accept  $x$  if  $u_3 \leq x^{\beta-1}$ ; otherwise go to 1);

The sequence of accepted values will simulate a sampling from  $Ga(x|1, \beta)$ . It is easy to see that the generation efficiency is

$$\epsilon(\beta) = \frac{e}{e + \beta} \Gamma(\beta + 1) \tag{875}$$

and  $\epsilon_{min}(\beta \simeq 0.8) \simeq 0.72$ .

- **Laplace:**  $La(x|\alpha, \beta); \alpha \in \mathcal{R}, \beta \in (0, \infty)$ :

$$p(x|\alpha, \beta) = \frac{1}{2\beta} e^{-|x-\alpha|/\beta} \mathbf{1}_{(-\infty, \infty)}(x) \quad \longrightarrow \quad x = \begin{cases} \alpha + \beta \ln(2u) & \text{if } u < 1/2 \\ \alpha - \beta \ln(2(1-u)) & \text{if } u \geq 1/2 \end{cases} \tag{876}$$

- **Logistic:**  $Lg(x|\alpha, \beta)$ ;  $\alpha \in \mathcal{R}$ ;  $\beta \in (0, \infty)$ :

$$p(x|\cdot) = \beta \frac{\exp\{-\beta(x - \alpha)\}}{(1 + \exp\{-\beta(x - \alpha)\})^2} \mathbf{1}_{(-\infty, \infty)}(x) \quad \longrightarrow \quad x = \alpha + \beta^{-1} \ln\left(\frac{u}{1 - u}\right) \quad (877)$$

- **Normal Distribution**  $N(\mathbf{x}|\boldsymbol{\mu}, \mathbf{V})$ .

There are several procedures to generate samples from a Normal Distribution. Let's start with the one-dimensional case  $X \sim N(x|\mu, \sigma)$  considering two independent standardized random quantities  $X_i \sim N(x_i|0, 1)$ ;  $i = 1, 2$  [?] with joint density

$$p(x_1, x_2) = \frac{1}{2\pi} e^{-(x_1^2 + x_2^2)/2} \quad (878)$$

After the transformation  $X_1 = R \cos\Theta$  and  $X_2 = R \sin\Theta$  with  $R \in [0, \infty)$ ;  $\Theta \in [0, 2\pi)$  we have

$$p(r, \theta) = \frac{1}{2\pi} e^{-r^2/2} r \quad (879)$$

Clearly, both quantities  $R$  and  $\Theta$  are independent and their distribution functions

$$F_r(r) = 1 - e^{-r^2/2} \quad \text{and} \quad F_\theta(\theta) = \frac{\theta}{2\pi} \quad (880)$$

are easy to invert so, using then the *Inverse Transform* algorithm:

- 1)  $u_1 \leftarrow Un(0, 1)$  and  $u_2 \leftarrow Un(0, 1)$ ;
- 2)  $r = \sqrt{-2 \ln u_1}$  and  $\theta = 2\pi u_2$ ;
- 3)  $x_1 = r \cos\theta$  and  $x_2 = r \sin\theta$ .

Thus, we get two independent samplings  $x_1$  and  $x_2$  from  $N(x|0, 1)$  and

- 4)  $z_1 = \mu_1 + \sigma_1 x_1$  and  $z_2 = \mu_2 + \sigma_2 x_2$

will be two independent samplings from  $N(x|\mu_1, \sigma_1)$  and  $N(x|\mu_2, \sigma_2)$ .

For the n-dimensional case,  $\mathbf{X} \sim N(\mathbf{x}|\boldsymbol{\mu}, \mathbf{V})$  and  $\mathbf{V}$  the covariance matrix, we proceed from the conditional densities

$$p(\mathbf{x}|\boldsymbol{\mu}, \mathbf{V}) = p(x_n|x_{n-1}, x_{n-2}, \dots, x_1; \cdot) p(x_{n-1}|x_{n-2}, \dots, x_1; \cdot) \cdots p(x_1|\cdot) \quad (881)$$

For high dimensions this is a bit laborious and it is easier if we do first a bit of algebra. We know from *Cholesky's Factorization Theorem* that if  $\mathbf{V} \in \mathcal{R}^{n \times n}$  is a symmetric positive defined matrix there is a unique lower triangular matrix  $\mathbf{C}$ , with positive diagonal elements, such that  $\mathbf{V} = \mathbf{C}\mathbf{C}^T$ . Let then  $\mathbf{Y}$  be an n-dimensional random quantity distributed as  $N(\mathbf{y}|\mathbf{0}, \mathbf{I})$  and define a new random quantity

$$\mathbf{X} = \boldsymbol{\mu} + \mathbf{C}\mathbf{Y} \quad (882)$$

Then  $\mathbf{V}^{-1} = [\mathbf{C}^{-1}]^T \mathbf{C}^{-1}$  and

$$\mathbf{Y}^T \mathbf{Y} = (\mathbf{X} - \boldsymbol{\mu})^T [\mathbf{C}^{-1}]^T [\mathbf{C}^{-1}] (\mathbf{X} - \boldsymbol{\mu}) = (\mathbf{X} - \boldsymbol{\mu})^T [\mathbf{V}^{-1}] (\mathbf{X} - \boldsymbol{\mu}) \quad (883)$$

After some algebra, the elements of the matrix  $\mathbf{C}$  can be easily obtained as

$$\mathbf{C}_{i1} = \frac{\mathbf{V}_{i1}}{\sqrt{\mathbf{V}_{11}}} \quad 1 \leq i \leq n \quad (884)$$

$$\mathbf{C}_{ij} = \frac{\mathbf{V}_{ij} - \sum_{k=1}^{j-1} C_{ik}C_{jk}}{C_{jj}} \quad 1 < j < i \leq n \quad (885)$$

$$\mathbf{C}_{ii} = \left( \mathbf{V}_{ii} - \sum_{k=1}^{i-1} C_{ik}^2 \right)^{1/2} \quad 1 < i \leq n \quad (886)$$

and, being lower triangular,  $C_{ij} = 0 \forall j > i$ . Thus, we have the following algorithm:

- 1) Get the matrix  $\mathbf{C}$  from the covariance matrix  $\mathbf{V}$ ;
- 2) Get  $n$  independent samplings  $z_i \leftarrow N(0, 1)$  with  $i = 1, \dots, n$ ;
- 3) Get  $x_i = \mu_i + \sum_{j=1}^n \mathbf{C}_{ij}z_j$

In particular, for a two-dimensional random quantity we have that

$$\mathbf{V} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \quad (887)$$

and therefore:

$$\mathbf{C}_{11} = \frac{\mathbf{V}_{11}}{\sqrt{\mathbf{V}_{11}}} = \sigma_1 \quad ; \quad \mathbf{C}_{12} = 0 \quad (888)$$

$$\mathbf{C}_{21} = \frac{\mathbf{V}_{21}}{\sqrt{\mathbf{V}_{11}}} = \rho\sigma_2 \quad ; \quad \mathbf{C}_{22} = (\mathbf{V}_{22} - \mathbf{C}_{21}^2)^{1/2} = \sigma_2\sqrt{1 - \rho^2} \quad (889)$$

so:

$$\mathbf{C} = \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{1 - \rho^2} \end{pmatrix} \quad (890)$$

Then, if  $z_{1,2} \leftarrow N(z|0, 1)$  we have that:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} z_1\sigma_1 \\ \sigma_2(z_1\rho + z_2\sqrt{1 - \rho^2}) \end{pmatrix} \quad (891)$$

- **Pareto:**  $Pa(x|\alpha, \beta); \alpha, \beta \in (0, \infty)$ :

$$p(x|\cdot) = \alpha\beta^\alpha x^{-(\alpha+1)} \mathbf{1}_{(\beta, \infty)}(x) \quad \longrightarrow \quad x = \beta u^{-1/\alpha} \quad (892)$$

- **Snedecor:**  $Sn(x|\alpha, \beta); \alpha, \beta \in (0, \infty)$ :

$$p(x|\cdot) \propto x^{\alpha/2-1} (\beta + \alpha x)^{-(\alpha+\beta)/2} \mathbf{1}_{(0, \infty)}(x) \quad \longrightarrow \quad x = \frac{x_1/\alpha}{x_2/\beta} \quad (893)$$

where  $x_1 \leftarrow Ga(x|1/2, \alpha/2)$  and  $x_2 \leftarrow Ga(x|1/2, \beta/2)$ .

- **Student**  $St(x|\nu); \nu \in (0, \infty)$ :

$$p(x|\cdot) \propto \frac{1}{(1 + x^2/\nu)^{(\nu+1)/2}} \mathbf{1}_{(-\infty, \infty)}(x) \quad \longrightarrow \quad x = \sqrt{\nu(u_1^{-2/\nu} - 1)} \sin(2\pi u_2) \quad (894)$$

where  $u_{1,2} \leftarrow Un(0, 1)$ .

- **Uniform**  $Un(x|a, b); a < b \in \mathcal{R}$

$$p(x|\cdot) = (b - a)^{-1} \mathbf{1}_{[a, b]}(x) \quad \longrightarrow \quad x = (b - 1) + a u \quad (895)$$



- **Weibull:**  $We(x|\alpha, \beta); \alpha, \beta \in (0, \infty)$ :

$$p(x|\cdot) = \alpha\beta^\alpha x^{\alpha-1} \exp\{-(x/\beta)^\alpha\} \mathbf{1}_{(0,\infty)}(x) \quad \longrightarrow \quad x = \beta (-\ln u)^{1/\alpha} \quad (896)$$

#### 4.4 Markov Chain Monte Carlo

With the methods we have used up to now we can simulate samples from distributions that are more or less easy to handle. Markov Chain Monte Carlo allows to sample from more complicated distributions. The basic idea is to consider each sampling as a state of a system that evolves in consecutive steps of a Markov Chain converging (asymptotically) to the desired distribution. In the simplest version were introduced by Metropolis in the 1950s and were generalized by Hastings in the 1970s.

