VOLUME 9

UNDERSTANDING THE BASIC COMPONENTS OF THE UNIVERSE, ITS STRUCTURE & EVOLUTION

Topic Coordinators
María José Costa & Rainer Schödel

CSIC SCIENTIFIC CHALLENGES: TOWARDS 2030

Challenges coordinated by:
Jesús Marco de Lucas & M. Victoria Moreno-Arribas

VOLUME 9

UNDERSTANDING THE BASIC COMPONENTS OF THE UNIVERSE, ITS STRUCTURE & EVOLUTION

Catálogo de publicaciones de la Administración General del Estado:
https://cpage.mpr.gob.es

Editorial CSIC:
http://editorial.csic.es (correo: *publ@csic.es*)

GOBIERNO DE ESPAÑA · MINISTERIO DE CIENCIA E INNOVACIÓN · CSIC CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS · EDITORIAL CSIC

ISBN Vol. 9: 978-84-00-10754-3
ISBN O.C.: 978-84-00-10736-9
e-ISBN Vol. 9: 978-84-00-10755-0
e-ISBN O.C.: 978-84-00-10734-5
NIPO: 833-21-157-5
e-NIPO: 833-21-155-4
DL: M-2426-2021

Diseño y maquetación: gráfica futura

VOLUME 9
UNDERSTANDING THE BASIC COMPONENTS OF THE UNIVERSE, ITS STRUCTURE & EVOLUTION

Topic Coordinators
María José Costa & Rainer Schödel

CSIC SCIENTIFIC CHALLENGES: TOWARDS 2030

What are the major scientific challenges of the first half of the 21st century? Can we establish the priorities for the future? How should the scientific community tackle them?

This book presents the reflections of the Spanish National Research Council (CSIC) on 14 strategic themes established on the basis of their scientific impact and social importance.

Fundamental questions are addressed, including the origin of life, the exploration of the universe, artificial intelligence, the development of clean, safe and efficient energy or the understanding of brain function. The document identifies complex challenges in areas such as health and social sciences and the selected strategic themes cover both basic issues and potential applications of knowledge. Nearly 1,100 researchers from more than 100 CSIC centres and other institutions (public research organisations, universities, etc.) have participated in this analysis. All agree on the need for a multidisciplinary approach and the promotion of collaborative research to enable the implementation of ambitious projects focused on specific topics.

These 14 "White Papers", designed to serve as a frame of reference for the development of the institution's scientific strategy, will provide an insight into the research currently being accomplished at the CSIC, and at the same time, build a global vision of what will be the key scientific challenges over the next decade.

VOLUMES THAT MAKE UP THE WORK

1 *New Foundations for a Sustainable Global Society*
2 *Origins, (Co)Evolution, Diversity and Synthesis of Life*
3 *Genome & Epigenetics*
4 *Challenges in Biomedicine and Health*
5 *Brain, Mind & Behaviour*
6 *Sustainable Primary Production*
7 *Global Change Impacts*
8 *Clean, Safe and Efficient Energy*
9 *Understanding the Basic Components of the Universe, its Structure and Evolution*
10 *Digital and Complex Information*
11 *Artificial Intelligence, Robotics and Data Science*
12 *Our Future? Space, Colonization and Exploration*
13 *Ocean Science Challenges for 2030*
14 *Dynamic Earth: Probing the Past, Preparing for the Future*

CSIC scientific challenges: towards 2030
Challenges coordinated by:
Jesus Marco de Lucas & M. Victoria Moreno-Arribas

Volume 9
Understanding the Basic Components of the Universe, its Structure & Evolution

Topic Coordinators:
María José Costa (IFIC, CSIC – UV) and Rainer Schödel (IAA, CSIC)

Chapters Coordinators:
Sven Heinemeyer (IFT, CSIC – UAM); Alicia Calderón (IFCA, CSIC-UC); Alberto Casas (IFT, CSIC – UAM); Avelino Vicente (IFIC, CSIC – UV); M. Pilar Hernandez (IFIC, CSIC – UV); Laura Tolos (ICE, CSIC – IEEC); Rocío Vilar (IFCA, CSIC – UC); Martin Crocce (ICE, CSIC – IEEC); Jose Manuel Vílchez (IAA, CSIC); J. Miguel Mas-Hesse (CAB, CSIC – INTA); Francisco Najarro (CAB, CSIC – INTA); Javier R. Goicoechea (IFF, CSIC); Tomás Ortín (IFT, CSIC – UAM); Gianluca Calcagni (IEM, CSIC); Carlos Lacasta (IFIC, CSIC – UV); Lourdes Fábrega (ICMAB, CSIC); Alberto Enciso (ICMAT, CSIC – UAM – UCM – UC3M); Fernando Marchesano (IFT, CSIC – UAM); Nanda Rea (ICE, CSIC – IEEC); Miguel Pérez-Torres (IAA, CSIC)

Participant Centers:
Centro de Astrobiología (CAB, CSIC-INTA)
Centro Nacional de Microelectrónica (IBM-CNM, CSIC)
Instituto de Astrofísica de Andalucía (IAA, CSIC)
Instituto de Ciencias del Espacio (ICE, CSIC)
Instituto de Nanociencia y Materiales de Aragón (INMA, CSIC-UZ)
Instituto de Ciencia de Materiales de Barcelona (ICMAB, CSIC)
Instituto de Ciencia de Materiales de Madrid (ICMM, CSIC)
Instituto de Ciencias de Matemáticas (ICMAT, CSIC-UAM-UC3M-UCM)
Instituto de Estructura de la Materia (IEM, CSIC)
Instituto de Física de Cantabria (IFCA, CSIC-UC)
Instituto de Física Corpuscular (IFIC, CSIC-UV)
Instituto de Física Fundamental (IFF, CSIC)
Instituto de Física Teórica (IFT, UAM-CSIC)

ABSTRACT

Understanding the basic components of the Universe, its structure, and evolution is one of the noblest and most ambitious undertakings of humanity. The fundamental laws of nature are the basis of all technology. Finding and understanding them requires exploring both the elementary components of matter at the smallest scales and the observable Universe at the largest scales. Only through understanding the laws that dictated the first instances of the Universe we will be able to find the ultimate reason for its actual appearance and future fate. Challenges in physics are intimately associated with technological challenges for the design and building of telescopes, space missions or accelerators, reactor and underground experiments, as well as developments in mathematics. Their resolution requires a worldwide, transdisciplinary effort and the orchestrated involvement of researchers, engineers, and technicians. The challenges described here are fully aligned with the priorities identified in international and European strategies.

KEYWORDS

Origin of mass | Symmetries and matter | QCD | Origin and fate of the Universe | Formation and evolution of galaxies | Cycle of matter in the Universe | Nature of gravity | Instruments and experimental techniques | Geometry and analysis | Matter and radiation under extreme conditions

INTRODUCTION

Understanding the basic components of the Universe, its structure and evolution is a complex of questions at the heart of humanity and has been one of its noblest and most ambitious undertakings. It can be traced back to the first natural philosophers of the antique world, but it gathered enormous impact and speed with the emergence of the scientific method in modern times. This quest has a double impact on society. On the one hand, there are the profound mental, emotional, philosophical, and potentially even political - if one thinks of the global environmental movement - changes that are prompted by our increasing understanding of our place in the Universe. On the other hand, fundamental mathematical and physical science is the basis of all of modern technology.

By its very nature, it is often impossible to predict where breakthrough discoveries will be made in fundamental science. So called “blue sky” research is one of its inherent, necessary features, combined with rigorous progress on precision and accuracy. Contrary to applied sciences, the way forward is rarely clear, with the exception of the need for permanent improvement of the experimental facilities, equipment, theoretical/numerical knowledge and skills. Time scales between fundamental discoveries and their direct technical application can be long. From past experience they can range up to ~100 years.

Understanding the Universe from the smallest to the largest scales, including its origin, evolution and fate, is a monumental challenge that requires a worldwide effort and the orchestrated involvement of theoretical and experimental researchers, together with engineers and technicians, across different fields: particle physics, cosmology, astrophysics, nuclear, molecular and materials science, as well as mathematics and microelectronics. A combination of abstract methods in theoretical physics and mathematics, technological developments and experimental scrutiny is equally needed.

Finding the fundamental laws of Nature requires us to explore both the elementary components of matter at the smallest scales and to observe the Universe, from our Solar System to galaxies across the visible Universe. Extreme objects and phenomena in the Universe can also provide key insights into environments that cannot be reproduced on Earth. Only through understanding the laws that dictated the first instances of the Universe we will be able to find the ultimate reason for its actual appearance and future fate.

María José Costa and Rainer Schödel (Topic Coordinators)

The Standard Model of particle physics is the theory that represents our best understanding of the fundamental components of matter and their interactions. It is the legacy of the twentieth century and describes (almost) all the experimental data, predicting the existence of new phenomena with highest precision. The discovery of the Higgs boson at the CERN Large Hadron Collider in 2012, fifty years after its prediction, was a seminal breakthrough and a major milestone in particle physics. However, the Standard Model cannot be considered a complete description of Nature as it does not explain some observed phenomena (such as the dominance of matter over antimatter in the Universe, the nature of dark matter or the non-zero neutrino masses) and does not include gravity. We are still confronted with many unresolved puzzles.

For example, we do not understand yet why there are three families of quarks and leptons and their mass hierarchy. Although the existence of the Higgs boson provides an answer to the question of how elementary particles acquire mass, our current understanding of the underlying principles behind this mechanism is very limited, and we cannot explain the values of the observed masses. While classical gravity is beautifully described by General Relativity, it clashes with Quantum Theory at the smallest scales, and attempts to reconcile those two fundamental, extremely successful physical theories have so far failed. The recent discovery of gravitational waves, 100 years after their prediction, opens a new window onto the physics of gravity and compact objects. The Standard Model is in principle a very predictive theory, but achieving a precise prediction of many observables, including proton and heavy ion collisions or basic properties of matter in extreme conditions, requires a deep understanding of the complex phenomena encompassed by the theory of the strong interactions.

The Standard Model of Cosmology, the so-called LambdaCDM model, can fit the observational data extremely well. Cosmological observations point to a spatially flat Universe containing 5% of baryonic matter, 26.8% of non-baryonic dark matter, and 68.3% of dark energy, which is responsible for the accelerated expansion of the Universe. However, embarrassingly, the natures of dark matter and dark energy are completely unknown. In current cosmological models inflation is the mechanism responsible for seeding the cosmic web of matter and galaxies, but so far we lack any observational evidence for it, such as the detection of a predicted stochastic background of gravitational waves. It is also extremely hard to explain the observed absence of antimatter in a Universe without a fine-tuned asymmetry in the initial conditions.

Understanding how the first luminous structures and galaxies arose from the dark ages about 300 million years after the Big Bang and subsequently evolved towards the diversity of galaxies in the actual Universe, requires to shed light on the mechanisms that regulate the connection between the dark matter halos originated from primordial fluctuations and the growth and evolution of cosmic filaments, galaxies, and galaxy (super)clusters. It is also crucial to understand the processes that form and transform nuclei, atoms, molecules and dust into stars and planets across the history of the Universe. A deep exploration of the properties of matter and radiation in the vicinity of compact objects (such as black holes or neutron stars) is essential to understand the most energetic and catastrophic events and the behavior of matter under the most extreme conditions in the Universe.

The open questions in our basic understanding of the Universe, from the smallest to the largest scales, are not independent, but deeply connected, reflecting the maturity of the achieved global understanding and the strong links between the involved fields.

Challenges in physics are intimately associated with technological challenges in the development of the appropriate tools for the design and building of telescopes, space missions or accelerators, reactor and underground experiments. These tools comprise a wide range of aspects that go beyond radiation and particle sensors and include, among others, readout electronics and data acquisition systems, intelligent data filtering in real time, mechanical structures, control of the environment, etc. Of equal importance is the development and management of large facilities.

Developments in the area of instrumentation enable tool-driven revolutions that can open the door to future discoveries. New mathematical developments in the fields of geometry and analysis as well as the application of cutting edge mathematics are also essential to face the above challenges in physics.

The efforts towards achieving a better fundamental understanding of the Universe have been divided in this document into the following ten challenges:

1. Origin of mass of elementary particles
2. Finding the underlying symmetries behind the fundamental components of matter
3. Solving Quantum Chromodynamics
4. Origin and fate of the Universe

5. Formation and evolution of large structures and galaxies
6. Understanding the cycle of matter in the Universe.
7. Understanding Gravity
8. New instrumentation and techniques for understanding the Universe, its structure and evolution.
9. New developments in geometry and mathematical analysis driven by the equations of physics.
10. Understanding matter and radiation under extreme conditions.