Let's start for simplicity with a discrete distribution. Suppose that we want a sampling of size  $n$  from the distribution

$$P(X = k) = \pi_k \quad \text{with} \quad k = 1, 2, \dots, N \quad (897)$$

that is, from the *probability vector*

$$\pi = (\pi_1, \pi_2, \dots, \pi_N) \quad ; \quad \pi_i \in [0, 1] \quad \forall i = 1, \dots, N \quad \text{and} \quad \sum_{i=1}^N \pi_i = 1 \quad (898)$$

and assume that it is difficult to generate a sample from this distribution by other procedures. Then, we may start from a sample of size  $n$  generated from a simpler distribution; for instance, a Discrete Uniform with

$$P_0(X = k) = \frac{1}{N} \quad ; \quad \forall k \quad (899)$$

and from the sample obtained  $\{n_1, n_2, \dots, n_N\}$ , where  $n = \sum_{i=1}^N n_i$ , we form the *initial sample probability vector*

$$\boldsymbol{\pi}^{(0)} = (\pi_1^{(0)}, \pi_2^{(0)}, \dots, \pi_N^{(0)}) = (n_1/n, n_2/n, \dots, n_N/n) \quad (900)$$

Once we have the  $n$  events distributed in the  $N$  classes of the sample space  $\Omega = \{1, 2, \dots, N\}$  we just have to redistribute them according to some criteria in different steps so that eventually we have a sample of size  $n$  drawn from the desired distribution  $P(X = k) = \pi_k$ .

We can consider the process of redistribution as an evolving system such that, if at step  $i$  the system is described by the probability vector  $\pi^{(i)}$ , the new state at step  $i + 1$ , described by  $\pi^{(i+1)}$ , depends only on the present state of the system ( $i$ ) and not on the previous ones; that is, as a Markov Chain. Thus, we start from the state  $\pi^{(0)}$  and the aim is to find a *Transition Matrix*  $\mathbf{P}$ , of dimension  $N \times N$ , such that  $\pi^{(i+1)} = \pi^{(i)} \mathbf{P}$  and allows us to reach the desired state  $\pi$ . The matrix  $\mathbf{P}$  is

$$\mathbf{P} = \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1N} \\ P_{21} & P_{22} & \cdots & P_{2N} \\ \vdots & \vdots & \cdots & \vdots \\ P_{N1} & P_{N2} & \cdots & P_{NN} \end{pmatrix} \quad (901)$$

where each element  $(\mathbf{P})_{ij} = P(i \rightarrow j) \in [0, 1]$  represents the probability for an event in class  $i$  to move to class  $j$  in one step. Clearly, at any step in the evolution the probability that an event in class  $i$  goes to any other class  $j = 1, \dots, N$  is 1 so

$$\sum_{j=1}^N (\mathbf{P})_{ij} = \sum_{j=1}^N P(i \rightarrow j) = 1 \quad (902)$$

and therefore is a *Probability Matrix*. If the Markov Chain is:

- 1) *irreducible*; that is, all the states of the system communicate among themselves;
- 2) *ergodic*; that is, the states are:
  - 2.1) *recurrent*: being at one state we shall return to it at some point in the evolution with probability 1;
  - 2.2) *positive*: we shall return to it in a finite number of steps in the evolution;
  - 2.3) *aperiodic*: the system is not trapped in cycles;

then there is a *stationary distribution*  $\pi$  such that:

- 1)  $\pi = \pi \mathbf{P}$  ;
- 2) Starting at any arbitrary state  $\pi^{(0)}$  of the system, the sequence

$$\begin{aligned} \pi^{(0)} \\ \pi^{(1)} &= \pi^{(0)} \mathbf{P} \\ \pi^{(2)} &= \pi^{(1)} \mathbf{P} = \pi^{(0)} \mathbf{P}^2 \\ &\vdots \\ \pi^{(n)} &= \pi^{(0)} \mathbf{P}^n \\ &\vdots \end{aligned}$$

converges asymptotically to the *fix vector*  $\pi$ ;

- 3)

$$\lim_{n \rightarrow \infty} \mathbf{P}^n = \begin{pmatrix} \pi_1 & \pi_2 & \cdots & \pi_N \\ \pi_1 & \pi_2 & \cdots & \pi_N \\ \vdots & \vdots & \dots & \vdots \\ \pi_1 & \pi_2 & \cdots & \pi_N \end{pmatrix} \quad (903)$$

There are infinite ways to choose the transition matrix  $\mathbf{P}$ . A *sufficient* (although not necessary) condition for this matrix to describe a Markov Chain with fixed vector  $\pi$  is that the *Detailed Balance* condition is satisfied (i.e.; a reversible evolution); that is

$$\pi_i (\mathbf{P})_{ij} = \pi_j (\mathbf{P})_{ji} \quad \iff \quad \pi_i P(i \rightarrow j) = \pi_j P(j \rightarrow i) \quad (904)$$

It is clear that if this condition is satisfied, then  $\pi$  is a fixed vector since:

$$\pi \mathbf{P} = \left( \sum_{i=1}^N \pi_i (\mathbf{P})_{i1}, \sum_{i=1}^N \pi_i (\mathbf{P})_{i2}, \dots, \sum_{i=1}^N \pi_i (\mathbf{P})_{iN} \right) = \pi \quad (905)$$

due to the fact that

$$\sum_{i=1}^N \pi_i (\mathbf{P})_{ik} = \sum_{i=1}^N \pi_k (\mathbf{P})_{ki} = \pi_k \quad \text{for} \quad k = 1, 2, \dots, N \quad (906)$$

Imposing the *Detailed Balance* condition, we have freedom to choose the elements  $(\mathbf{P})_{ij}$ . We can obviously take  $(\mathbf{P})_{ij} = \pi_j$  so that it is satisfied trivially ( $\pi_i \pi_j = \pi_j \pi_i$ ) but this means that being at class  $i$  we shall select the new possible class  $j$  with probability  $P(i \rightarrow j) = \pi_j$  and, therefore, to sample directly the desired distribution that, in principle, we do not know how to do. The basic idea of Markov Chain Monte Carlo simulation is to take

$$(\mathbf{P})_{ij} = q(j|i) \cdot a_{ij} \quad (907)$$

where

$q(j|i)$ : is a probability law to select the possible new class  $j = 1, \dots, N$  for an event that is actually in class  $i$ ;

$a_{ij}$ : is the probability to accept the proposed new class  $j$  for an event that is at  $i$  taken such that the *Detailed Balance* condition is satisfied for the desired distribution  $\pi$ .

Thus, at each step in the evolution, for an event that is in class  $i$  we propose a new class  $j$  to go according to the probability  $q(j|i)$  and accept the transition with probability  $a_{ij}$ . Otherwise, we reject the transition and leave the event in the class where it was. The Metropolis-Hastings [?] algorithm consists in taking the acceptance function

$$a_{ij} = \min \left\{ 1, \frac{\pi_j q(i|j)}{\pi_i q(j|i)} \right\} \quad (908)$$

It is clear that this election of  $a_{ij}$  satisfies the *Detailed Balance* condition. Indeed, if  $\pi_i q(j|i) > \pi_j q(i|j)$  we have that:

$$a_{ij} = \min \left\{ 1, \frac{\pi_j q(i|j)}{\pi_i q(j|i)} \right\} = \frac{\pi_j q(i|j)}{\pi_i q(j|i)} \quad \text{and} \quad a_{ji} = \min \left\{ 1, \frac{\pi_i q(j|i)}{\pi_j q(i|j)} \right\} = 1 \quad (909)$$

and therefore:

$$\pi_i (\mathbf{P})_{ij} = \pi_i q(j|i) a_{ij} = \pi_i q(j|i) \frac{\pi_j \cdot q(i|j)}{\pi_i \cdot q(j|i)} = \quad (910)$$

$$= \pi_j q(i|j) = \pi_j q(i|j) a_{ji} = \pi_j (\mathbf{P})_{ji} \quad (911)$$

The same holds if  $\pi_i q(j|i) < \pi_j q(i|j)$  and is trivial if both sides are equal. Clearly, if  $q(i|j) = \pi_i$  then  $a_{ij} = 1$  so the closer  $q(i|j)$  is to the desired distribution the better.

A particularly simple case is to choose a symmetric probability  $q(j|i) = q(i|j)$  [?]

$$a_{ij} = \min \left\{ 1, \frac{\pi_j}{\pi_i} \right\} \quad (912)$$

In both cases, it is clear that since the acceptance of the proposed class depends upon the ratio  $\pi_j/\pi_i$ , the normalization of the desired probability is not important.

The previous expressions are directly applicable in the case we want to sample an absolute continuous random quantity  $X \sim \pi(x)$ . If reversibility holds,  $p(x'|x)\pi(x) = p(x|x')\pi(x')$  and therefore

$$\int_X p(x'|x) dx' = 1 \quad \longrightarrow \quad \int_X p(x'|x)\pi(x) dx = \pi(x') \int_X p(x|x') dx = \pi(x') \quad (913)$$

The transition kernel is expressed as

$$p(x'|x) \equiv p(x \rightarrow x') = q(x'|x) \cdot a(x \rightarrow x') \quad (914)$$

and the acceptance probability given by

$$a(x \rightarrow x') = \min \left\{ 1, \frac{\pi(x') q(x|x')}{\pi(x) q(x'|x)} \right\} \quad (915)$$

Let's see one example.