The CSIC institutes participating in these activities are:

- Centro de Astrobiología (CAB, CSIC-INTA)
- Centro Nacional de Microelectrónica (IMB-CNM, CSIC)
- Instituto de Astrofísica de Andalucía (IAA, CSIC)
- Instituto de Ciencias del Espacio (ICE, CSIC)
- Instituto de Nanociencia y Materiales de Aragón (INMA, CSIC-UZ)
- Instituto de Ciencia de Materiales de Barcelona (ICMAB, CSIC)
- Instituto de Ciencia de Materiales de Madrid (ICMM, CSIC)
- Instituto de Ciencias de Matemáticas (ICMAT, CSIC-UAM-UC3M-UCM)
- Instituto de Estructura de la Materia (IEM, CSIC)
- Instituto de Física de Cantabria (IFCA, CSIC-UC)
- Instituto de Física Corpuscular (IFIC, CSIC-UV)
- Instituto de Física Fundamental (IFF, CSIC)
- Instituto de Física Teórica (IFT, UAM-CSIC)

The resolution of the challenges described in this chapter can only be achieved within a worldwide collaborative effort and are also identified as main goals in the corresponding international and European strategies. The CSIC institutes involved have groups with an internationally recognized expertise. Only in this way CSIC (and Spain) can play an important (and sometimes leading) role in future ground-breaking discoveries yielding a fundamental improvement in knowledge and technology, including future applications.

By their very nature, all the challenges laid out here lean on large, international research collaborations and major international facilities, as well as technological progress. For example, CSIC researchers can be world-leading in their fields, even though 90% of their collaborators work outside CSIC and even outside of Spain. It is therefore difficult to identify focused challenges that can be approached at the sub-global, sub-national, and even institutional level. Given the complexity and longtime scales of the tasks there are, however, several overarching, general requirements that show up in all challenges:

- The necessity to have teams of critical mass and sufficient time coherence.
- Outstanding researchers with a high international profile are joining CSIC in temporary positions with some of the new excellence programs at the national and regional levels. A career path for these researchers should be a priority.
- The participation in large, international research facilities should be supported decisively. For example, the creation of a CSIC group at CERN would strengthen the CSIC position and visibility there or a point of contact at CSIC to liaise with the ESA's Science Program would also be very helpful for the different groups developing instrumentation for ESA missions.
- The long time scales of technology developments call for an urgent need of a separate technological career at CSIC, both for engineers and scientists.
- Postdoctoral scientists tend to be fundamental to progress because they can focus on science and are not burdened by administrative tasks and tend to move a lot, thus bringing fresh ideas. A CSIC specific program for both postdoctoral researchers and PhD students would be highly beneficial, which could also include a permanent assignment of necessary positions to the research groups.
- The strengthening of international relations by facilitating and funding long exchange visits in both directions.
- Making CSIC a more welcoming institution for researchers from other countries by minimizing obstacles that arise from problems in information flow, bureaucracy, or language.
- A coordinated action towards strengthening the ties of the CSIC groups involved in these challenges would be important to take advantage of the diverse and complementary expertise available at CSIC in different disciplines. The creation of transversal research groups including researchers and engineers of various centers, also external ones, could be the way to achieve this. The support of CSIC to existing national networks (such as the "Redes Temáticas") and coordinating bodies such as CPAN (Centro Nacional de Física de Partículas, Astropartículas y Nuclear) will also be instrumental.

CHAPTER 1

ABSTRACT

One fundamental question in physics is the origin of elementary particle masses. The discovery of a Higgs boson at the LHC indicates mass generation through electroweak symmetry breaking. Our knowledge about the underlying physics is still very limited; details of the Higgs sector remain a conceptual mystery. The extremely light neutrino mass scale is also unexplained. While oscillation experiments prove non-zero neutrino masses, their values remain to be measured.

KEYWORDS

Elementary particles | Mass | Higgs | Neutrino | Top quark | Vacuum

CHAPTER 1

ORIGIN OF MASS OF ELEMENTARY PARTICLES

Coordinators
Sven Heinemeyer
(IFT, CSIC – UAM, Coordinator)
Alicia Calderón
(IFCA, CSIC-UC, Adjunct Coordinator)

Researchers and Centers (in alphabetical order)
Francisco Javier Castander
(ICE, CSIC – IEEC)
Stefano Gariazzo (IFIC, CSIC – UV)
Michele Maltoni (IFT, CSIC – UAM)
Olga Mena (IFIC, CSCI – UV)
María Moreno Llácer
(IFIC, CSIC – UV)
Michel Sorel (IFIC, CSIC – UV)

1. INTRODUCTION

1.1. The origin of mass

One fundamental question in the particle physics field is the origin of mass of elementary particles. To our current knowledge the fundamental particles, as summarized in the Standard Model (SM) of particle physics, are six quarks and six leptons, see Fig. 1.1. Besides gravity, see Sec. 7, the forces between these particles are transmitted by the photon (electromagnetic force), the W and Z (weak force) and the gluon (strong force). Within the SM the electromagnetic and the weak force are unified to the electroweak force. The discovery of a Higgs boson at the LHC clearly indicates the mass generation through electroweak (EW) symmetry breaking (EWSB): the state with the lowest energy ("the vacuum") has a non-zero energy. Our current experimental knowledge of this mechanism is still very limited, and the Higgs sector remains a conceptual mystery with numerous important questions to be investigated. It does not explain the values of the observed particle masses, and it cannot explain the extremely light neutrino mass scale, as will be discussed below. These are unambiguous evidences that the SM needs to be extended.

The discovery of a Higgs boson at the LHC represents the discovery of a new and special particle: it has no spin and no charge, unlike any other particle in the SM. The Higgs mechanism provides the interaction that generates the masses of elementary particles through the spontaneous EWSB. Within the experimental and theoretical uncertainties, the Higgs measurements at the

LHC agree with the predictions of the SM. The LHC has measured its mass to be 125.10 GeV with a precision of 0.12%, and the spin and parity/CP properties are (within experimental and theoretical uncertainties) consistent with those predicted by the SM [Tanabashi et al., 2018]. Moreover, the LHC has measured the strength of its interactions to the electroweak gauge bosons and the third generation of charged fermions [Aad et al., 2020], but the Higgs self-couplings and its couplings with the lighter fermion generations are still undetermined (at best limits far from the SM predictions exist). The current accuracy of the measured Higgs couplings is moderate ($\sim 10 - 20\%$) and leaves ample room for beyond the SM (BSM) interpretations.

Numerous important questions remain to be investigated. Is it the Higgs of the SM or the first member of a new type of fundamental particles? Is it elementary or a composite object, signaling a new layer of substructure? What stabilizes the hierarchy between the Planck and the EW scale (referred to as the "hierarchy problem")? What determines the pattern of the observed particle masses? In particular, if new physics occurs at a higher mass scale, there is no explanation for why the quantum corrections to the Higgs-boson mass are much larger than the Higgs-boson mass itself. There exist indeed alternative models of electroweak symmetry breaking, satisfying all current experimental constraints, and giving answer to some of the above open questions problems, which need to be carefully scrutinized.

Moreover, the Higgs field itself has unique properties that may have allowed it to play a central role in the evolution of the early Universe which led our universe to be entirely made of matter and no anti-matter. Current (SM) predictions are inconsistent with the observed matter-antimatter asymmetry, requiring new BSM physics, see Sec. 4. The temperature at which EWSB took place in the early universe is a priori unknown, but this process may have left imprints in the gravitational wave background, see Sec. 4. Exploring the Higgs-boson interactions and searching for potential anomalies is crucial. In this line, the interaction with the top quark, the heaviest known elementary particle, plays a key role. The reasons why the top quark mass is heavier than the Higgs and electroweak bosons, and much heavier than any other fermion, are unknown; the top Yukawa coupling (i.e. the strength of the interaction between the Higgs boson and the top quark) is predicted and measured (with an uncertainty of $\sim 10\%$) to be 1, the only coupling with a natural value, which is an intriguing fact.

In a broader context, the challenge is connected to the identification of BSM physics that is known to exist. High-precision measurements of the Higgs properties will be needed, as well as precise determinations of the EW sector of the SM, which is deeply connected to the Higgs sector and the (so far undiscovered) BSM physics. Therefore, improving and completing the measurement of all elementary particle masses is also an important part of this challenge. The Higgs-boson and the top quark masses need to be more accurately determined as they are fundamental probes of EWSB. Their current values seem to indicate an unbounded Higgs potential at high scales, above 10^{10} GeV, with profound implications for the stability of our universe.

A particular role in the "origin of mass" is played by neutrinos. While the charged lepton masses are rather well known, the masses of the neutral fermions, the neutrinos, still remain to be measured. Neutrinos are now known to have non-zero masses, thanks to neutrino oscillation experiments that determine their squared mass differences, but, so far, all attempts to measure the absolute values of the neutrino masses have been unsuccessful. Dedicated direct neutrino mass measurements only provide upper limits, that demonstrate that the absolute scale of neutrino masses is much smaller than the one of the other fermions in the SM. During the next few decades strong efforts will be undertaken to measure the neutrino masses. Being the neutrinos the most abundant matter particle known in the Universe, determining their masses is particularly important to understand the large-scale structure formation in our Universe. After they decouple from the photon- baryon plasma in the early universe, the free streaming of the neutrinos leaves imprints on the galaxy clustering power spectrum as well as on the cosmic microwave background (CMB) radiation that can be measured and that depends on their total mass. While cosmological measurements can provide the strongest upper limits on the sum of neutrino masses, those limits are model-dependent. It is therefore important to pursue neutrino mass measurements also in dedicated terrestrial experiments over the next decades.

1.2. Status of the field

Status of fundamental particle mass measurements

In Fig. 1.1 we summarize the current knowledge of the elementary particle masses [Tanabashi et al., 2018] and couplings to the Higgs boson [Aad et al., 2020]. The upper plot shows (in logarithmic scale) the values of the particle masses, where for the neutrinos only upper limits exist. The lower panels show the relative experimental uncertainties of the measurements or the masses (left) and for the Higgs couplings (right).

Higgs, Top and Electroweak physics

• *Measurement of the Higgs-boson couplings to the SM particles.*
The ATLAS and CMS experiments at the LHC have performed measurements of the Higgs-boson couplings (coupling modifiers) to the gauge bosons and third generation fermions based on the κ-framework [Heinemeyer et al., 2013], requiring a modest input of theoretical assumptions (here $|\kappa_{W,Z}| \leq 1$). Current experimental precision is at the 10-20% level [Aad et al., 2020], as indicated as dashed lines in the lower right part of Fig. 1.1. In case of the muon coupling modifier, the uncertainty is at the level of 100%, while for invisible particles is of the order of 20-30% [Aad et al., 2020]. The extraction of Higgs-boson couplings requires the (SM) prediction of the Higgs-boson production cross sections and decay probabilities, which are currently available roughly at the same precision as the experimental uncertainty. Thus, the overall uncertainty includes experimental (statistical and systematic) and theoretical uncertainty. Possible deviations from the SM predictions would give clear indications for BSM physics. For the interpretation of these data w.r.t. new physics, theory predictions of the Higgs boson in BSM models is required at the same level of accuracy, which is available so far only in selected models, such as Supersymmetry (SUSY).

• *Higgs-boson self-coupling*
The direct measurement of the trilinear self-coupling λ_3, and thus of the shape of the Higgs potential, is possible via the study of Higgs-boson pair production. This is experimentally very challenging due to its small cross section. Complementary, indirect measurements provide a way to constrain the self-coupling from precision single Higgs-boson measurements. Currently, bounds of order $O(10)$ are set on λ_3, relative to its SM value [ATLAS Collaboration, 2019]. As above, theoretical predictions for the di-Higgs production and subsequent decay are required, which are currently substantially more precise than the experimental errors. No bounds on the quartic self-coupling, λ_4, exist.

• *Top-quark and Higgs-boson masses*
These are free, but fundamental parameters in the SM. The Higgs-boson mass is known at the level of ~1 permille, the top quark mass at the level of ~5 permille. Concerning the latter, theoretical issues concerning the definition of the top mass are yet to be resolved. The Higgs-boson mass within BSM models such as SUSY is a predicted, calculable quantity, where theory predictions currently are about 10 times less accurate than the experimental errors. These predictions can yield important insights into possible BSM models and restrict their parameter space.

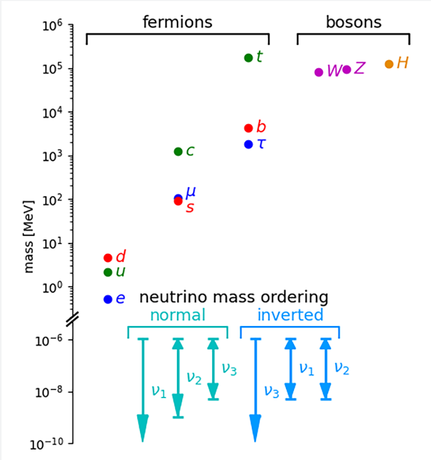
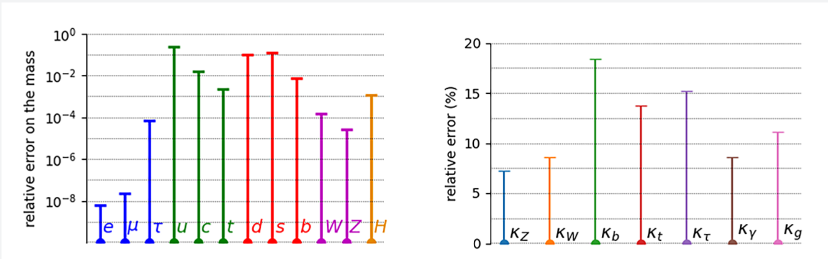


FIGURE 1.1–Status of measurements of fundamental particles: absolute mass values (top panel) and relative errors (bottom left panel). For the neutrino masses normal (inverted) ordering is shown in light green (light blue); relative experimental uncertainty of Higgs couplings (bottom right, with the theory assumption of $|\kappa_{W,Z}| \leq 1$).

Neutrinos are not included in the discussion here.



The current measurements of the Higgs and top mass indicate that the EW vacuum is only meta-stable [Tanabashi et al., 2018]. Similarly, in the determination of the lifetime of the Universe [Andreassen et al., 2018], the uncertainty is related to the top quark mass. It is therefore of paramount importance to determine these parameters with the highest possible precision, estimating reliably all sources of uncertainty, and compare experimental data with theory predictions.

- *Precision physics in the EW and top sector*

Studying precisely the top quark properties will give insight into the origin of the mass and the SM flavour puzzle. Together with EW measurements of the masses and properties of the SM gauge bosons, stringent tests of the consistency of the SM are performed [Baak et al., 2014]. Corresponding required theory predictions of these quantities are available at the same level of the current experimental accuracy, or better. These tests, while overall in agreement with the SM, indicate small tensions, in particular w.r.t. the Higgs-boson mass measurement. Such (potential) inconsistencies can give hints towards new, not directly accessible physics scales.

• *Direct searches for BSM physics*
Besides indirect searches via Higgs couplings or top/EW precision measurements, direct searches for a plethora of BSM physics has been performed at the ATLAS and CMS experiments. These searches are performed in specific BSM models, proposed to resolve some of the above-mentioned open issues of the SM, particularly the hierarchy problem. In this context, owing to the large top quark mass, many BSM models predict "heavy top partners", stimulating a lot of experimental effort. The BSM models under investigation comprise SUSY, composite Higgs models and extra dimensions, but also additional vector-like fermions and long-lived particles. So far, none of these searches could establish a BSM signal [Tanabashi et al., 2018], and limits are set roughly at the TeV scale. However, these searches are made under specific model assumptions and cannot be seen as absolute limits. Moreover, some interesting anomalies have been observed that will be further studied in the upcoming LHC runs. The interpretation of the searches requires first the formulation of a well-defined and motivated BSM model. In a second step, the production and decay properties of the BSM physics have to be calculated. Theory predictions are so far mostly restricted to tree-level calculations (with the exception of SUSY and pure QCD corrections).

Neutrinos

• *Absolute scale of neutrino masses*
In order to determine the absolute scale of neutrino masses, several types of efforts have been undertaken. Direct detection measurements, which are using the kinematics of beta decay, provide the least model-dependent results. The KATRIN experiment has recently provided its first neutrino mass result, with an upper limit of 1.1 eV on the absolute mass scale of neutrinos, which are expected to be improved by a factor of 5 in the future. If neutrinos are Majorana particles (see below), their absolute masses can be probed using neutrinoless double beta nuclear decays. Several isotopes and experimental isotopes have been used, with all searches to date yielding null results. Currently, the best upper limits on the effective neutrino Majorana mass are at the level of $O(0.1\ \text{eV})$, from ^{136}Xe and ^{76}Ge experiments. Several efforts in this direction are expected in the following years, with expected improvements

of at least one order of magnitude on the decay half-life bounds. Cosmological probes, instead, are interpreted within the framework of the standard cosmological model and try to single out the effect of hot dark matter on various observables. The most constraining cosmological upper bounds to date on neutrino masses are obtained combining CMB fluctuations with different large scale structure observations, yielding an upper limit on the sum of neutrino masses of 0.12 eV. Cosmological observations should be able to determine the contribution of the neutrino mass to the matter-energy content of the universe within a decade.

• *Mass ordering*
Neutrino oscillation experiments provide essential input to the determination of neutrino masses. Neutrino squared mass differences have been measured with an accuracy of 3% or better. Neutrino oscillation experiments can also establish the ordering among the three known neutrino mass states: they can determine whether the most electron-rich mass state (ν_1) is the lightest (normal neutrino ordering, similar to quarks) or not (inverted neutrino ordering, "opposite" to quarks). Current global neutrino fits indicate a preference for normal ordering at the 3σ level.

• *Dirac or Majorana neutrinos*
Neutrino masses also relate to theoretical challenges. It is currently unknown whether neutrinos are Dirac particles, like all the other fermions, or Majorana ones. Particles of the latter type are identical to their own antiparticles, which means only neutral fermions can be Majorana. Neutrinos are the only known particle that could in principle be of this type. If this fact is realized in nature, neutrinos would constitute a new form of matter, different from all other matter fermions, and a rare decay process which violates the conservation of lepton number would be possible: neutrinoless double beta decay. As discussed above, no evidence for this process has been found to date.

• *Origin of neutrino masses*
The smallness of the neutrino masses with respect to the ones of the other fermions and bosons in the SM raises also a theoretical problem: it is unlikely that the Higgs mechanism that gives mass to the charged fermions is also responsible for the mass of neutrinos, which is several orders of magnitudes smaller (see Fig. 1.1). A different theoretical explanation is probably yet to be discovered. The favorite theoretical framework to explain the origin of neutrino masses is the seesaw mechanism: the small but non-zero neutrino masses are here connected to a new physics scale, which would naturally lie at

energies far beyond the direct collider reach, and is always connected to the presence of additional fermion states. Such states, which may be numerous, are also referred to as “sterile neutrinos”, i.e. they do not have EW interactions, in contrast to the three “active” neutrinos that are present in the SM. The expected mass scale for the sterile neutrinos ranges from below to much heavier than the EW scale.

1.3. The overall challenge and the connection to other challenges

Understanding the origin of mass is a key point to the overall challenge: to find the theory that describes the most fundamental building blocks of nature and their interactions. This overall challenge connects the various individual challenges:

1. Origin of mass of elementary particles,
2. finding the underlying symmetries behind the fundamental components of matter,
3. solving Quantum Chromodynamics,
4. origin and fate of the Universe,
5. gravity,
6. new instrumentation and techniques for understanding the Universe, its structure and evolution.

Only a combined effort in all these challenges will make it possible to fulfill the overall challenge.

2. IMPACT ON BASIC SCIENCE AND POTENTIAL APPLICATIONS

Research towards the origin of mass has direct and important impact in fundamental and applied science, as well as in industry and medicine.

The overall aim of the challenge is to find the theory that describes the most fundamental building blocks of nature and their interactions. This answers the most basic questions of mankind. What is our world made of, and how does it “work”?

Investigating and answering these questions has an extraordinary impact on fundamental science. The knowledge of the values of the Higgs Yukawa couplings is essential to our understanding of the deep structure of matter, from the stability of the nuclei as a result of the up-and down-quark Yukawas, to the size of the atoms set by the electron Yukawa and the lifetime of the electroweak vacuum determined by the top-quark Yukawa. The Higgs self-coupling controlled the thermodynamics of the electroweak phase transition that occurred shortly after the Big Bang and that might be responsible for the matter imbalance within the universe. This imbalance might also have been generated by an asymmetry between matter and anti-matter in the very early universe (leptogenesis), see Sec. 4. Neutrinos also influence the later phases of evolution of the universe: they play a crucial role around Big Bang Nucleosynthesis (BBN), at the time of the decoupling of the CMB and during the formation of structures.

Direct applications of fundamental physics have always revealed several years later, being impossible to predict them in the present moment. However, all our current technology is based on similar basic research 50 to 100 years ago. On the other hand, there are always direct spin-offs, leading to direct applications. A prime example is that particle physicists developed an internet based system to facilitate the easy exchange of data. This system is commonly known at the "World Wide Web", one of the backbones of our society nowadays. Another important direct application of particle physics detectors and accelerators exists for the fields of medicine (hadron therapy) and radiological protection. CSIC is directly involved in a number of projects showcasing the direct applications of particle physics to society. For example, CSIC R&D in high-gradient radio-frequency cavities toward more compact electron linear accelerators for basic science are also finding applications in hadron therapy using protons or light ions (TULIP project). The technology of xenon scintillation light detection using silicon photomultipliers (SiPMs), developed at CSIC for neutrino physics research, is being adapted to develop xenon-based positron emission tomography (PET) scanners for medical imaging applications (PETALO project). A similar SiPM technology has also been exported by CSIC to the development of a detector designed to measure the concentrations of tritium in water, with applications in nuclear safety (TRITIUM project).

Carrying out the challenge requires new technology and developments at the forefront of industrial and computational research that have a great amount of applications in other fundamental and applied research areas.

- The experimental challenge requires the design and construction of high-luminosity particle accelerators, as well as ultra-performant and low-background detectors based on novel sensors, high-speed electronics and specialized mechanics, see Sec. 8. Strong teams of engineers and technicians have been trained at CSIC in these forefront technologies. This, in combination with the existence of significant infrastructures, has allowed CSIC to acquire the knowhow to tackle the development of novel detectors from conceptual design to final production. It will be necessary to go beyond the state-of-art of current technologies and new developments are foreseen. As a specific example from our challenge in the area of neutrino mass measurement, for the PTOLEMY proposal (first direct detection of relic neutrinos) new micro-calorimetric techniques and storage of graphene layers will be required, with new possible applications in electronics. There is furthermore a global effort for the development of new technologies for future particle detectors. Among them, CSIC is involved in new silicon radiation, fast-timing detectors and photo-detectors (dedicated to particle tracking and timing measurements for the HL-LHC), as well as on fiber optics detectors (for structural and environment monitoring of HEP detectors).

- The required future high-luminosity accelerators will increase the data volume by a factor of 10 with respect to the current LHC. The Square Kilometer Array (SKA) [Bull et al., 2020], under development with active participation of many CSIC institutes, will be collecting 8 Terabits per second of data. The full and efficient exploitation of such data will impose the need to implement new technologies that will allow to handle such impressive data rates and store all the relevant information, but also to apply new analysis techniques, such as new deep learning algorithms. Moreover, even more efficient computing will be required to find faster ways to analyze such incoming data and perform beyond the current state-of-the art theoretical calculations. This will have a direct impact on quantum computing (see chapter on Digitalization in this White Book).

3. KEY CHALLENGES

Understanding the origin of the mass of the fundamental particles is at the core of some of the remaining fundamental questions in particle physics and cosmology. As discussed above, understanding the origin of mass is part of the most fundamental challenge: finding the theory that describes the most fundamental building blocks of nature and their interactions. The discovery of a Higgs boson at the LHC has made this connection even more compelling and intriguing. In the following we will describe the key challenges related to the origin of mass. The strategic plan to resolve them will be given in Sec. 1.5.

1. Understand the EW symmetry breaking mechanism: our current experimental knowledge is still very limited, with many fundamental questions:

   a)Extended Higgs sectors:
   Is there only one Higgs boson as predicted by the SM, or is it one of many Higgs bosons, as predicted by (many) BSM theories?

   b) Substructure:

   Is the Higgs boson a fundamental particle or does it have an internal structure?
   Hierarchy problem:
   Why are quantum corrections to the Higgs mass much larger that its mass itself? What stabilizes the huge hierarchy between the Higgs and the Planck mass?

   c) Shape of the Higgs potential:
   The shape of the Higgs potential directly related to the (so far unknown) Higgs self-couplings.

2. Precise determination of the fundamental masses and parameters that have an important role in the SM and BSM models. This includes in particular the masses and couplings of the Higgs boson and the top quark.

3. Determine the absolute scale of neutrino masses and their ordering. This challenge can be broken into two main tasks:

   a) The unambiguous determination of the ordering (normal or inverted) of the three neutrino mass states and its explanation in terms of flavour physics models.

   b) An accurate and robust (model-independent) measurement of the mass of the lightest neutrino mass state (ν_1 or ν_3).

4. Probing the origin of neutrino mass. We identify the following tasks:

   a) Determination of whether neutrinos are Majorana or Dirac fermions.

   b) If neutrinos are Majorana particles: explanation of the neutrino mass generation in the framework of the seesaw mechanism. Determination of the new physics scale.

c) If neutrinos are Dirac particles: explanation of the neutrino mass generation in the framework of the electroweak symmetry breaking mechanism. Identification of corresponding BSM physics.