**Example 4.11: The Binomial Distribution.** Suppose we want a sampling of size  $n$  of the random quantity  $X \sim Bi(x|N, \theta)$  Since  $x = 0, 1, \dots, N$  we have  $i = 1, 2, \dots, N + 1$  classes and the desired probability vector, of dimension  $N + 1$ , is

$$\pi = (p_0, p_1, \dots, p_N) \quad \text{where} \quad p_k = P(X = k|\cdot) = \binom{N}{k} \theta^k (1 - \theta)^{N-k} \quad (916)$$

Let's take for this example  $N = 10$  (that is, 11 classes),  $\theta = 0.45$  and  $n = 100,000$ . We start from a sampling of size  $n$  from a uniform distribution (figure ??-1). At each step of the evolution we swap over the  $n$  generated events. For an event that is in bin  $j$  we choose a new possible bin to go  $j = 1, \dots, 10$  with uniform probability  $q(j|i)$ . Suppose that we look at an event in bin  $i = 7$  and choose  $j$  with equal probability among the 10 possible bins. If, for instance,  $j = 2$ , then we accept the move with probability

$$a_{72} = a(7 \rightarrow 2) \min \left( 1, \frac{\pi_2 = p_2}{\pi_7 = p_7} \right) = 0.026 \quad (917)$$

if, on the other hand, we have  $j = 6$ ,

$$a_{76} = a(7 \rightarrow 6) \min \left( 1, \frac{\pi_6 = p_6}{\pi_7 = p_7} \right) = 1. \quad (918)$$

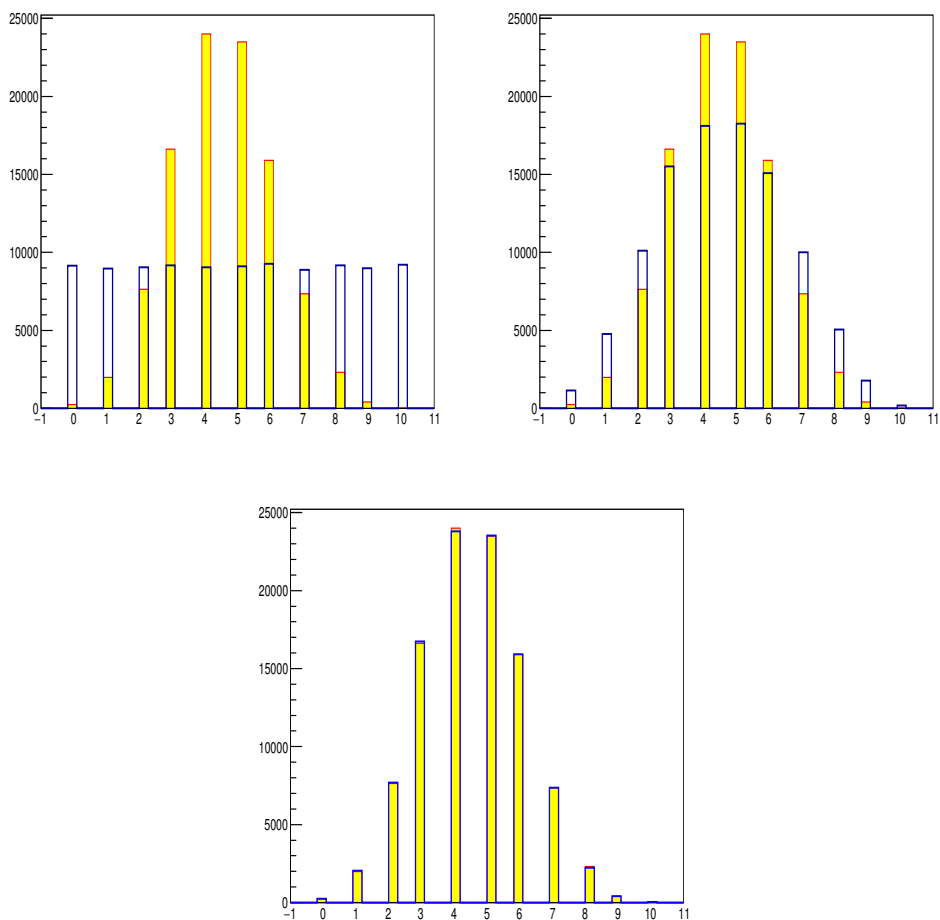
so we make the move of the event. After two swaps over all the sample we have the distribution shown in figure ??-2 and after 100 swaps that shown in figure ??-3, both compared to the desired distribution:

$$\begin{aligned} \pi^{(0)} &= (0.091, 0.090, 0.090, 0.092, 0.091, 0.091, 0.093, 0.089, 0.092, 0.090, 0.092) \\ \pi^{(2)} &= (0.012, 0.048, 0.101, 0.155, 0.181, 0.182, 0.151, 0.100, 0.050, 0.018, 0.002) \\ \pi^{(100)} &= (0.002, 0.020, 0.077, 0.167, 0.238, 0.235, 0.159, 0.074, 0.022, 0.004, 0.000) \\ \pi &= (0.000, 0.021, 0.076, 0.166, 0.238, 0.234, 0.160, 0.075, 0.023, 0.004, 0.000) \end{aligned}$$

The evolution of the moments, in this case the mean value and the variance with the number of steps is shown in figure ?? together with the Kullback-Leibler logarithmic discrepancy between each state and the new one defined as

$$\delta_{KL}\{\pi|\pi^n\} = \sum_{k=1}^{10} \pi_k^{(n)} \ln \frac{\pi_k^{(n)}}{\pi_k} \quad (919)$$

As the previous example shows, we have to let the system evolve some steps (i.e.; some initial sweeps for “burn-out” or “thermalization”) to reach the stable running conditions and get close to the stationary distribution after starting from an arbitrary state. Once this is achieved, each step in the evolution will be a sampling from the desired distribution so we do not have necessarily to generate a sample of the desired size to start with. In fact, we usually don't do that; we choose one admissible state and let the system evolve. Thus, for instance if we want a sample of  $X \sim p(x|\cdot)$  with  $x \in \Omega_X$ , we may start with a value  $x_0 \in \Omega_X$ . At a given step  $i$  the system will be in the state  $\{x\}$  and at the step  $i + 1$  the system will be in a new state  $\{x'\}$  if we accept the change  $x \rightarrow x'$  or in the state  $\{x\}$  if we do not accept it. After thermalization, each trial will simulate a sampling of  $X \sim p(x|\cdot)$ . Obviously, the sequence of states of the



**Fig. 19:** Distributions at steps 0, 2 and 100 (figs. 1,2,3; blue) of the Markov Chain with the desired Binomial distribution superimposed in yellow.

system is not independent so, if correlations are important for the evaluation of the quantities of interest, it is a common practice to reduce them by taking for the evaluations one out of few steps.

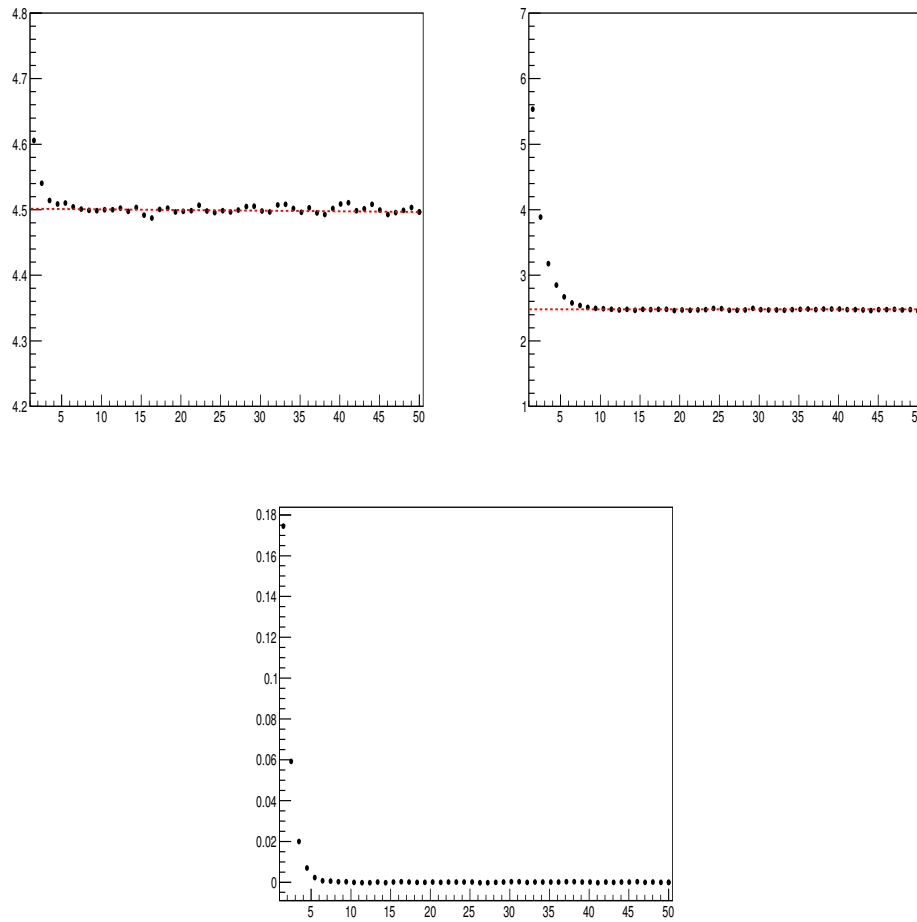
As for the *thermalization* steps, there is no universal criteria to tell whether stable conditions have been achieved. One may look, for instance, at the evolution of the discrepancy between the desired probability distribution and the probability vector of the state of the system and at the moments of the distribution evaluated with a fraction of the last steps. More details about that are given in [?]. It is interesting also to look at the acceptance rate; i.e. the number of accepted new values over the number of trials. If the rate is low, the proposed new values are rejected with high probability (are far away from the more likely ones) and therefore the chain will mix slowly. On the contrary, a high rate indicates that the steps are short, successive samplings move slowly around the space and therefore the convergence is slow. In both cases we should think about tuning the parameters of the generation.

**Example 4.12: The Beta distribution.**

Let’s simulate a sample of size  $10^7$  from a Beta distribution  $Be(x|4, 2)$ ; that is:

$$\pi(x) \propto x^3 (1 - x) \quad \text{with} \quad x \in [0, 1] \tag{920}$$

In this case, we start from the admissible state  $\{x = 0.3\}$  and select a new possible state  $x'$  from the density



**Fig. 20:** Distributions of the mean value, variance and logarithmic discrepancy vs the number of steps. For the first two, the red line indicates what is expected for the Binomial distribution.