The main challenge of understanding the Higgs sector and mechanism, as well as the related BSM physics that is needed to explain the observed phenomena (baryon asymmetry, dark matter (DM), hierarchy between the Planck and EW scale, ultra-light neutrino masses) is a world wide effort.

Proposed BSM theories naturally predict the existence of new phenomena (new particles, forces or extra space-time dimensions). These theories must be tested at the high energy frontier: the LHC, its (approved) high-luminosity upgrade, HL-LHC, and at future colliders. The complementary strategy, direct search for new particles and indirect search via precision Higgs, top quark and EW physics measurements or the search for rare processes where new particles and forces can manifest, will be of utmost importance. The future measurements must be matched with theoretical predictions at the same level of accuracy, covering both SM and BSM predictions. The exploitation of the huge amount of data expected at the LHC and the HL-LHC will be crucial. Future colliders, currently being discussed as part of the European Strategy 2020 update, will bring very significant improvements.

This challenge requires the design, construction and exploitation of data from high luminosity particle accelerators, ultra-performant detectors, efficient computing, new sophisticated analysis techniques and beyond the current state-of-the art theoretical calculations and analyses. also more advanced statistical techniques/analysis.

In the neutrino sector, it is crucial to determine the neutrino mass ordering by exploiting matter effects in neutrino oscillation experiments using man-made and atmospheric neutrinos, and to determine for the first time the absolute neutrino mass scale with the precise measurement of the galaxy distribution power spectrum and the cosmic microwave background. Ideally, we would measure neutrino mass also with completely independent laboratory-based techniques. We should also establish experimentally whether massive neutrinos are Majorana fermions via the observation of neutrinoless double beta decay. In the absence of such a signal, the unambiguous determination of the Dirac nature would be more complicated but not impossible. In this case, the interplay between all neutrino mass observables (oscillations, cosmology, nuclear beta decay, neutrinoless double beta decay) will be critical.

The above key challenges can be fulfilled only in the interplay of experimental and theoretical physics.

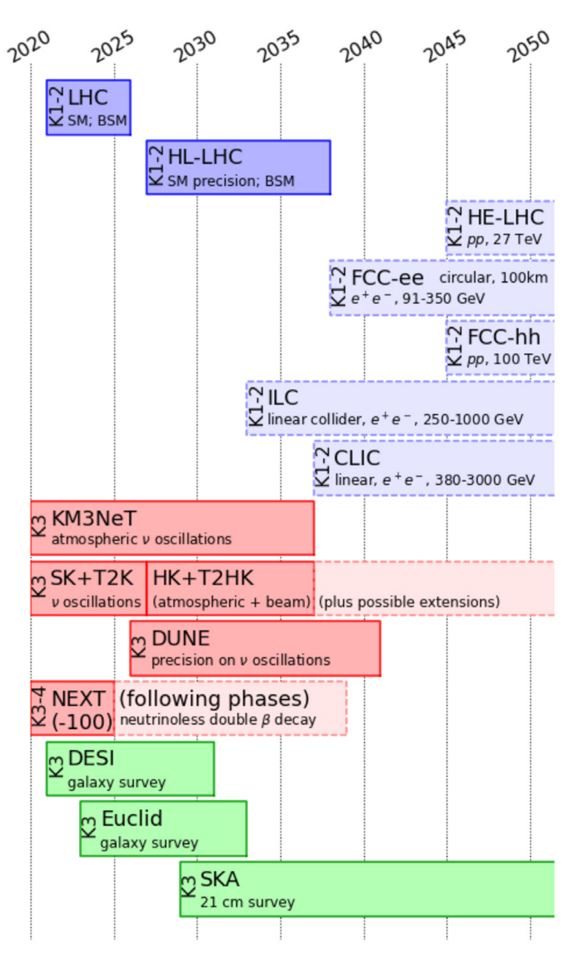
The experimental accuracy is dictated by the requirement to find the underlying theory, i.e. understanding the origin of mass. This sets the precision goal for the theoretical calculations for the (HL-) LHC operations. This also sets the precision goal on a larger time scale for future collider experiments, where substantially higher experimental precision is anticipated and consequently correspondingly more accurate predictions must be provided.

The plan to fulfill the challenge of "origin of mass" includes the participation in current and (possible) future experiments, as will be detailed below. In Fig. 1.2 we give an overview about current and planned experiments. Dark shaded boxes indicate approved experiments, while light shaded ones indicate planned (possible future) experiments. Blue, red and green refer to high-energy physics (Higgs/Top/EW) experiments, neutrino experiments and cosmological observation experiments, respectively.

Concerning the Higgs/Top/EW experiments, the HL-LHC is approved and will substantially extend our knowledge in Higgs and the physics of EW symmetry breaking, see below. Overall consensus in the community was expressed (https://indico.cern.ch/event/808335/contributions/3365090/) that an e^+e^- collider should be the next big collider experiment following the (HL-)LHC. Several options concerning geometry and location exist: the ILC (http://www.linearcollider.org/) and CLIC (https://clic.cern) are linear colliders to be constructed in Japan and at CERN, respectively, while the FCC-ee is a circular machine located at CERN, possibly followed by the FCC-hh. Similar proposals in China exist, but are not shown here.

In the neutrino sector, see the lower part of Fig. 1.2, large underground/underwater neutrino observatories studying atmospheric neutrino oscillations are either operating (Super-Kamiokande, KM3NeT) or being planned/constructed (Hyper-Kamiokande). The T2K long-baseline neutrino oscillation experiment is taking data in Japan. The approved Deep Underground Neutrino Experiment (DUNE) in the USA will also use accelerator neutrinos to study neutrino oscillations with unprecedented accuracy. Similarly, the Hyper-Kamiokande detector in Japan will also be exposed to accelerator neutrinos (T2HK). The NEXT-100 detector, whose goal is to search for the neutrinoless double beta decay of 136Xe, is currently under construction at the LSC (Spain). Larger double beta decay detectors (NEXT-HD, NEXT-BOLD) have been proposed.

FIGURE 1.2—Current and (possible) future experiments relevant for the "origin of mass". Approved (planned) experiments are shown in dark (light) shaded boxes. The timeline indicates the planned time of operation and data taking. The key challenges addressed are indicated for each experiment. The end dates correspond to the end of data taking, but not to the end of the data analysis.



In the field of cosmology, DESI has just been commissioned and will start operation soon (Arizona, USA), while ESA's Euclid space telescope is expected to be launched in 2022. Both will observe the distribution of galaxies and their clustering properties with high precision. The approved SKA project in Australia will probe earlier epochs w.r.t. galaxy surveys via the 21 cm emission from neutral hydrogen. Each of the three experiments will allow a study of the effect of neutrino masses on the cosmological evolution.

Concrete plan for the future

Concerning the key challenges of Higgs/Top/EW physics, namely challenge 1 "understanding EWSB" (search for extended Higgs sectors, for Higgs structure, solve the hierarchy problem and measure the Higgs potential) and challenge 2 "the determination of fundamental parameters", the joint concrete plan involves the following lines of action:

- exploitation of the LHC Run 3 (2021 - 2024) - exp. side and theory side, exploitation of the (approved) HL-LHC (2027 - 2038) - exp. side and theory side:
  The participation in these experiments is crucial to obtain a substantially more precise knowledge of the fundamental parameters of the model and to explore the various possibilities of BSM physics. This will give indirect and direct access, respectively, to extended Higgs sectors, possible Higgs substructures and BSM physics at the TeV scale to explain the hierarchy problem.
  The LHC Run 3, starting in 2021, is expected to provide ~300 fb^{-1} of data for analysis, more than a factor two increase with respect current luminosity. Later on, around 2027, an even higher (5-7 times) luminosity era, HL-LHC, is planned and will allow to increase the dataset up to ~3000 fb^{-1}.
  This increase in luminosity will be very valuable in the measurement of those parameters currently limited by the statistical uncertainty and will allow to explore new phase space regions and carry out differential measurements. It will also help to understand better the detector and with that improve the experimental systematic uncertainties.
  During the full HL-LHC run, we expect the production of ~170M Higgs bosons and ~120k HH pairs. Such amount of data will allow precision measurements of Higgs couplings to bosons or taus with uncertainties ≤ 2% level, being dominant the uncertainty from theory. In case of heavy quarks, like t and b, such precision will be ~ 3−4%. Similar precision is expected for the coupling to muons. We will also have access to the coupling to other SM particles, like to the quark c, reaching an upper limit < 2.5 − 5 x SM at 95% CL. In these two cases the main uncertainty comes from statistics. The expectations are summarized as red bars in all the subpanels in Fig. 1.3 and should be compared to the respective current precision as given in the lower right plot of Fig. 1.1. The trilinear Higgs self-coupling (and thus access to the shape of the Higgs potential) is expected to be determined at the HL-LHC with a precision of about 50%; Invisible and untagged Higgs-boson decays will be probed at the few-%

level. Direct searches at the HL-LHC for extended Higgs sectors and Higgs compositness scales will enter the multi-TeV region [A. Dainese et al. Report on the Physics at the HL-LHC, and Perspectives for the HE-LHC. CERN-2019-007 Yellow Report, 2019], deeply probing these ideas of EWSB and solving the hierarchy problem.

Much work is still necessary to understand better the top quark mass and reduce the sources of uncertainties. In particular, progress in Monte Carlo studies and QCD calculations for top production and decay, as well as in theoretical work concerning top mass definitions (which are at the core of investigations within the top physics community), are required. Investigations along these lines are already in progress and “alternative” methods for the determination of the top quark mass using more suitably defined observables (e.g. total production cross section or peaks/endpoints of differential distributions) are being explored. Regarding EW precision observables and top quark couplings, measurements of differential cross-sections and angular distributions are at the core of the (HL-) LHC physics program.

Direct searches for new physics, aiming to understand the hierarchy problem, including SUSY, composite Higgs models and extra dimensions, but also for additional vector-like fermions and long-lived particles, among other models, form part of the strategic lines of investigation.

From the theory side the experimental accuracies must be matched by theory predictions (for cross sections, branching ratios, masses, asymmetries, etc. within the SM and the relevant BSM models) at the same level of accuracy, and must thus equally be pursued.

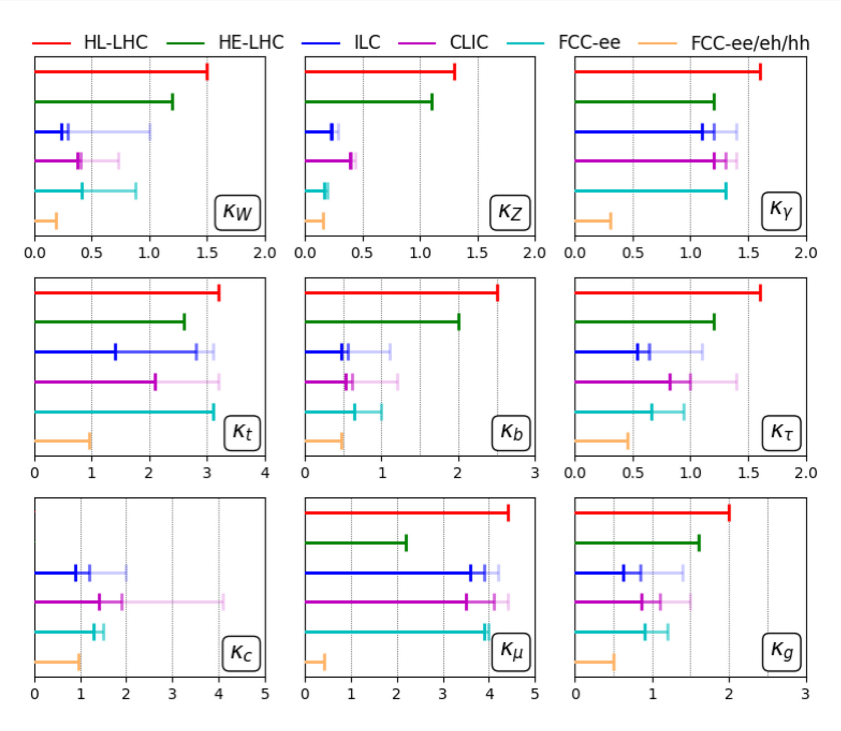
Combining the experimental results (also from astro-particle physical and DM measurements, gravitational wave experiments and from low-energy and flavour experiments, see Sec. 1.1.3) with the theory predictions will allow to perform stringent consistency tests of the SM or any other (possibly favored) BSM model. These fits will bring us closer to solve the overall challenge, the identification of the most basic constituents of our world.

- participation in the planning and execution of the next future collider experiment.

  Future particle accelerators will extend the search for new particles and interactions, as well as enable precision studies of the Higgs/Top physics and the EW sector properties.

  Forming part of the next large e^+e^- collider (from the experimental as well as from the theoretical side) at this stage comprises the active participation in phenomenological and technological studies by CSIC. The

FIGURE 1.3—Future anticipated experimental precisions of Higgs coupling (modifiers) [J. de Blas et al. Higgs Boson Studies at Future Particle Colliders. JHEP, 01:139, 2020]. All numbers refer to relative uncertainties in percent. In red are shown the results for the approved HL-LHC. Green, blue, pink, teal and yellow show the accuracies for the HE-LHC, the ILC, CLIC, FCC-ee and FCC-ee/eh/hh, respectively. Dark (light) lines indicate the final (first) stage precisions. (All "pure" pp accuracies assuming $|\kappa_{W,Z}| < 1$).



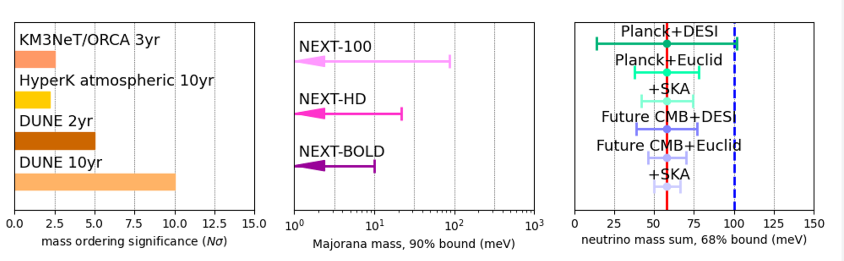
anticipated precision in the Higgs-boson couplings from the various future collider experiments are summarized in Fig. 1.3 [J. de Blas et al. Higgs Boson Studies at Future Particle Colliders. JHEP, 01:139, 2020] (see caption). These precisions are extremely high and explore mass scales in BSM models in the multi-TeV range. This covers largely our current ideas of EWSB and the hierarchy problem. Due to the very high anticipated accuracies, accompanying theory calculations will have to improve by more than one order of magnitude. This constitutes a huge theoretical challenge and requires the development of new calculational techniques.

Concerning the question which collider should be pursued, CSIC will have to follow the global endeavor, and in particular align its efforts with the European Update of the Strategy for Particle physics (to be released later this year) [B. Heinemann et al. Physics briefing book : Input for the european strategy for particle physics update 2020. CERN-ESU-004; arXiv:1910.11775, 2019]. Irrespectively of this choice, the participation in the experimental challenge will be crucial for CSIC to maintain its strong role in the investigation of the “origin of mass”.

Particularly important is the trilinear Higgs self-coupling. Future linear e^+e^- colliders will be able to measure it at the level of ~ 10%. The most optimistic expectations at future pp colliders (FCC-hh) go down to ~ 5%, requiring improvements in theoretical precisions by two orders of magnitude (relying on the development of new, so far unknown, calculational techniques). Eventually, we could get even first access to the quartic Higgs coupling. These improvements will allow us to probe the “origin of mass” at the deepest and most fundamental level.

With the strongly improved precisions in Higgs/Top/EW from future e^+e^- collider experiments, we will be able to probe new physics scales far beyond the 10 TeV range. This will have a profound impact on our knowledge on EWSB and the hierarchy problem. As before, global fits, will play a crucial role here.

FIGURE 1.4—Future anticipated experimental reach in neutrino mass ordering, effective Majorana mass and sum of the neutrino masses from cosmology.



Concerning the challenges related to neutrino physics, namely determining the absolute scale of neutrino masses and their ordering and probing the origin of neutrino mass the joint concrete plan involves the following four lines of action:

- Exploitation of neutrino oscillation experiments
  The exploitation of neutrino oscillation experiments is key to unambiguously determine the neutrino mass ordering over the next 5–10 years. The prospects through experiments with existing or possible CSIC

participation is summarized in the left panel of Fig. 1.4. CSIC participates in the KM3NeT-ORCA experiment, currently being installed. After three years of operation, the neutrino mass ordering should be determined in KM3NeT-ORCA alone with at least 2.5 σ significance. This information should be incorporated by CSIC phenomenologists in global analyses of neutrino oscillation and cosmological measurements, together with additional data available by then. In the longer term, the approved DUNE accelerator-based experiment (where CSIC also participates) will provide far better mass ordering determination, thanks to its 1300 km long baseline between neutrino production and detection locations. After 2 (10) years of operation, DUNE alone will determine the neutrino mass ordering with a significance of at least 5 (10) sigmas. This research line also involves the necessary improvements in the modelling of neutrino-nucleus interactions by nuclear theorists, as well as the development of predictive flavour physics models capable of explaining the observed mass ordering.

- Exploitation of neutrinoless double beta decay experiments
  CSIC has developed the technology of high-pressure xenon gas detectors for neutrinoless double beta decay searches in ^{136}Xe.
  The prospects of this technology are shown in the middle panel of Fig. 1.4 for various detector stages. The sensitivity is shown as a function of the effective neutrino Majorana mass: the longer the neutrinoless double beta decay half-life that can be observed, the lower the Majorana mass values that can be probed. The imminent NEXT-100 detector should reach a half-life sensitivity of 10^{26} yr after 5 years of operation, corresponding to a Majorana mass reach of 86 meV for typical nuclear physics assumptions. In the longer term, the proposed NEXT-HD and NEXT-BOLD ton-scale detectors should push the sensitivity reach down to 21 and 9 meV, respectively, for the same nuclear physics assumptions. Such ultimate sensitivities should be sufficient to guarantee the discovery of neutrinoless double beta decay in the case of Majorana neutrinos and inverted mass ordering, and to provide good discovery prospects also in the case of normal ordering. CSIC nuclear theorists should continue to improve nuclear matrix element calculations for neutrinoless double beta decay. The observation of this process would determine that neutrinos are Majorana particles. On the contrary, if sufficiently strong bounds on neutrinoless double beta decay signal can be obtained, CSIC phenomenologists should combine all neutrino mass observables to try to establish the Dirac nature of neutrinos.

Right-handed neutrinos with Majorana masses below the electroweak scale should also be directly searched for at colliders, for example through displaced dilepton vertices, and at beam dump experiments.

- Exploitation of cosmological surveys
  Cosmology surveys, tracing the early photons of the CMB and sampling the matter distribution through the observation of galaxies and neutral hydrogen throughout cosmic time, are now becoming comprehensive enough to deliver precise measurements and produce competitive neutrino mass determinations when combined together. CSIC is well placed in this quest with its strong involvement in the most relevant future galaxy (DESI, Euclid) and 21 cm wavelength (SKA) surveys.
  The right panel of Fig. 1.4 shows how well the sum of neutrino masses can be measured through the combination of various cosmological datasets: a neutrino mass sum of 60 meV (red vertical line) is assumed, roughly what one expects for a normal neutrino mass ordering in which the lightest state is massless. The combination of future CMB data with DESI, Euclid and SKA should allow us to measure the neutrino mass sum with an uncertainty of 20, 12 and 8 meV, respectively. This precision could also allow us to determine the mass ordering through cosmology, considering that the neutrino mass sum cannot be lower than about 100 meV in the inverted ordering case (blue vertical line). CSIC should also explore the robustness of these results for a number of plausible, and non-minimal, cosmological scenarios. Given the involvement in both communities, CSIC should exploit its strategic advantage in combining neutrino oscillation and cosmological data. On the theory side, CSIC should continue to develop models of cosmic structure formation based on perturbation theory.

- Development of neutrino mass models
  Theorists from CSIC should continue to develop models capable of reproducing small neutrino masses in agreement with the upcoming measurements discussed above. Models for neutrino mass generation fall naturally into two categories, depending on whether neutrinos are Majorana or Dirac particles. If neutrinos are Majorana particles, seesaw-type neutrino mass models will be developed by CSIC theorists, and the phenomenology between seesaw models and low-energy neutrino observables will be explored. Leptogenesis in the context of seesaw models will also be studied to resolve the matter-antimatter asymmetry. In the presence of a neutrinoless double beta decay signal, the impact of lepton

number-violating BSM mechanisms alternative to the standard light Majorana neutrino exchange will need to be carefully scrutinized. If neutrinos are Dirac particles, the main theoretical challenge will be the development of EWSB BSM models yielding very small Yukawa couplings (e.g. radiative neutrino masses or extra-dimension models).

CHAPTER 2

ABSTRACT

Unraveling the structure behind the components of matter and the fundamental laws governing their interactions is the ultimate goal of particle physics. One of the central questions in the Standard Model is the origin of the fermion family replication, which may point towards a hidden flavour symmetry. Many experiments will soon be addressing this issue by exploring the flavour sector with impressive sensitivities. Understanding other symmetries, such as CPT and lepton or baryon numbers, is also crucial, since their violation would have dramatic consequences for the shape of the underlying physics. Other open questions, such as the origin of the chiral character of the weak interactions, the stability of the particle that may constitute the dark matter of the Universe and the possible unification of the fundamental interactions at high energies, may also be associated with the existence of new symmetries and dynamics, possibly better described with a new mathematical language.

KEYWORDS

Symmetries | Flavour

Beyond the Standard Model | Gauge theory

CHAPTER 2

FINDING THE UNDERLYING SYMMETRIES BEHIND THE FUNDAMENTAL COMPONENTS OF MATTER

Coordinators
Alberto Casas
(IFT, CSIC – UAM, Coordinator)
Avelino Vicente
(IFIC, CSIC – UV, Adjunct Coordinator)

Researchers and Centers
Guillermo Ballesteros
(IFT, CSIC – UAM)
Alicia Calderón (IFCA, CSIC – UC)
Enrique Fernández
(IFT, CSIC – UAM)
Óscar García-Prada
(ICMAT, CSIC – UAM- UCM – UC3M)
Sven Heinemeyer (IFT, CSIC – UA,)
Carlos Mariñas (IFIC, CSIC – UA)
Salvador Marti (IFIC, CSIC – UV)
María Moreno (IFIC, CSIC – UA)
Carlos Muñoz (IFT, CSIC – UAM)
Jose Miguel No (IFT, CSIC – UAM)
Antonio Pich (IFIC, CSIC – UA)
Emma Torró (IFIC, CSIC – UA)
Bryan Zaldívar (IFT, CSIC – UAM)

1. INTRODUCTION

1.1. The flavour puzzle

The fermionic components of matter display an intriguing family structure that is not yet understood. With just one family of fermions, the up and down quarks, the electron and its corresponding neutrino, the Standard Model (SM) would be a beautiful and very economical explanation of the world. These four particles are enough to describe the rich structure of ordinary matter (atoms, nuclei and radioactivity) with the SM interactions. Besides the values of their masses and electric charges, the SM would only need five inputs to describe the world: three gauge couplings and the two parameters of the Higgs potential. However, two more fermionic families with increasing masses do exist in Nature and we are totally ignorant about the reasons of this replication.

The three fermion generations only differ by their masses. Thus, we have now a new "periodic table" of elementary particles with four columns and three rows, where the three elements of any given column have almost identical interactions. Furthermore, the four fermion masses of the one-family model get

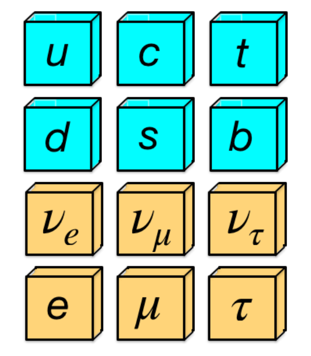


FIGURE 2.1—The three fermion generations in the Standard Model.

converted into four 3-dimensional complex matrices in flavour space, giving rise to a rich pattern of fermion masses and mixings that are empirically determined. There should be some fundamental dynamical explanation of this "flavour" structure, in the same way that the (at the time unknown) atomic structure was behind the periodic table of the XIX century. Finding out such explanation, which must be beyond the SM, and understanding how the pattern of fermion masses and mixings emerges from some underlying flavour dynamics, is a primary goal of fundamental physics.

Symmetries are usually the key to uncover hidden dynamics. In the absence of Yukawa couplings to the Higgs (and thus of masses), the SM Lagrangian has a huge $U(3)^6$ flavour symmetry (including right-handed neutrinos) that is explicitly broken by the Yukawa matrices, which point into specific directions in the flavour space. Flavour experiments indicate a quite distinctive pattern of symmetry breaking that is different in the quark and lepton sectors. Improving our knowledge of these parameters is a challenging task of great importance, since they give us precious hints on the underlying flavour dynamics.