$q(x'|x) = 2x'$ ; not symmetric and independent of  $x$ . Thus we generate a new possible state as

$$F_q(x') = \int_0^{x'} q(s|x) ds = \int_0^{x'} 2s ds = x'^2 \quad \longrightarrow \quad x' = u^{1/2} \quad \text{with} \quad u \leftarrow Un(0, 1) \quad (921)$$

The acceptance function will then be

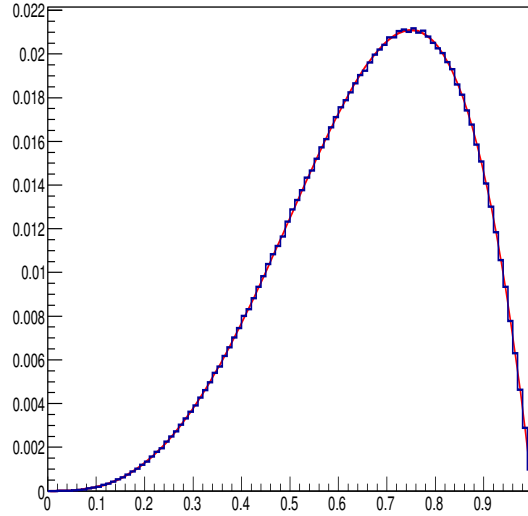
$$a(x \rightarrow x') = \min \left\{ 1, \frac{\pi(x') \cdot q(x|x')}{\pi(x) \cdot q(x'|x)} \right\} = \min \left\{ 1, \frac{x'^2(1-x')}{x^2(1-x)} \right\} \quad (922)$$

depending on which we set at the state  $i + 1$  the system in  $x$  or  $x'$ . After evolving the system for thermalization, the distribution is shown in figure ?? where we have taken one  $x$  value out of 5 consecutive ones. The red line shows the desired distribution  $Be(x|4, 2)$ .

**Example 4.13: Path Integrals in Quantum Mechanics.**

In Feynman’s formulation of non-relativistic Quantum Mechanics, the probability amplitude to find a particle in  $x_f$  at time  $t_f$  when at  $t_i$  was at  $x_i$  is given by

$$K(x_f, t_f|x_i, t_i) = \int_{paths} e^{i/\hbar S[x(t)]} D[x(t)] \quad (923)$$



**Fig. 21:** Sampling of the Beta distribution  $Be(x|4, 2)$  (blue) of the example 4.12 compared to the desired distribution (continuous line).

where the integral is performed over all possible trajectories  $x(t)$  that connect the initial state  $x_i = x(t_i)$  with the final state  $x_f = x(t_f)$ ,  $S[x(t)]$  is the classic action functional

$$S[x(t)] = \int_{t_i}^{t_f} L(\dot{x}, x, t) dt \quad (924)$$

that corresponds to each trajectory and  $L(\dot{x}, x, t)$  is the Lagrangian of the particle. All trajectories contribute to the amplitude with the same weight but different phase. In principle, small differences in the trajectories cause big changes in the action compared to  $\hbar$  and, due to the oscillatory nature of the phase, their contributions cancel. However, the action does not change, to first order, for the trajectories in a neighborhood of the one for which the action is extremal and, since they have similar phases (compared to  $\hbar$ ) their contributions will not cancel. The set of trajectories around the extremal one that produce changes in the action of the order of  $\hbar$  define the limits of classical mechanics and allow to recover laws expressed as the *Extremal Action Principle*.

The transition amplitude ( *propagator* ) allows to get the wave-function  $\Psi(x_f, t_f)$  from  $\Psi(x_i, t_i)$  as:

$$\Psi(x_f, t_f) = \int K(x_f, t_f | x_i, t_i) \Psi(x_i, t_i) dx_i \quad \text{for } t_f > t_i \quad (925)$$

In non-relativistic Quantum Mechanics there are no trajectories evolving backwards in time so in the definition of the propagator a Heaviside step function  $\theta(t_f - t_i)$  is implicit. Is this clear from this equation that  $K(x_f, t | x_i, t) = \delta(x_f - x_i)$ .

For a local Lagrangian (additive actions), it holds that:

$$K(x_f, t_f | x_i, t_i) = \int K(x_f, t_f | x, t) K(x, t | x_i, t_i) dx \quad (926)$$

analogous expression to the Chapman-Kolmogorov equations that are satisfied by the conditional probabilities of a Markov process. If the Lagrangian is not local, the evolution of the system will depend on the intermediate states and this equation will not be true. On the other hand, if the Classical Lagrangian has no explicit time dependence the propagator admits an expansion ( *Feynman-Kac Expansion Theorem* ) in terms of a complete set of eigenfunctions  $\{\phi_n\}$  of the Hamiltonian as:

$$K(x_f, t_f | x_i, t_i) = \sum_n e^{-i/\hbar E_n (t_f - t_i)} \phi_n(x_f) \phi_n^*(x_i) \quad (927)$$

where the sum is understood as a sum for discrete eigenvalues and as an integral for continuous eigenvalues. Last, remember that expected value of an operator  $A(x)$  is given by:

$$\langle A \rangle = \int A[x(t)] e^{i/\hbar S[x(t)]} D[x(t)] / \int e^{i/\hbar S[x(t)]} D[x(t)] \quad (928)$$

Let's see how to do the integral over paths to get the propagator in a one-dimensional problem. For a particle that follows a trajectory  $x(t)$  between  $x_i = x(t_i)$  and  $x_f = x(t_f)$  under the action of a potential  $V(x(t))$ , the Lagrangian is:

$$L(\dot{x}, x, t) = \frac{1}{2} m \dot{x}(t)^2 - V(x(t)) \quad (929)$$

and the corresponding action:

$$S[x(t)] = \int_{t_i}^{t_f} \left( \frac{1}{2} m \dot{x}(t)^2 - V(x(t)) \right) dt \quad (930)$$

so we have for the propagator:

$$\begin{aligned} K(x_f, t_f | x_i, t_i) &= \int_{Tr} e^{i/\hbar S[x(t)]} D[x(t)] = \\ &= \int_{Tr} \exp \left\{ \frac{i}{\hbar} \int_{t_i}^{t_f} \left( \frac{1}{2} m \dot{x}(t)^2 - V(x(t)) \right) dt \right\} D[x(t)] \end{aligned} \quad (931)$$

where the integral is performed over the set  $Tr$  of all possible trajectories that start at  $x_i = x(t_i)$  and end at  $x_f = x(t_f)$ . Following Feynman, a way to perform this integrals is to make a partition of the interval  $(t_i, t_f)$  in  $N$  subintervals of equal length  $\epsilon$  (figure ??); that is, with

$$\epsilon = \frac{t_f - t_i}{N} \quad \text{so that} \quad t_j - t_{j-1} = \epsilon ; \quad j = 1, 2, \dots, N \quad (932)$$

Thus, if we identify  $t_0 = t_i$  and  $t_N = t_f$  we have that

$$[t_i, t_f] = \cup_{j=0}^{N-1} [t_j, t_{j+1}] \quad (933)$$

On each interval  $[t_j, t_{j+1})$ , the possible trajectories  $x(t)$  are approximated by straight segments so they are defined by the sequence

$$\{x_0 = x_i = x(t_i), x_1 = x(t_1), x_2 = x(t_2), \dots, x_{N-1} = x(t_{N-1}), x_N = x_f = x(t_f)\}$$

Obviously, the trajectories so defined are continuous but not differentiable so we have to redefine the velocity. An appropriate prescription is to substitute  $\dot{x}(t_j)$  by

$$\dot{x}(t_j) \longrightarrow \frac{x_j - x_{j-1}}{\epsilon} \quad (934)$$

so the action is finally expressed as:

$$S_N[x(t)] = \epsilon \sum_{j=1}^N \left[ \frac{1}{2} m \left( \frac{x_j - x_{j-1}}{\epsilon} \right)^2 - V(x_j) \right] \quad (935)$$

Last, the integral over all possible trajectories that start at  $x_0$  and end at  $x_N$  is translated in this space with discretized time axis as an integral over the quantities  $x_1, x_2, \dots, x_{N-1}$  so the differential measure for the trajectories  $D[x(t)]$  is substituted by

$$D[x(t)] \longrightarrow A_N \prod_{j=1}^{j=N-1} dx_j \quad (936)$$

with  $A_N$  a normalization factor. The propagator is finally expressed as:

$$K_N(x_f, t_f | x_i, t_i) =$$



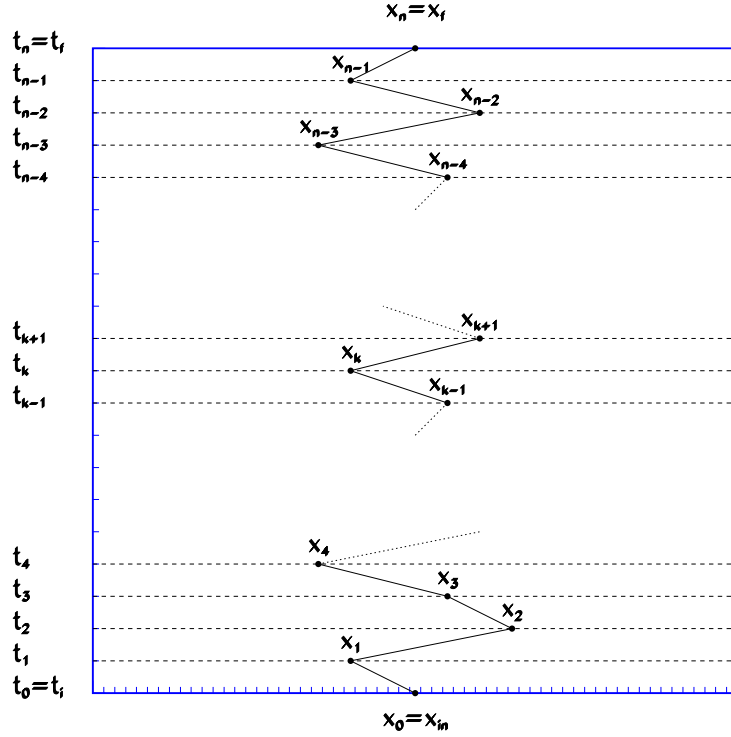


Fig. 22: Trajectory in a space with discretized time.

$$= A_N \int_{x_1} dx_1 \cdots \int_{x_{N-1}} dx_{N-1} \exp \left\{ \frac{i}{\hbar} \sum_{j=1}^N \left[ \frac{1}{2} m \left( \frac{x_j - x_{j-1}}{\epsilon} \right)^2 - V(x_j) \right] \cdot \epsilon \right\} \quad (937)$$

After doing the integrals, taking the limit  $\epsilon \rightarrow 0$  (or  $N \rightarrow \infty$  since the product  $N\epsilon = (t_f - t_i)$  is fixed) we get the expression of the propagator. Last, note that the interpretation of the integral over trajectories as the limit of a multiple Riemann integral is valid only in Cartesian coordinates.