The Standard Model predicts a strong suppression for certain *rare* processes, which would only take place at very low rates. This is the case for many flavour violating transitions, such as the $K \to \pi\nu\bar{\nu}$ and $B \to K\mu^+\mu^-$ decays, to mention a couple of representative examples. Several reasons are behind the SM suppression, including accidental symmetries and small mass ratios. However, these suppression mechanisms are "fragile", and any tiny variation of the theoretical framework may drastically alter the predictions. Similarly, there are also flavour processes which are *forbidden* in the SM but may take place in other scenarios, like $\mu \to e\gamma$ and $H \to \mu\tau$. In summary, large deviations from the SM predictions may be found in rare and forbidden processes, which makes their study one of the most powerful probes of new physics. In fact, the search for these processes will allow us to probe energy scales about a thousand times higher than those directly accessible at the CERN LHC.

Several experiments are already (or will be very soon) actively exploring the flavour sector with unprecedented sensitivities. The LHCb detector at the LHC is the most prominent example at present, but the list also includes experiments looking for the violation of individual lepton flavours, such as MEG-II and Mu3e (in Europe) or Mu2e (in USA), experiments specialized in processes involving quark flavour transitions, such as NA62 (in Europe) or KOTO (in Japan), and experiments with a broader coverage of flavour observables, like the aforementioned LHCb (in Europe) or Belle II (in Japan). The two multi-purpose LHC detectors, ATLAS and CMS, also play a complementary role in the study of flavour observables. This worldwide effort makes the question of the fermion replication in the SM more timely than ever.

1.2. Symmetries and the fundamental structure of matter

Other symmetries are behind other fundamental, and not yet understood, facts.

This is the case of the CP symmetry. Owing to the chiral structure of the SM gauge group, the fermionic couplings to the weak gauge bosons break parity (P) and charge- conjugation (C) in a maximal way, but the product of these two discrete transformations (CP) remains an exact symmetry with one or two families of fermions. Therefore, the violation of CP in hadrons seems to be directly related with the three-fold fermion replication. On the other hand, these symmetries are related to a fundamental cosmological problem. As it is well known, C and CP violation is a key ingredient to explain the observed and intriguing matter-antimatter asymmetry of the universe. Therefore the investigation of these issues directly connects and can give valuable insights to the research in other challenges, particularly “Origin and fate of the Universe” (chapter 4).

On the other hand, the apparent conservation of CP in the strong interactions is another problem in the SM. In order to solve this open question a new symmetry is often postulated, a Peccei-Quinn symmetry, naturally predicting the existence of the axion, a hypothetical particle that may constitute the dark matter (DM) of the Universe. Furthermore, due to the CPT theorem, violations of CP imply corresponding violations of time reversal (T). Understanding the realization and violation of these symmetries in nature is another fundamental challenge to be faced.

The mixings among fermion families break the conservation of the individual lepton and quark flavour quantum numbers, but the total lepton and baryon numbers remain as conserved quantities, because of accidental global symmetries of the dimension-4 SM Lagrangian. The breaking of these symmetries (even if it is tiny) has dramatic consequences. Namely, some violation of the baryon number is also a necessary ingredient to generate the above-mentioned matter-antimatter asymmetry of the Universe. Likewise, lepton number violation might be at the origin of neutrino masses (as well as at the root of the baryon asymmetry of the Universe). In fact, the unification of fundamental interactions might imply the breaking of baryon and lepton numbers at some high-energy scale. Hence, precise tests of these symmetries have the potential to be sensitive to new physics interactions at higher scales and contribute to the understanding of other challenges ("Origin of mass of elementary particles" (chapter 1) and "Origin and fate of the Universe (chapter 4)".

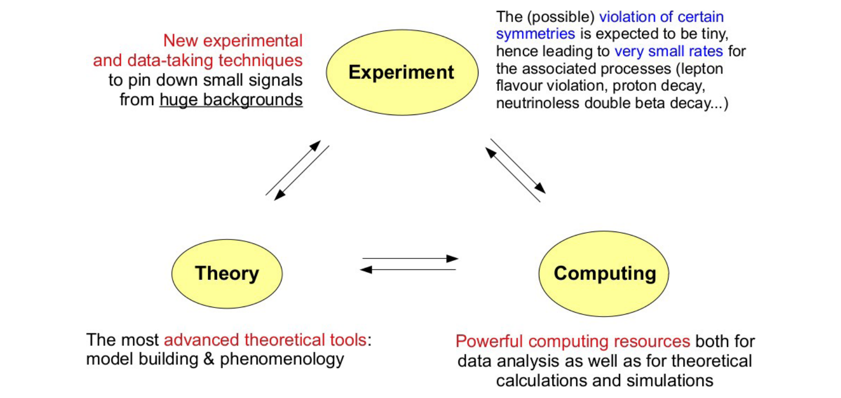
Finally, the physics associated to the mysterious dark matter of the universe, in particular the symmetry that could be behind its stability, is another crucial basic question which also connects to the just-mentioned challenges.

All these questions are going to be probed at different levels in the forthcoming years. Namely, CP violation is going to be explored in Higgs interactions at the LHC and in hadron physics in the above-mentioned experiments, e.g. LHCb and Belle II. Likewise, the peculiar flavour pattern in the neutrino sector, including potential CP violation, will be explored in DUNE. In addition, different candidates to dark matter are going to be probed in a plethora of space-based and ground-based experiments, using different techniques, e.g. XENON, CDMS, PANDA, CTA, ADMX, etc.

1.3. Other open questions

The chirality of the SM gauge group (why left is different from right?) is another open question that is lacking a convincing dynamical answer. Can the SM be embedded at higher energies in a larger symmetry group with corresponding left and right fermionic sectors? Does this group allow for unification of the different interactions? How does the larger symmetry get broken at low energies? Does it stabilize the particle that constitutes the dark matter of the Universe? Can gravity be merged with the other interactions in a consistent theoretical framework? Does supersymmetry play any role? What are the mathematics describing such super-unification?

FIGURE 2.2—This challenge requires a multidisciplinary approach, combining theoretical, experimental and computing developments.



2. IMPACT ON BASIC SCIENCE AND POTENTIAL APPLICATIONS

The issues described in the previous section are among the most prominent indications of physics beyond the Standard Model (BSM). Therefore, any progress in their theoretical or experimental status, even if it is small, will be of crucial relevance for fundamental physics. Let us mention here that, in the past, flavour-related physics has already been instrumental to show the existence of new physics before its direct observation. For instance, the study of the $K-\bar{K}$ system was the key to predict the existence of the charm quark and its mass, the measurement of ε_K (a parameter that quantifies the CP violation in kaon mixing) indicated the existence of a third generation, neutrino flavour oscillations showed that neutrinos are massive, etc.

Part of the previous achievements and the potential success of the research in this area is due to the fact that the measurement of rare (or forbidden) processes is extremely sensitive to the existence of new BSM physics. In this sense, any real progress must involve experimental results as the main ingredient, but theory is also key to refine the SM predictions and to interpret the results in a theoretically consistent way that may lead to new and relevant predictions. At present, the most promising areas where this progress may take place involve searches for flavour and CP anomalies at the LHC and other accelerators, anomalous decays of heavy mesons, forbidden lepton decays, electric dipole moments and anomalous Higgs Yukawa-couplings. These can all be found among the main goals of this challenge. In addition, other promising areas include direct and indirect detection of dark matter, production of new particles at the LHC and future colliders, experiments of ν–less double β–decay and neutrino beams, and QCD-corrections to flavour and DM processes. A signal of new physics emerging from any of these research fronts would have an importance comparable to the discovery of the Higgs boson at the LHC.

On the other hand, although this challenge mainly deals with basic science questions, it also comes with a variety of applied science consequences. Developments in computing, by the experimental collaborations or by the theoretical groups performing new physics fits, and in data analysis, both with traditional (supercomputing, parallelization) and artificial intelligence (machine learning) techniques, are expected. The experimental groups also face very specific technological challenges, due to the high precision (and intensity) they are aiming at. In fact, this challenge requires a multidisciplinary approach, combining new data-taking techniques to pin down small signals from huge backgrounds, powerful computing resources and the most advanced theoretical tools. Fig. 2.2 illustrates this multidisciplinarity and highlights the relevance of high-luminosity experiments in the search for rare processes, a promising gate to new physics due to their high sensitivity to BSM contributions.

3. KEY CHALLENGES

In short, the basic challenges in this area of fundamental research are

(i)

Identify new theories and symmetries that can improve our understanding of some very fundamental problems, namely

- The origin of flavour structure (family pattern, masses and mixing angles of elementary particles)
- Sources of CP violation (necessary to understand the matter-antimatter asymmetry)
- Nature and origin of the dark matter
- Other symmetries that might play a crucial role to solve other fundamental puzzles: Supersymmetry, Peccei-Quinn symmetry, left-right symmetry, etc.

(ii)

Derive novel strategies to explore (and hopefully discover or discriminate among) these theories/symmetries in the experimental/observational front, namely:

- LHC (including the forthcoming high-luminosity LHC, HL-LHC)
- Other accelerator experiments, such as Belle II, ...; as well as future projects, such as e^+e- colliders.
- Neutrino oscillation experiments, including the forthcoming DUNE. Also $0\nu2\beta$ experiments.
- Decay experiments.
- DM direct-detection experiments including axion detectors; also DM indirect- detection experiments.
- Study of the large scale structure of the Universe.

(iii)

Establish fruitful connections between the experimental and the theoretical fronts, such as

- Perform theoretical calculations for SM backgrounds or precise BSM predictions, in order to improve the discovery power of the experiments.
- Use novel statistical and machine-learning techniques to extract signals from background in complex experimental data, like those from the LHC or indirect-detection experiments.
- Study the performance of future experiments, like new accelerators, neutrino facilities and DM experiments.

Experiment	Location	Begins data-taking in
LHCb	Europe	Ongoing
Belle II	Japan	Ongoing
MEG-II	Europe	2020-2022
Mu3e	Europe	2022
DeeMe	Japan	2020-2022
COMET	Japan	2020-2022
Mu2e	USA	2022
NA62	Europe	Ongoing
KOTO	Japan	Ongoing
Muon g-2	USA	Ongoing
ATLAS & CMS	Europe	Ongoing

TABLE 2.1: Current and near future flavour experiments.

More precisely, some key challenges which should combine the effort and expertise of different groups and researchers are the following:

1. Flavour experiments
A most natural arena to explore the physics behind the flavour structure of the SM is the physics of mesons. For instance, the exploration of lepton flavour universality (LFU) and flavour violating decays of K and B mesons can lead to a major discovery, giving a first signal of BSM physics. This is at present a hot issue, where flavour factories, especially LHCb and Belle II, are involved.

The experimental aspect of this challenge is to strengthen the presence of CSIC in these crucial experiments by bolstering the existing experimental groups.

On the theoretical side, one of the main technical difficulties to explore flavour is the dominant role of the strong interactions in the quark and hadron dynamics. The challenge here is to use novel techniques to evaluate those contributions, which leads to a natural synergy between the experimental and theoretical sides of this enterprise. A particularly appropriate scheme to cope with this problem is Lattice QCD, which connects this challenge to the "Solving Quantum Chromodynamics" challenge (chapter 3).

Let us mention that the search for charged lepton flavour violating processes is also going to be central in the next years, with an intense experimental activity led by flavour factories and other new low-energy experiments, such as Mu3e and Mu2e, as shown in Table 2.1. In this case, the challenge on the theoretical side will be the interpretation of the experimental results, which may require the development of global fits and specific models that accommodate them.

Finally, we notice that neutrino physics is a line (also represented in challenge "Origin of mass of elementary particles", chapter 1), which is intimately connected to flavour physics. In this sense, the (CSIC) NEXT project or the CSIC participation in DUNE will be instrumental in the forthcoming years to explore fundamental symmetries in the leptonic sector, namely lepton number and leptonic CP violation. The latter is particularly relevant at the moment, after the first hints of a non-vanishing Dirac CP violating phase have been found and can have dramatic consequences to understand the matter-antimatter asymmetry of the universe, which connects to the "Origin and fate of the Universe" challenge (chapter 4).

2. Higgs, top and flavour

The physics of flavour is intimately related to that of the Higgs. In this sense, there are fundamental measurements which are still pending and will shed light on basic questions. In particular, while the Higgs Yukawa couplings to third-generation SM fermions are rather well measured, the values of the Yukawas for the first and second generation SM fermions remain very weakly constrained, yet their measurement is key to confirm that the SM Higgs gives mass to the first two generations of matter. This goal is central to challenge "Origin of mass of elementary particles" (chapter 1), but it also constitutes a fundamental input for this challenge, since it concerns the flavour structure of the matter fermions and the Higgs properties, both related to the Yukawa couplings.

The aim here is to develop novel approaches to probe the SM Higgs mass generation mechanism for light quarks and leptons, which are complementary to existing methods and can help improve the precision of the LHC in measuring these couplings. This requires identifying LHC processes which are particularly sensitive to a departure of the light quark/lepton Yukawas from their SM values. This is the case of Higgs production processes which are directly proportional to such Yukawas. A known and studied case is Higgs production in association with a charm-tagged jet; but there are other yet unexplored possibilities, like Higgs production in association with a photon.