To derive the propagator from path integrals is a complex problem and there are few potentials that can be treated exactly (the “simple” Coulomb potential for instance was solved in 1982). The Monte Carlo method allows to attack satisfactorily this type of problems but before we have first to convert the complex integral in a positive real function. Since the propagator is an analytic function of time, it can be extended to the whole complex plane of  $t$  and then perform a rotation of the time axis (Wick’s rotation) integrating along

$$\tau = e^{i\pi/2} t = i t \quad ; \text{ that is, } t \longrightarrow -i\tau \quad (938)$$

Taking as prescription the analytical extension over the imaginary time axis, the oscillatory exponentials are converted to decreasing exponentials, the results are consistent with those derived by other formulations (Schrodinger or Heisenberg for instance) and it is manifest the analogy with the partition function of Statistical Mechanics. Then, the action is expressed as:

$$S[x(t)] \longrightarrow i \int_{\tau_i}^{\tau_f} \left( \frac{1}{2} m \dot{x}(t)^2 + V(x(t)) \right) dt \quad (939)$$

Note that the integration limits are real as corresponds to integrate along the imaginary axis and not to just a simple change of variables. After partitioning the time interval, the propagator is expressed as:

$$K_N(x_f, t_f | x_i, t_i) = A_N \int_{x_1} dx_1 \cdots \int_{x_{N-1}} dx_{N-1} \exp \left\{ -\frac{1}{\hbar} S_N(x_0, x_1, \dots, x_N) \right\} \quad (940)$$

where

$$S_N(x_0, x_1, \dots, x_N) = \sum_{j=1}^N \left[ \frac{1}{2} m \left( \frac{x_j - x_{j-1}}{\epsilon} \right)^2 + V(x_j) \right] \cdot \epsilon \quad (941)$$

and the expected value of an operator  $A(x)$  will be given by:

$$\langle A \rangle = \frac{\int \prod_{j=1}^{j=N-1} dx_j A(x_0, x_1, \dots, x_N) \exp \left\{ -\frac{1}{\hbar} S_N(x_0, x_1, \dots, x_N) \right\}}{\int \prod_{j=1}^{j=N-1} dx_j \exp \left\{ -\frac{1}{\hbar} S_N(x_0, x_1, \dots, x_N) \right\}} \quad (942)$$

Our goal is to generate  $N_{gen}$  trajectories with the Metropolis criteria according to

$$p(x_0, x_1, \dots, x_N) \propto \exp \left\{ -\frac{1}{\hbar} S_N(x_0, x_1, \dots, x_N) \right\} \quad (943)$$

Then, over these trajectories we shall evaluate the expected value of the operators of interest  $A(x)$

$$\langle A \rangle = \frac{1}{N_{gen}} \sum_{k=1}^{N_{gen}} A(x_0, x_1^{(k)}, \dots, x_{N-1}^{(k)}, x_N) \quad (944)$$

Last, note that if we take  $(\tau_f, x_f) = (\tau, x)$  and  $(\tau_i, x_i) = (0, x)$  in the Feynman-Kac expansion we have that

$$K(x, \tau | x, 0) = \sum_n e^{-1/\hbar E_n \tau} \phi_n(x) \phi_n^*(x) \quad (945)$$

and therefore, for sufficiently large times

$$K(x, \tau | x, 0) \approx e^{-1/\hbar E_0 \tau} \phi_0(x) \phi_0^*(x) + \dots \quad (946)$$

so basically only the fundamental state will contribute.

Well, now we have everything we need. Let's apply all that first to an harmonic potential

$$V(x) = \frac{1}{2} k x^2 \quad (947)$$

(see figure ??-left) so the discretized action will be:

$$S_N(x_0, x_1, \dots, x_N) = \sum_{j=1}^N \left[ \frac{1}{2} m \left( \frac{x_j - x_{j-1}}{\epsilon} \right)^2 + \frac{1}{2} k x_j^2 \right] \cdot \epsilon \quad (948)$$

To estimate the energy of the fundamental state we use the *Virial Theorem*. Since:

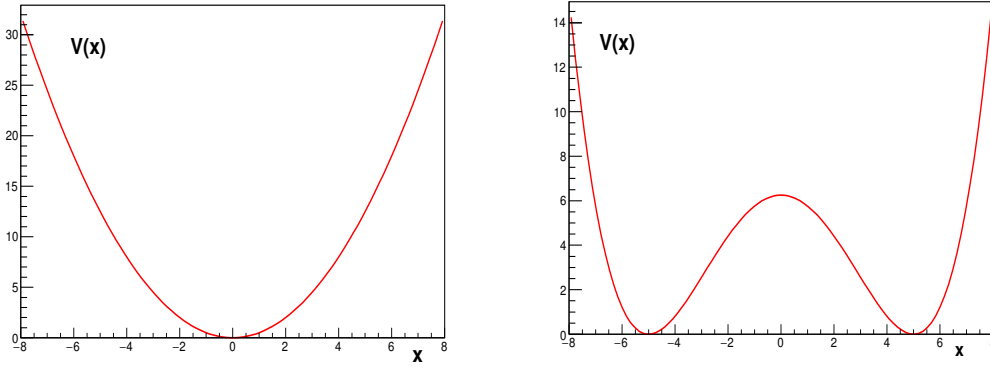
$$\langle T \rangle_\Psi = \frac{1}{2} \langle \vec{x} \cdot \vec{\nabla} V(\vec{x}) \rangle_\Psi \quad (949)$$

we have that  $\langle T \rangle_\Psi = \langle V \rangle_\Psi$  and therefore

$$\langle E \rangle_\Psi = \langle T \rangle_\Psi + \langle V \rangle_\Psi = k \langle x^2 \rangle_\Psi \quad (950)$$

In this example we shall take  $m = k = 1$ .

We start with an initial trajectory from  $x_0 = x(t_i) = 0$  to  $x_f = x(t_f) = 0$  and the intermediate values  $x_1, x_2, \dots, x_{N-1}$  drawn from  $Un(-10., 10.)$ , sufficiently large in this case since their support is  $(-\infty, \infty)$ . For the parameters of the grid we took  $\epsilon = 0.25$  and  $N = 2000$ . The parameter  $\epsilon$  has to be small enough so that the results we obtain are close to what we should have for a continuous time and  $N$  sufficiently large so that  $\tau = N\epsilon$  is large enough to isolate the contribution of the fundamental state. With this election we have that  $\tau = 2000 \cdot 0.25 = 500$ . Obviously, we have to check the stability of the result varying both parameters. Once the grid is fixed, we sweep



**Fig. 23:** Potential wells studied in the example 4.13.

over all the points  $x_1, x_2, \dots, x_{N-1}$  of the trajectory and for each  $x_j, j = 1, \dots, N-1$  we propose a new candidate  $x'_j$  with support  $\Delta$ . Then, taking  $\hbar = 1$  we have that:

$$P(x_j \rightarrow x'_j) = \exp \left\{ -S_N(x_0, x_1, \dots, x'_j, \dots, x_N) \right\} \tag{951}$$

and

$$P(x_j \rightarrow x_j) = \exp \left\{ -S_N(x_0, x_1, \dots, x_j, \dots, x_N) \right\} \tag{952}$$

so the acceptance function will be:

$$a(x_j \rightarrow x'_j) = \min \left\{ 1, \frac{P(x_j \rightarrow x'_j)}{P(x_j \rightarrow x_j)} \right\} \tag{953}$$

Obviously we do not have to evaluate the sum over all the nodes because when dealing with node  $j$ , only the intervals  $(x_{j-1}, x_j)$  and  $(x_j, x_{j+1})$  contribute to the sum. Thus, at node  $j$  we have to evaluate

$$a(x_j \rightarrow x'_j) = \min \left\{ 1, \exp \left\{ -S_N(x_{j-1}, x'_j, x_{j+1}) + S_N(x_{j-1}, x_j, x_{j+1}) \right\} \right\} \tag{954}$$

Last, the trajectories obtained with the Metropolis algorithm will follow eventually the desired distribution  $p(x_0, x_1, \dots, x_N)$  in the asymptotic limit. To have a reasonable approximation to that we shall not use the first  $N_{term}$  trajectories ( *thermalization* ). In this case we have taken  $N_{term} = 1000$  and again, we should check the the stability of the result. After this, we have generated  $N_{gen} = 3000$  and, to reduce correlations, we took one out of three for the evaluations; that is  $N_{used} = 1000$  trajectories, each one determined by  $N = 2000$  nodes. The distribution of the accepted values  $x_j$  will be an approximation to the probability to find the particle at position  $x$  for the fundamental state; that is,  $|\Psi_0(x)|^2$ . Figure ?? shows the results of the simulation compared to

$$|\Psi_0(x)|^2 \propto e^{-x^2} \tag{955}$$

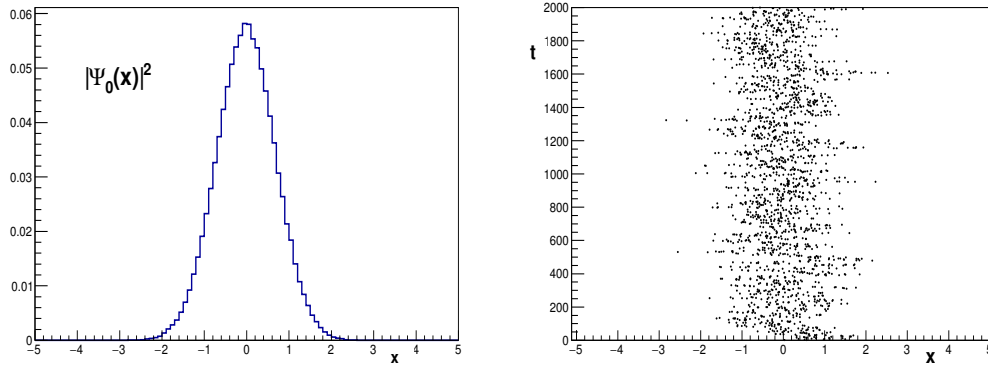
together with one of the many trajectories generated. The sampling average  $\langle x^2 \rangle = 0.486$  is a good approximation to the energy of the fundamental state  $E_0 = 0.5$ .