Central to our challenge is the search for Higgs couplings to pairs of fermions of different families and/or containing non-vanishing complex phases. In particular, flavour anomalies in Higgs (and also Z) decays, such as $H \to \mu\tau$, and other flavour violating processes involving the Higgs boson, e.g. the top quark decay $t \to h\ c/u$, would represent a major discovery, with profound implications for particle physics. Similarly, finding CP violation in the Higgs sector, and its potential interplay with flavour, would constitute a major breakthrough. These experimental challenges involve the existing ATLAS and CMS groups at CSIC.

Finally, one should keep in mind that the ultimate goal of this challenge is to fully understand the origin of flavour. It is therefore necessary to combine the experimental information with theoretical developments. In particular, an important effort in model building and phenomenological study of novel setups involving flavour symmetries is required. This theoretical challenge takes advantage of the expertise of the theory groups at IFT and IFIC.

3. Higgs, top and CP

CP violation in the Higgs sector remains a possible source of the baryon asymmetry of the Universe. Measuring the amount of CP violation in the Higgs sector is one of the key tasks of the LHC and a crucial ingredient for precision studies, for example through effective field theory coefficients. Differential measurements of signed angular distributions in Higgs boson production provide a general experimental probe of the CP structure of Higgs boson interactions.

However, new LHC search strategies to probe CP violation in the Higgs sector, complementary to existing ones, would be of major relevance for this fundamental issue. A possible new avenue to search for CP violation would be to identify and target Higgs decay modes forbidden by CP.

Similar studies could be realized in connection to CP violation in top physics. This challenge would involve both experimental and theoretical research groups.

4. Search for new symmetries

There are other symmetries, such as supersymmetry, left-right symmetry or Peccei- Quinn symmetry, that have been postulated to solve some of the most intriguing aspects of the SM. An important issue in particle physics, and one of the main goals of the LHC, is to search for signals of this new physics.

For supersymmetry, the theoretical goal is the identification of viable models, provide calculations to test these models and perform those tests, in particular global fits that could lead to a cornering of the actual supersymmetric scenario (if it is really there). This connects with the experimental searches in ATLAS and CMS, where several CSIC groups are involved. These groups will be deeply involved in the HL-LHC, expected to begin in 2026. This new phase will provide a huge amount of data, hence being sensitive to rare signatures with lower number of events. In the precision frontier it is clear that a new machine will be required in the future, i.e. an e^+e^- collider. The studies for the performance of such machine require also the collaboration between theoreticians and experimentalists.

Concerning the left-right symmetry, the goal is to determine whether such symmetry does exist. This can be explored in the front of high-energy experiments (ATLAS and CMS, with the help of theoretical work to perform simulations) and in the front of flavour physics, e.g. studying the angular distribution in an eventual observation of $\mu \to e\gamma$, which will give information about the muon polarization and thus the left-right structure of the theory behind flavour physics.

In the context of dark matter research, a key experimental goal that spans over many of the challenges in this chapter, is to discover the particle(s) that constitutes the dark matter of the Universe. This will be a clear window to physics beyond the standard model, shedding light on the properties of the unknown dark sector and hopefully finding out the symmetries behind its dynamics, e.g. the possible presence of a new symmetry that protects the stability of the DM particle. This connection goes also in the opposite direction: the theoretical construction of plausible models of DM gives hints and suggests experimental techniques to detect DM. An example of this is supersymmetry with R-parity conservation, which has stable candidates for dark matter such as the neutralino or the sneutrino; whereas without R-parity the axino or the gravitino, having a lifetime longer than the age of the Universe, are also interesting candidates. Actually, in the latter case, the direct connection with gravity allows to explore a framework where all interactions of nature are unified. Other decaying candidates such as the axion (or the axino) are related to the Peccei-Quinn symmetry and the solution of the strong CP problem. Hence, our goal here is connected to other challenges, in particular "Origin and fate of the Universe", "Origin of mass of elementary particles" and "Gravity" (chapters 4, 1 and 7, respectively), and also to our key-challenge 7 below.

5. Axion physics

The existence of axion particles is a classical prediction of theories with a Peccei- Quinn symmetry able to solve the strong CP problem. In addition, these particles are very well motivated candidates for DM. In recent times, axion-like particles emerging from a more generic Peccei-Quinn symmetry with free parameters have been also intensively considered.

Study of axions and axion-like particles require the collaboration of experiment (high-energy colliders and other kinds of experiments) and theory (particle and astroparticle physics as well as cosmology). Thus, it is a highly interdisciplinary area.

One interesting challenge that illustrates these synergies is the following.

An open question in axion physics nowadays consists in quantifying precisely the axion strings contribution to dark matter (if the Peccei-Quinn symmetry is broken after inflation). Current estimates indicate that it is at least as significant as that of the misalignment mechanism (which is more standard). The difficulty in obtaining this contribution precisely stems from the fact that very large lattice simulations are required to describe the cosmological decay of the strings into axions. Much theoretical and numerical effort is being put into addressing this problem. Sheer computational power alone is at the moment insufficient, hence progress on this topic requires bold ideas and a multidisciplinary team of axion physics, lattice and cosmology experts (like the ones CSIC may reunite).

Axions and axion-like particles may also reveal themselves in low-energy flavour experiments. This can happen in two ways: by being produced in particle decays and by mediating rare or flavour violating processes. Many such flavour signatures are known and in fact an increasingly large fraction of the axion models currently under study have flavour predictions. This is another example of the remarkable interdisciplinarity of axion physics, which also plays a relevant role in flavour factories.

6. Machine-learning techniques

The application of state-of-the-art statistical methods and artificial intelligence (machine learning, ML) techniques is a promising way to improve the efficiency of analyses that involve complex sets of data. This applies in particular to LHC and DM searches.

One example of this is indirect detection, which is one of the leading techniques to explore the nature of (WIMP-like) dark matter. Similarly, ML techniques are currently being used on a regular basis by the ATLAS and CMS collaborations in order to discriminate signals from backgrounds. Due to their complexity, the analyses of data from these experiments are quite challenging from the statistical point of view.

A proposal in this sense is to work with data and simulations from the most sensitive (present and near-future) experiments of Indirect Detection and Collider searches (where several CSIC groups are involved), applying different ML paradigms as Supervised, Unsupervised and Semi-Supervised Learning. Far from replacing the physics of interest (in our case, the physics of the DM), such tools would allow us to estimate in an agnostic way the effect of the "nuisance parameters" (i.e. the background). Similar techniques can be applied to direct detection of DM.

The successful accomplishment of this challenge would imply a radical qualitative improvement of the searches for DM, which so far do not exploit in general the capabilities of modern statistical modeling. Improvements are also welcome in LHC searches, which are already using these techniques but would become more sensitive to very rare events. It will also tighten the collaborations between different CSIC institutes and groups across the country.

A similar strategy can be used, in collaboration with (CSIC) experimental LHC groups, in order to improve background estimates and discriminate BSM signals.

Last but not least, these collaborations will provide powerful tools to be used in many other branches of physics.

7. Determine the underlying mathematical structure behind the fundamental symmetries

It is a common belief that there is some fundamental reason for the presence of symmetries in nature. These symmetries include the examples of previous sections and other ones, such as gauge symmetries.

Indeed, the gauge theory of Yang-Mills and other gauge theories involving Higgs fields as well as spinors are central in the study of the basic properties of particles. From a mathematical points of view these theories lead to the study of moduli spaces, parametrizing gauge equivalence classes of solutions to the corresponding gauge equations, that are extremely rich, both from the physical and mathematical point of view. The study of the topology and geometry of the moduli spaces of instantons, monopoles, vortices, Higgs bundles and other soliton-type objects is indeed a very challenging problem of great interest in the mathematics and physics community. It is in fact an area of great interaction between these two communities.

CHAPTER 3

ABSTRACT

The aim of this challenge is to develop the conceptual and methodological tools necessary to achieve a complete understanding of the phenomena encompassed by QCD, including the ability of computing observable magnitudes sensitive to strong interactions, with essential applications in particle physics, nuclear physics, astrophysics and cosmology.

KEYWORDS

Perturbation Theory

Effective Field theory | Lattice QCD

Quantum Information Technologies

CHAPTER 3

SOLVING QUANTUM CHROMO-DYNAMICS

Coordinators

M. Pilar Hernández
(IFIC, CSIC – UV, Coordinator)

Laura Tolos
(ICE, CSIC - IEEC, Adjunct Coordinator)

Researchers and Centers (in alphabetical order)

José F. Barbón (IFT, CSIC – UAM)

Margarita García-Pérez
(IFT, CSIC – UAM)

Oscar García Prada
(ICMAT, CSIC – UAM – UCM –UC3M)

Antonio González-Arroyo
(IFT, CSIC – UAM)

Esperanza López (IFT, CSIC – UAM)

Cristina Manuel (ICE, CSIC - IEEC)

Juan Nieves (IFIC, CSIC – UV)

Carlos Pena (IFT, CSIC – UAM)

Alberto Ramos (IFIC, CSIC – UV)

Germán Rodrigo (IFIC, CSIC – UV)

Agustín Sabio Vera (IFT, CSIC – UAM)

1. INTRODUCTION

Quantum Chromodynamics (QCD) is the theory that describes the strong force responsible for confining quarks and gluons into protons and neutrons, and ultimately for the nuclear forces that bind them in nuclei. The structure of the basic constituents of matter is therefore rooted in the properties of this complex theory.

QCD is a quantum field theory that has a simple formulation in terms of colored quarks and gluons, and the gauge symmetry principle that dictates their interactions. Its solution is however beautifully complex, giving rise to a plethora of emerging phenomena, such as quark confinement. The elementary degrees of freedom are confined into composite neutral states, the hadrons, whose masses cannot be accounted for by that of its constituents, but by the energy density of the confining force. This is the so-called mass gap. Confinement and other geometrical and topological properties of gauge theories, like QCD, have made them extremely interesting mathematical constructs that have led to spectacular progress in mathematics.

Lacking an analytical solution of QCD, a multi-prong strategy has been implemented that includes a variety of tools: The perturbative treatment whenever high-momentum scales are involved, the construction of effective field theories that exploit symmetries in a maximal way, and a first-principles formulation, based on the introduction of a space- time lattice, that allows, in particular, for a numerical solution. Notwithstanding, a formal solution to the theory continues to be a key challenge for the future both in physics and mathematics.

Solving QCD could also be instrumental for progress in related challenges such as finding the extension of the Standard Model that can explain its shortcomings and open questions, or the correct theory of quantum gravity.

1.1. What can we learn by solving QCD?

Testing the Standard Model and revealing new laws of physics

The Standard Model (SM) accurately explains the dynamics of the basic constituents of matter. This theory is extremely predictive and has been subjected to numerous experimental tests, but the precise theoretical prediction of many observables that involve the strong interaction remains extremely challenging. Even basic questions such as what is the mass difference between the proton and the neutron, which underlies the stability of matter, is highly non-trivial to answer. A first-principles prediction of the most basic observables, such as hadron masses, requires a non-perturbative method.

More generally, QCD interactions are ubiquitous in High-Energy Physics (HEP) experiments. The Large Hadron Collider (LHC), the most powerful particle accelerator, collides strongly-interacting protons up to 7 TeV energies, providing an exquisite view of fundamental particle interactions at the shortest distances or highest energies. It has so far confirmed the validity of the SM of particle physics up to unprecedented energy scales, and with great precision in the sectors of strong and electroweak interactions, the Higgs boson, and flavour physics including top quark properties. With a ten-fold larger data set, the HL-LHC will extend the sensitivity to new physics in direct and indirect searches for processes with low production cross sections and harder signatures, and notably to the still unknown self-couplings of the Higgs boson. The properties of heavy hadrons, such as B and D mesons, will be measured with unprecedented accuracy in the LHCb experiment, as well as in the Belle-II B-meson factory complex in Japan. These measurements will provide some of the most stringent tests of the SM. An ultimate test of precision is also provided by the measurement of the magnetic moment of the muon, $g - 2$, where tiny quantum QCD effects are key for an adequate SM prediction, or that of the electric dipole moments of subatomic particles.

Future colliders can potentially further unlock deeper answers about the nature of the fundamental interactions of matter. In addition, a considerable improvement is expected in precise measurements that will reach the percent level. One common thread underlying present and future efforts is that new physics might first appear as a subtle deviation from the SM predictions. Their potential therefore hinges on the limited precision with which we presently understand the involved QCD effects: Significant improvements in the theoretical understanding of the quantum structure of QCD at high energies are a prerequisite for the success of the experimental programs.