As a second example, we have considered the potential well

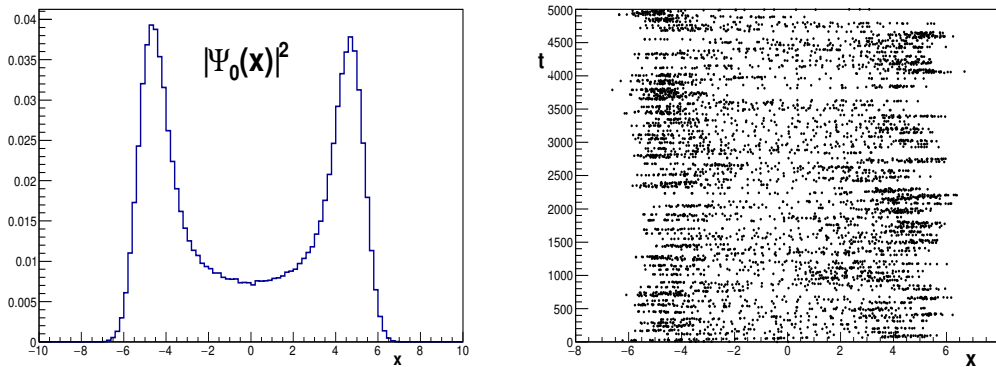
$$V(x) = \frac{a^2}{4} \left[ (x/a)^2 - 1 \right]^2 \tag{956}$$

(see figure ??-right) and, again from the Virial Theorem:

$$\langle E \rangle_\Psi = \frac{3}{4a^2} \langle x^4 \rangle_\Psi - \langle x^2 \rangle_\Psi + \frac{a^2}{4} \tag{957}$$



**Fig. 24:** Squared norm of the fundamental state wave-function for the harmonic potential and one of the simulated trajectories.



**Fig. 25:** Squared norm of the fundamental state wave-function for the quadratic potential and one of the simulated trajectories.

We took  $a = 5$ , a grid of  $N = 9000$  nodes and  $\epsilon = 0.25$  (so  $\tau = 2250$ ), and as before  $N_{term} = 1000$ ,  $N_{gen} = 3000$  and  $N_{used} = 1000$ . From the generated trajectories we have the sample moments  $\langle x^2 \rangle = 16.4264$  and  $\langle x^4 \rangle = 361.4756$  so the estimated fundamental state energy is  $\langle E_0 \rangle = 0.668$  to be compared with the exact result  $E_0 = 0.697$ . The norm of the wave-function for the fundamental state is shown in figure ?? together with one of the simulated trajectories exhibiting the tunneling between the two wells.

#### 4.4.1 Sampling from Conditionals and Gibbs Sampling

In many cases, the distribution of an n-dimensional random quantity is either not known explicitly or difficult to sample directly but sampling the conditionals is easier. In fact, sometimes it may help to introduce an additional random quantity and consider the conditional densities (see the example 4.14). Consider then the n-dimensional random quantity  $\mathbf{X} = (X_1, \dots, X_n)$  with density  $p(x_1, \dots, x_n)$ , the usually simpler conditional densities

$$p(x_1|x_2, x_3, \dots, x_n) \quad p(x_2|x_1, x_3, \dots, x_n) \quad \dots \quad p(x_n|x_1, x_2, \dots, x_{n-1}) \quad (958)$$

and an arbitrary initial value  $\mathbf{x}^0 = \{x_1^0, x_2^0, \dots, x_n^0\} \in \Omega_{\mathbf{X}}$ . If we take the approximating density  $q(x_1, x_2, \dots, x_n)$  and the conditional densities

$$q(x_1|x_2, x_3, \dots, x_n) \quad q(x_2|x_1, x_3, \dots, x_n) \quad \dots \quad q(x_n|x_1, x_2, \dots, x_{n-1}) \quad (959)$$

we generate for  $x_1$  a proposed new value  $x_1^1$  from  $q(x_1|x_2^0, x_3^0, \dots, x_n^0)$  and accept the change with probability

$$\begin{aligned} a(x_1^0 \rightarrow x_1^1) &= \min \left\{ 1, \frac{p(x_1^1, x_2^0, x_3^0, \dots, x_n^0)q(x_1^0, x_2^0, x_3^0, \dots, x_n^0)}{p(x_1^0, x_2^0, x_3^0, \dots, x_n^0)q(x_1^1, x_2^0, x_3^0, \dots, x_n^0)} \right\} = \\ &= \min \left\{ 1, \frac{p(x_1^1|x_2^0, x_3^0, \dots, x_n^0)q(x_1^0|x_2^0, x_3^0, \dots, x_n^0)}{p(x_1^0|x_2^0, x_3^0, \dots, x_n^0)q(x_1^1|x_2^0, x_3^0, \dots, x_n^0)} \right\} \end{aligned} \quad (960)$$

After this step, let's denote the value of  $x_1$  by  $x_1'$  (that is,  $x_1' = x_1^1$  or  $x_1' = x_1^0$  if it was not accepted). Then, we proceed with  $x_2$ . We generate a proposed new value  $x_2^1$  from  $q(x_2|x_1', x_3^0, \dots, x_n^0)$  and accept the change with probability

$$\begin{aligned} a(x_2^0 \rightarrow x_2^1) &= \min \left\{ 1, \frac{p(x_1', x_2^1, x_3^0, \dots, x_n^0)q(x_1', x_2^0, x_3^0, \dots, x_n^0)}{p(x_1', x_2^0, x_3^0, \dots, x_n^0)q(x_1', x_2^1, x_3^0, \dots, x_n^0)} \right\} = \\ &= \min \left\{ 1, \frac{p(x_2^1|x_1', x_3^0, \dots, x_n^0)q(x_2^0|x_1', x_3^0, \dots, x_n^0)}{p(x_2^0|x_1', x_3^0, \dots, x_n^0)q(x_2^1|x_1', x_3^0, \dots, x_n^0)} \right\} \end{aligned} \quad (961)$$

After we run over all the variables, we are in a new state  $\{x_1', x_2', \dots, x_n'\}$  and repeat the whole procedure until we consider that stability has been reached so that we are sufficiently close to sample the desired density. The same procedure can be applied if we consider more convenient to express the density

$$p(x_1, x_2, x_3, \dots, x_n) = p(x_n|x_{n-1}, \dots, x_2, \dots, x_1) \cdots p(x_2|x_1) p(x_1) \quad (962)$$

Obviously, we need only one one admissible starting value  $x_1^0$ .

Gibbs sampling is a particular case of this approach and consists on sampling sequentially all the random quantities directly from the conditional densities; that is:

$$q(x_i|x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) = p(x_i|x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \quad (963)$$

so the acceptance factor  $a(x \rightarrow x') = 1$ . This is particularly useful for Bayesian inference since, in more than one dimension, densities are usually specified in conditional form after the ordering of parameters.

**Example 4.14:** Sometimes it may be convenient to introduce additional random quantities in the problem to ease the treatment. Look for instance at the Student's distribution  $X \sim St(x|\nu)$  with

$$p(x|\nu) \propto (1 + x^2/\nu)^{-(\nu+1)/2} \quad (964)$$

Since

$$\int_0^\infty e^{-au} u^{b-1} du = \Gamma(b) a^{-b} \quad (965)$$

we can introduce an additional random quantity  $U \sim Ga(u|a, b)$  in the problem with  $a = 1 + x^2/\nu$  and  $b = (\nu+1)/2$  so that

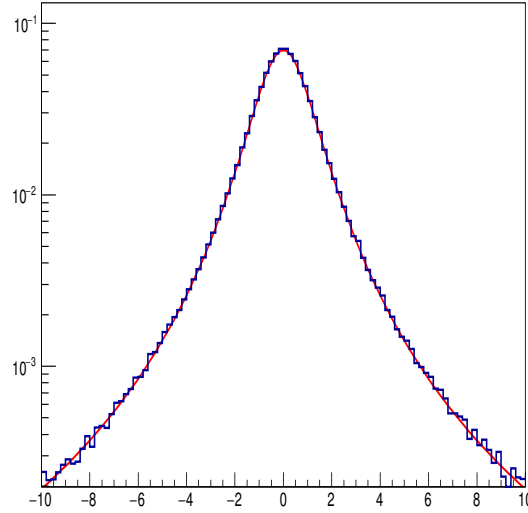
$$p(x, u|\nu) \propto e^{-au} u^{b-1} \quad \text{and} \quad p(x) \propto \int_0^\infty p(x, u|\nu) du \propto a^{-b} = (1 + x^2/\nu)^{-(\nu+1)/2} \quad (966)$$

The conditional densities are

$$p(x|u, \nu) = \frac{p(x, u|\nu)}{p(u|\nu)} = N(x|0, \sigma) \quad ; \quad \sigma^2 = \nu(2u)^{-1} \quad (967)$$

$$p(u|x, \nu) = \frac{p(x, u|\nu)}{p(x|\nu)} = Ga(u|a, b) \quad ; \quad a = 1 + x^2/\nu, \quad b = (\nu + 1)/2 \quad (968)$$

so, if we start with an arbitrary initial value  $x \in \mathcal{R}$ , we can



**Fig. 26:** Sampling of the Student's Distribution  $St(x|2)$  (blue) compared to the desired distribution (red).

- 1) Sample  $U|X: u \leftarrow Ga(u|a, b)$  with  $a = 1 + x^2/\nu$  and  $b = (\nu + 1)/2$
- 2) Sample  $X|U: x \leftarrow N(x|0, \sigma)$  with  $\sigma^2 = \nu(2u)^{-1}$

and repeat the procedure so that, after equilibrium,  $X \sim St(x|\nu)$ . We can obviously start from  $u \in \mathcal{R}$  and reverse the steps 1) and 2). Thus, instead of sampling from the Student's distribution we may sample the conditional densities: Normal and a Gamma distributions. Following this approach for  $\nu = 2$  and  $10^3$  thermalization sweeps (far beyond the needs), the results of  $10^6$  draws are shown in figure ?? together with the Student's distribution  $St(x|2)$ .

**Example 4.15:** We have  $j = 1, \dots, J$  groups of observations each with a sample of size  $n_j$ ; that is  $\mathbf{x}_j = \{x_{1j}, x_{2j}, \dots, x_{n_j j}\}$ . Within each of the  $J$  groups, observations are considered an exchangeable sequence and assumed to be drawn from a distribution  $x_{i,j} \sim N(x|\mu_j, \sigma^2)$  where  $i = 1, \dots, n_j$ . Then:

$$p(\mathbf{x}_j|\mu_j, \sigma) = \prod_{i=1}^{n_j} N(x_{ij}|\mu_j, \sigma^2) \propto \sigma^{-n_j} \exp \left\{ -\sum_{i=1}^{n_j} \frac{(x_{ij} - \mu_j)^2}{2\sigma^2} \right\} \quad (969)$$

Then, for the  $J$  groups we have the parameters  $\mu = \{\mu_1, \dots, \mu_J\}$  that, in turn, are also considered as an exchangeable sequence drawn from a parent distribution  $\mu_j \sim N(\mu_j|\mu, \sigma_\mu^2)$ . We reparameterize the model in terms of  $\eta = \sigma_\mu^{-2}$  and  $\phi = \sigma^{-2}$  and consider conjugated priors for the parameters considered independent; that is

$$\pi(\mu, \eta, \phi) = N(\mu|\mu_0, \sigma_0^2) Ga(\eta|c, d) Ga(\phi|a, b) \quad (970)$$

Introducing the sample means  $\bar{x}_j = n_j^{-1} \sum_{i=1}^{n_j} x_{ij}$ ,  $\bar{x} = J^{-1} \sum_{j=1}^J \bar{x}_j$  and defining  $\bar{\mu} = J^{-1} \sum_{j=1}^J \mu_j$  and defining After some simple algebra it is easy to see that the marginal densities are:

$$\mu_j \sim N \left( \frac{n_j \sigma_\mu^2 \bar{x}_j + \mu_0 \sigma_0^2}{n_j \sigma_\mu^2 + \sigma_0^2}, \frac{\sigma_\mu^2 \sigma_0^2}{n_j \sigma_\mu^2 + \sigma_0^2} \right) \quad (971)$$

$$\mu \sim N \left( \frac{\sigma_\mu^2 \mu_0 + \sigma_0^2 J \bar{\mu}}{\sigma_\mu^2 + J \sigma_0^2}, \frac{\sigma_\mu^2 \sigma_0^2}{\sigma_\mu^2 + J \sigma_0^2} \right) \quad (972)$$

$$\eta = \sigma_\mu^{-2} \sim Ga \left( \frac{1}{2} \sum_{j=1}^J (\mu_j - \mu)^2 + c, \frac{J}{2} + d \right) \quad (973)$$

$$\phi = \sigma^{-2} \sim Ga\left(\frac{1}{2} \sum_{j=1}^J \sum_{i=1}^{n_j} (x_{ij} - \mu_j)^2 + a, \frac{1}{2} \sum_{j=1}^J n_j + b\right) \quad (974)$$

Thus, we set initially the parameters  $\{\mu_0, \sigma_0, a, b, c, d\}$  and then, at each step

- 1 Get  $\{\mu_1, \dots, \mu_J\}$  each as  $\mu_j \sim N(\cdot, \cdot)$
- 2 Get  $\mu \sim N(\cdot, \cdot)$
- 3 Get  $\sigma_\mu = \eta^{-1/2}$  with  $\eta \sim Ga(\cdot, \cdot)$
- 4 Get  $\sigma = \phi^{-1/2}$  with  $\phi \sim Ga(\cdot, \cdot)$

and repeat the sequence until equilibrium is reached and samplings for evaluations can be done.

#### 4.5 Evaluation of definite integrals

A frequent use of Monte Carlo sampling is the evaluation of definite integrals. Certainly, there are many numerical methods for this purpose and for low dimensions they usually give a better precision when fairly compared. In those cases one rarely uses Monte Carlo... although sometimes the domain of integration has a very complicated expression and the Monte Carlo implementation is far easier. However, as we have seen the uncertainty of Monte Carlo estimations decreases with the sampling size  $N$  as  $1/\sqrt{N}$  regardless the number of dimensions so, at some point, it becomes superior. And, besides that, it is fairly easy to estimate the accuracy of the evaluation. Let's see in this section the main ideas.

Suppose we have the  $n$ -dimensional definite integral

$$I = \int_{\Omega} f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \quad (975)$$

where  $(x_1, x_2, \dots, x_n) \in \Omega$  and  $f(x_1, x_2, \dots, x_n)$  is a Riemann integrable function. If we consider a random quantity  $\mathbf{X} = (X_1, X_2, \dots, X_n)$  with distribution function  $P(\mathbf{x})$  and support in  $\Omega$ , the mathematical expectation of  $\mathbf{Y} = g(\mathbf{X}) \equiv f(\mathbf{X})/p(\mathbf{X})$  is given by

$$E[\mathbf{Y}] = \int_{\Omega} g(\mathbf{x}) dP(\mathbf{x}) = \int_{\Omega} \frac{f(\mathbf{x})}{p(\mathbf{x})} dP(\mathbf{x}) = \int_{\Omega} f(\mathbf{x}) d\mathbf{x} = I \quad (976)$$

Thus, if we have a sampling  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ , of size  $N$ , of the random quantity  $\mathbf{X}$  under  $P(\mathbf{x})$  we know, by the Law of Large Numbers that, as  $N \rightarrow \infty$ , the sample means

$$I_N^{(1)} = \frac{1}{N} \sum_{i=1}^N g(\mathbf{x}_i) \quad \text{and} \quad I_N^{(2)} = \frac{1}{N} \sum_{i=1}^N g^2(\mathbf{x}_i) \quad (977)$$

converge respectively to  $E[Y]$  (and therefore to  $I$ ) and to  $E[Y^2]$  (as for the rest, all needed conditions for existence are assumed to hold). Furthermore, if we define

$$SI^2 = \frac{1}{N} \left( I_N^{(2)} - (I_N^{(1)})^2 \right) \quad (978)$$

we know by the Central Limit Theorem that the random quantity

$$\mathbf{Z} = \frac{I_N^{(1)} - I}{SI} \quad (979)$$

is, in the limit  $N \rightarrow \infty$ , distributed as  $N(x|0, 1)$ . Thus, Monte Carlo integration provides a simple way to estimate the integral  $I$  and a quantify the accuracy. Depending on the problem at hand, you can envisage several tricks to further improve the accuracy. For instance, if  $g(\mathbf{x})$  is a function "close" to  $f(\mathbf{x})$  with the same support and known integral  $I_g$  one can write

$$I = \int_{\Omega} f(\mathbf{x}) d\mathbf{x} = \int_{\Omega} (f(\mathbf{x}) - g(\mathbf{x})) d\mathbf{x} + I_g \quad (980)$$

and in consequence estimate the value of the integral as

$$\tilde{I} = \frac{1}{N} \sum_{i=1}^N (f(\mathbf{x}_i) - g(\mathbf{x}_i)) + I_g \quad (981)$$

reducing the uncertainty.

## References

- [1] Bogachev, V.I. (2006); *Measure Theory*; Springer.
- [2] Gut, A.(2013); *Probability: A Graduate Course*; Springer Texts in Statistics.
- [3] Lebesgue, H.L. (1926); "*Sur le développement de la notion d'intégrale*".
- [4] Mana, C. (2017); *Probability and Statistics for Particle Physics*; Springer; UNITEXT for Physics.
- [5] D'Agostini G. (2003); *Bayesian Reasoning in Data Analysis*; World Scientific Publishing.
- [6] James F. (2006); *Statistical Methods in Experimental Physics*; World Scientific Pub. Co.
- [7] Bernardo J.M. (1996); The Concept of Exchangeability and its Applications; Far East Journal of Mathematical Sciences 4, 111-121 ([www.uv.es/~bernardo/Exchangeability.pdf](http://www.uv.es/~bernardo/Exchangeability.pdf)).
- [8] Bernardo J.M., Smith A.F.M (1994); *Bayesian Theory*; John Wiley & Sons;
- [9] Kass R.E., Wasserman L. (1996); *The Selection of Prior Distributions by Formal Rules*; J. Am. Stat. Assoc. V91, No 453, 1343-1370.
- [10] Jeffreys H. (1939); *Theory of Probability*; Oxford Univ. Press.
- [11] Jaynes E.T. (1964); *Prior Probabilities and Transformation Groups*,NSF G23778.
- [12] Stone M. (1965); *Right Haar Measures for Convergence in Probability to Invariant Posterior Distributions*; Annals of Math. Stat. 36, 440-453.
- [13] Stone M. (1970); *Necessary and Sufficient Conditions for Convergence in Probability to Invariant Posterior Distributions*; Annals of Math. Stat. 41, 1349-1353.
- [14] Raiffa, H. and Schlaifer R. (1961); *Applied Statistical Decision Theory*; Cambridge MA: Harvard University Press.
- [15] Dalal S.R., Hall, W.J. (1983); J. Roy. Stat. Soc. Ser. B, 45, 278-286.
- [16] Welch, B.; Pears H. (1963) J. Roy. Stat. Soc. B25, 318-329.
- [17] Gosh M., Mukerjee R. (1984) Biometrika 84, 970-975.
- [18] Datta G.S., Ghosh M. (1996); Ann. Stat. 24, No 1, 141-159.
- [19] Datta G.S., Mukerjee R. (2004); *Probability Matching Priors and Higher Order Asymptotics*, Springer, New York.
- [20] Bernardo J.M. (1979); J. Roy. Stat. Soc. Ser. B 41, 113-147.
- [21] Berger J.O., Bernardo J.M., Sun D. (2009); Ann. Stat; 37, No. 2; 905-938.
- [22] Bernardo J.M., Ramón J.M. (1998); The Statistician 47; 1-35.
- [23] Berger J.O., Bernardo J.M., Sun D. (2012); Objective Priors for Discrete Parameter Spaces; J. Am. Stat. Assoc. 107, No 498, 636-648.
- [24] O'Hagan A. (1995); J. Roy. Stat. Soc. B57, 99-138.
- [25] Berger, J.O., Pericchi L.R. (1996); J. Am. Stat. Assoc. V91, No 433, 109-122.
- [26] Kass R.E., Raftery A.E. (1995); J. Am. Stat. Assoc. V90, No 430, 773-795.
- [27] Schwarz, G. (1978); Ann. Stat. 6, 461-464.
- [28] Feldman G.J. and Cousins R.D. (1997); arXiv:physics/9711021v2.
- [29] Berger, J.O., Pericchi L.R. (2004); Ann. Stat. V32, No 3, 841-869.
- [30] James F. (1980); *Monte Carlo theory and practice*, Rep. Prog. Phys. 43, 1145-1189.
- [31] Knuth D.E. (1981); *The Art of Computer Programming*, vol 2.; Addison-Wesley.



- [32] James F., Hoogland J., Kleiss R. (1999); *Comput. Phys. Comm* 2-3, 180-220.
- [33] L'Ecuyer P. (1998); *Handbook of Simulations*, Chap. 4, 93-137; Wiley.
- [34] Marsaglia G., Zaman A. (1987); *Toward a Universal Random Number Generator*, Florida State University Report FSU-SCRI-87-50.
- [35] Rubin D.B. (1981); *Ann. Stat.* 9, 130-134.
- [36] Box G.E.P., Müller M.E. (1958); *A Note on the Generation of Random Normal Deviates*, *Ann. Math. Stat.* 29, No 2, 610-611.
- [37] Hastings W.K. (1970); *Biometrika* 57, 97-109.
- [38] Metropolis N., Rosenbluth A.W., Rosenbluth M.W., Teller A.H., Teller E. (1953); *Journal of Chemical Physics* 21, 1087-1092.
- [39] Gelman A.B., Carlin J.S., Stern H.S. and Rubin D.B. (1995); *Bayesian Data Analysis*; Chapman & Hall.



## Organizing Committee

M. Aguilar (CIEMAT, Spain)  
L. Álvarez-Gaumé (CERN)  
E. Carrera (USFQ, Ecuador)  
M. Cerrada (CIEMAT, Spain)  
C. Dib (UTFSM, Chile)  
M.T. Dova (UNLP, Argentina)  
J. Ellis (King's College London, UK and CERN)  
N. Ellis (Schools Director, CERN (Chair))  
A. Gago-Medina (PUCP, Peru)  
M. Gandelman (UFRJ, Brazil)  
P. Garcia (CIEMAT, Spain)  
M. Losada (UAN, Colombia)  
M. Mulders (Schools Deputy-Director, CERN)  
K. Ross (Schools Administrator, CERN)  
G. Zanderighi (CERN)  
A. Zepeda (Cinvestav, Mexico)

## Local Organizing Committee

A. Aranda (FC-UCOL, Mexico)  
A. Ayala (ICN-UNAM, Mexico)  
L. Díaz-Cruz (FCFM-BUAP, Mexico)  
A. Fernandez (FCFM-BUAP and DPyC-SMF, Mexico)  
G. Herrera-Corral (CINVESTAV, Mexico)  
G. López-Castro (CINVESTAV, Mexico)  
G. Murguía (FC-UNAM, Mexico)  
M. Napsuciale (IF-UGTO, Mexico)  
A. Raya (IFM-UMSNH, Mexico)  
M. Tejeda-Yeomans((chair) DF-USON, Mexico)  
G. Toledo (IF-UNAM, Mexico)  
A. Zepeda (CINVESTAV and MCTP, Mexico)

## Lecturers

A. Ayala (ICN-UNAM, Mexico)  
L. Da Rold (CAB, CONICET/IB, Argentina)  
C. Garcia-Canal (UNLP, Argentina)  
G. Gonzalez (Louisiana State U., USA)  
C. Mañá (CIEMAT, Spain)  
M. Mangano (CERN)  
M. Mondragon (UNAM, Mexico)  
A. Pich (IFIC, U. Valencia - CSIC, Spain)  
R. Rosenfeld (IFT-UNESP, Brazil)  
F. Sanchez (IFAE/BIST, Spain)  
R. Shellard (CBPF, Brazil)  
P. Sphicas (CERN and U. of Athens, Greece)

## Discussion Leaders

J. Cobos-Martinez (LFTC, Universidade Cruzeiro do Sul, Brazil)

G. von Gersdorff (PUC Rio, Brazil)

R. Hernandez-Pinto (UAS, Mexico)

F. von der Pahlen (U. of Antioquia, Colombia)

A. Szykman (UNLP, Argentina)

## Students

Suxer Lázara ALFONSO GARCÍA  
Rafael AOUDE  
Rose Elisabeth ARDELL  
Denys Yen ARREBATO VILLAR  
Marvin ASCENCIO  
Hector BELLO MARTINEZ  
Anthony CALATAYUD  
Laura Randa CALCAGNI  
Giacomo CARIA  
Jose Carlos CASTILLO FALLAS  
Nhell Heder CERNA VELAZCO  
Alberto CHAVARRÍA  
Gilson CORREIA SILVA  
Juan Carlos DE HARO SANTOS  
María DI DOMENICO FRANCO  
Edgar DOMINGUEZ ROSAS  
Martha Cecilia DURAN OSUNA  
Sofia ESCOBAR  
Mariel ESTEVEZ  
Ignacio FABRE  
Matías FERNÁNDEZ LAKATOS  
Maria FRAIRE  
Vinicius FRANCO LIMA  
Jorge Luis GUTIÉRREZ SANTIAGO  
Simen HELLESUND  
Luis HERNANDEZ  
Matheus HOSTERT  
Joaquin HOYA  
Carlos JIMÉNEZ  
José Carlos JIMÉNEZ APAZA  
Christian LANER OGILVY  
Santeri LAURILA  
Otto LIPPMANN  
Edgardo MARBELLO SANTRICH  
Gino MARCECA  
Jorge MARTÍNEZ ARMAS  
Ivania MATURANA  
Annie MENESES GONZALEZ

Pedro Roberto MERCADO LOZANO  
Ignacio Alberto MONROY CANON  
Jose Andres MONROY MONTANEZ  
Manuel MORGADO  
Marcus O'FLAHERTY  
Roberto PADRON STEVENS  
Julia PENA  
Karla PENA RODRIGUEZ  
Adriana PÉREZ MARTÍNEZ  
Daniel Alejandro PÉREZ NAVARRO  
Elias Natanael POLANCO EUAN  
Jose Manuel QUIMBAYO GARCIA  
Raul Iraq RABADAN TREJO  
Norma Selomit RAMIREZ URIBE  
Rogelio REYES ALMANZA  
Pablo RIVADENEIRA  
David RIVERA  
Mayra Alejandra RIVERA RUIZ  
Juan Cristobal RIVERA VERGARA  
Jamerson RODRIGUES  
Arturo RODRÍGUEZ  
Antonio ROJAS RAMOS  
Camilo SALAZAR  
Jordi SALINAS SAN MARTÍN  
Marxil SÁNCHEZ GARCÍA  
Danilo SILVA DE ALBUQUERQUE  
Luiz Gustavo SILVA DE OLIVEIRA  
Freja THORESEN  
Stoyan TRILOV  
Luis VALENZUELA CAZARES  
Paulina Fernanda VALENZUELA CORONADO  
Adriana VASQUEZ  
Omar VAZQUEZ RUEDA  
Jorge Alberto VENZOR VALDIVIA  
Azarael YEBRA PÉREZ  
Rosa Luz ZAMORA PEINADO  
De Hua ZHU

# Posters

Presenter	Poster title
AOUDE, R. T.	A study of the doubly Cabibbo-suppressed $D^+ \rightarrow K^- K^+ K^+$ decay
ARDELL, R.	Measurements of top-quark pair differential cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector
BELLO MARTINEZ, H.	Revealing the source of the radial flow patterns in pp collisions using hard probes
CALATAYUD, A. ; ASCENCIO, M.	Invisible Neutrino Decay at DUNE
CALCAGNI, L.	Electromagnetic Shower Simulations
CORREIA SILVA, G.	Search for Long-Lived Gluinos in Compressed SUSY Scenarios
DE HARO, J.	Commissioning of the Fast Cooling System of the Cryogenic Underground Observatory for Rare Events
ESTEVEZ M.	A look through the $pp \rightarrow t\bar{t}b\bar{b}$ process in search of new physics at the LHC
FABRE, I.	Higgs pair production as a probe for BSM physics in the SM EFT framework at NNLO
FRANCO LIMA, V.	VELO Upgrade Sensor R&D
HERNANDEZ, L. A.	The magnetized shiny pre-equilibrium QGP
HOSTERT, M.	NuSTORM Decoherence and CP Violation
HOYA, J.	ATLAS $e/\gamma$ Trigger Performance in Run-2
JIMÉNEZ, C.	A method to get current algebra
JIMÉNEZ, J.	Quark Matter Effects on Neutron Stars
LANER, C.	Search for new phenomena in final states with jets and missing energy at CMS

<b>Presenter</b>	<b>Poster title</b>
MARTÍNEZ, J.	How much space does the top quark occupy?
MERCADO, P.	Magnetic Catalysis of a Finite-Size Pion Condensate
MONROY, I.A.	Charmed Baryon Spectroscopy at the LHCb experiment
O'FLAHERTY, M.	Neutron Yields from Neutrino Interactions on Oxygen
OLIVEIRA, L.	Central Exclusive Production on LHCb Experiment
PÉREZ MARTÍNEZ, A.	Higgs Potential with $S_3$ symmetry
POLANCO, E.	Dibaryon octet strong decay coupling constants sum rules with first-order SU(3) symmetry breaking
QUIMBAYO GARCIA, J.M.	Fermionic Dark Matter in a U(1)' model
RODRIGUES, J.	Inflation, axion dark matter and seesaw in a 3-3-1 model
ROJAS RAMOS, A.	Hadronic contributions to muon g-2
RIVADENEIRA, P.	Monte Carlo simulation chain for a scalar extension of the Standard Model
RIVERA RUIZ, M.A.	Dirac form factors and electric charge radii of baryons in the combined chiral and $1/N_c$ expansions
SALINAS, J.	Chiral asymmetry during the EWPT from CP-violating scattering of bubble walls
SÁNCHEZ GARCÍA, S.	The role of Majorana neutrinos in $M^+ \rightarrow M^- \ell^+ \ell^+$
THORESEN, F.	Flow analysis with the ALICE detector
TRILOV, S.	The CEDAR Detector Control And Safety Systems at NA62 – a rare Kaon decay experiment at the CERN SPS
VÁSQUEZ, A.	The Muon Telescope project
YEBRA, A.	Neutrino-electron scattering with a new source of CP violation
ZHU, D.	Detector Modules for the CMS Pixel Phase 1 Upgrade