Extreme QCD in astrophysics and the Early Universe

QCD interactions are also essential in the understanding of matter in extreme conditions of high temperature and density. Heavy ion collisions, such as those produced at RHIC or in the ALICE experiment at the LHC, recreate in the laboratory the conditions shortly after the Big Bang, where the Universe was filled with a hot plasma of quarks and gluons. On the other hand, the understanding of astrophysical compact objects such as neutron stars involve matter at extremely high densities, where new exotic phases of matter, such as color superconducting phases, can appear. Solving QCD will bring a quantitative understanding of the QCD phase diagram depicted in Fig. 3.1, by determining the equation of state and transport properties of extreme matter. This will be a major achievement in our understanding of the Early Universe, and the dynamics and evolution of astrophysical compact objects, that are starting to be studied with unprecedented precision via multi-messenger astrophysics.

Strongly-coupled Beyond the Standard Model Theories

Solving QCD in the perturbative and non-perturbative regimes could have an impact also on model building beyond the Standard Model (BSM) of particle physics. Indeed, many BSM theories that address open questions, such as the hierarchy problem, the nature of dark matter or the origin of flavour, involve new interactions that lead to QCD-like phenomenon such as confinement and/or chiral symmetry breaking. Indeed, the existence of a mass gap in QCD is the most beautiful example of an emergent *natural* scale. If the Higgs mass were dynamically generated, as the QCD mass gap, or in other words if the Higgs were a composite particle made of more elementary constituents,

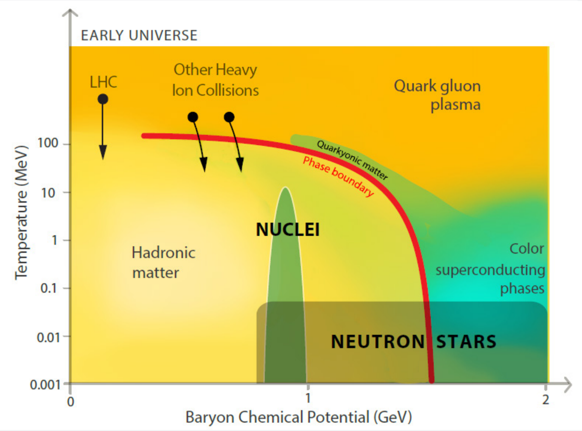


FIGURE 3.1—A sketch of the expected QCD phase diagram. Source: Watts et al., Rev. Mod. Phys. 88 (2016) 021001

similarly to the known hadrons, this would solve the hierarchy problem of the SM. The same methods that help us solve QCD could also be essential to address new physics challenges.

Gauge theories and gravity

The intriguing connection of gauge theories and gravity via the holography principle might result in further applications of QCD, and other strongly coupled gauge theories, to gravity and string theory and vice versa.

The idea that some sort of string theory should emerge in the limit of a large gauge group (i.e. large number of colors) was already present in the original work by Hooft, as well as many phenomenological hints in the physics of meson resonances, going back to the pre-QCD days. The celebrated AdS/CFT correspondence of Maldacena is a particular proposal for such a string solution of the large color number (*Nc*) limit, for the particular case of the maximally supersymmetric cousin of QCD. The answer turned out to involve the same string theories which were developed as models of quantum gravity and unification. The correspondence is 'holographic' in the sense that the non-gravitational QCD-like description uses degrees of freedom on a boundary, while the gravitational, string-like description lives in the bulk of a spacetime with negative vacuum energy.

An enormous theoretical effort has been devoted to the exploration of these ideas in recent times. In one direction, one uses classical gravitational methods as a computational tool in strongly coupled QFTs. In another direction, the holographic aspect holds deep connections with the theory of quantum black holes, such as the mystery of what microscopic degrees of freedom are being counted by the Bekenstein-Hawking entropy formula.

These results have demonstrated that the richness of quantum field theories in general, and QCD in particular, goes much beyond what was originally envisioned. In some sense, all these findings are examples of hidden duality symmetries of QCD-like theories, and the known examples are likely to be just the tip of an iceberg. This expectation is backed by the continued discovery of novel mathematical structures, such as the recently found connections between positive projective geometry and scattering amplitudes in gauge theories, very interesting results pointing in the direction of generalizations of the notion of holography for the S-matrix in flat spacetime.

Applications in Mathematics

Many of the recent connections between physics and mathematics are related to the geometrical and topological properties of Yang-Mills theory. With its mathematical roots in the theory of connections on principal bundles and the notion of curvature, Yang-Mills theory irrupted into mathematics in the 1970s. One of the most interesting applications of the Atiyah–Singer index theorem was in the computation by Atiyah-Hitchin-Singer in 1978 of the dimension of the moduli space of instantons — self-dual solutions to the Yang-Mills equations. The quest for a constructive method for instantons of arbitrary topological charge was participated by both physicists and mathematicians, and ended with the Atiyah-Drinfeld-Hitchin-Manin construction. This was followed by the spectacular work of Simon Donaldson to prove the existence of exotic differentiable structures in R^4 using the moduli space of instantons. This won him the Fields medal in 1986. Moduli spaces of instantons and their dimensional reductions like monopoles, vortices and Higgs bundles have been of great importance in lower dimensional geometry and topology, linking with fundamental classification problems in algebraic geometry. Even on Riemann surfaces these gauge theories, involving both, gauge fields and Higgs fields, are central in the study of Mirror symmetry and Langlands duality —the mathematical theory generalizing electric-magnetic duality.

How can we solve QCD?

Perturbation Theory

In the high-energy and ultra-relativistic regime, the perturbative treatment of QCD can accurately describe the complex hard scattering processes taking place at high-energy colliders, such as the LHC, and the future colliders taken at consideration in the recent update of the European Strategy for Particle Physics. Achieving extreme precision from theory is of paramount importance for the full exploitation of these experiments, and requires the accomplishment of higher quantum orders in the perturbative expansion. Theoretical predictions at third order (next-to-next-to-leading order), and even at one order higher for e.g. the production cross-section of the Higgs boson, are now becoming standard, but will not be sufficient to describe the data of the HL-LHC. Furthermore, the resummation of enhanced contributions to all orders turns out to be essential to achieve reliable predictions in certain kinematical limits. The development of new methods beyond the current state of the art is compulsory to face the enormous challenge posed by the demands of present and future experiments at high-energy colliders, which also requires prominent computer resources and necessarily a deeper understanding of the mathematical proper- ties of multi-loop scattering amplitudes.

Exploiting symmetries: Effective Field Theories

Besides numerical methods, analytical tools based on dispersion relations and/or effective field theory techniques that exploit symmetries in a maximal way (ChPT, HQET, LEET, NRQCD, SCET, OSEFT, etc.) provide complementary approaches to face the strong-coupling regime of QCD interactions. These methods require experimental input or lattice input to ensure model-independent predictions. For very complex hadronic systems such as exotic heavy hadrons, nuclear physics, and hadronization processes they remain nowadays indispensable. Effective field theories have also been successfully adapted and used to study QCD at finite baryon density.

Lattice QCD and High-performance computing

In most cases, however, we need to deal with QCD interactions at energies that are not much higher than the mass gap. In this situation, perturbation theory breaks down and a non-perturbative approach is mandatory. The most promising method to solve QCD in this regime is the discretization of QCD on a space-time lattice, which makes the theory solvable by numerical methods. Significant progress has been achieved in recent years on the determination of the light hadron spectrum, the QCD coupling or the determination of light quark masses [Aoki et al., 2020]. Further efforts are needed to reach the same level of precision in other observables of phenomenological interest (e.g. the anomalous magnetic moment of the muon, structure functions, hadron decays, etc), that are of greater complexity and/or involve heavy quarks.

It is important to stress that computing infrastructures are an essential tool in lattice QCD (LQCD), and the advent of exascale systems in the EuroHPC initiative[1] will allow to solve many of these problems, as long as there is an effort in code and algorithm adaptation dedicated to exploit them maximally.

On the other hand, the study of systems with high-baryon density or out-of-equilibrium dynamics are affected by the so-called sign-problem, which makes them intractable with standard Lattice Monte Carlo techniques. The LQCD community is intensively looking for alternative tools to deal with these problems, among them complex Langevin or integration over Lefschetz thimbles.

Quantum Information Technologies

There is nowadays growing interest towards the study of QCD-like systems with Quantum Information (QI) technologies –see for example [Bañuls et al., 2019] and references therein. The purpose is, in this case, twofold: on the one hand, gauge theories are a very useful test bed for quantum simulators and quantum computation; on the other, quantum technologies can provide a novel tool to address the sign-problem. Although a real impact for QCD can only be expected on the long term, this cross-fertilization between the LQCD and the Quantum information communities opens up a window of opportunity that is al- ready producing very interesting outputs and can be very fruitful for the evolution of the two disciplines.

Formal Methods

Solving QCD analytically has been a long-standing goal in fundamental physics and mathematics. The Clay Mathematics Institute in 2000 established a list of mathematical problems that includes some of the most challenging mathematical unsolved problems. In this Millennium Problems list (as they call it), there is the Yang–Mills and mass gap problem, that is to give a mathematical proof of the existence of a Yang–Mills theory in four dimensions with a mass gap[2]. Progress in establishing this will require the introduction of fundamental new ideas both in physics and in mathematics

1 Accessible at: https://eurohpc-ju.europa.eu

2 Accessible at: http://www.claymath.org/millennium-problems/yang-mills-and-mass-gap

2. IMPACT ON BASIC SCIENCE AND POTENTIAL APPLICATIONS

As explained above one of the major motivations to solve QCD is to provide accurate predictions of the observables that will be studied in present and future particle colliders. Similarly, understanding the evolution of the Early Universe and its recreation in heavy ion collisions or the dynamics of neutron stars, requires that we understand the properties of matter in extreme conditions of high temperature and baryon density. The solution of QCD could also provide some clues as to what may lie beyond the SM or what is the theory of quantum gravity. Finally, the progress towards solving QCD could result in cross-pollination in inter-connected fields ranging from numerical methods in statistical physics, quantum information technologies and mathematics.

Particle physics at the HL-LHC

After Run 3 of the LHC ends in 2025, the world largest collider will enter a two-year long shutdown. Starting in 2027 the HL-LHC will provide a better potential to see rare processes and a reduced uncertainty in many measurements. Being primarily a proton- beam collider, QCD uncertainties enter in the evaluation of basically all LHC results. Given that the center-of-mass energies reached by LHC will remain unparalleled for a long time, many of our hopes to better understand the nature of the fundamental laws of Nature rely on being able to make the most of this experimental facility.

In order to exploit the full potential of the HL-LHC physics program [Dainese et al., 2019], the high precision of experimental data must be confronted with theoretical predictions of comparable accuracy. Theory is already the main limitation factor in many current LHC analyses, and the gap between theory and experiment will become soon even more evident unless further progress in the understanding of the quantum properties of QCD at high energies is achieved. Prospects for future colliders generally assume that improving the theory uncertainties by a factor 2 to 4 will be necessary.

Flavour physics

The SM of particle physics provides a simple and accurate description of fundamental forces in terms of gauge interactions. The flavour structure and the Higgs sector, on the other hand, remain cumbersome and mysterious, with over twenty unconstrained parameters and poorly-understood large hierarchies. It is tempting to interpret this as a suggestion that flavour physics might be the portal to a more fundamental organizing principle, that underlies explanations to experimental and observational findings outside the scope of the SM (dark matter, neutrino masses, matter/antimatter asymmetry, etc.) or to its fine tunings, such as the electroweak (EW) hierarchy problem, the strong CP problem, or the favor puzzle itself.

In the next decade, the existing tensions in various B-meson decay modes in the LHCb and Belle-II experiments will be extensively explored. This is then expected to provide a platform for even more ambitious, future enterprises, e.g., reaching integrated luminosities of 250 ab^{-1} with a SuperKEK upgrade. Any discovery of new physics will result from a comparison between high-precision experimental results and theoretical predictions; thus, experimental efforts must be matched by theory. LQCD determinations of the strong-interaction contributions to flavour physics are a mandatory ingredient of this strategy.

Proton structure

Revealing the internal structure of the proton and how its mass and spin emerge from the interactions of the fundamental quarks and gluons is a long-sought dream. Future experimental facilities, like the eRHIC project at RHIC, plan to probe the internal structure of the proton via high-energy electrons. This experimental effort should come hand in hand with a theoretical effort in order to gain a better understanding. The potential impact of such combined effort is momentous: the origin of the proton spin crisis, solving the proton radius puzzle, clarifying the role of quarks and gluons in the atomic nucleus structure, or discovering new forms of matter like the color-glass condensate, are just a few of the many targets.

Equation-of-state and transport properties of matter under extreme conditions

The study of matter at extreme conditions, as those found in heavy ion collisions or neutron stars, can be done using effective field theories (combined with unitarization techniques in certain cases). Recent success has been achieved in the dense regime of neutron stars by addressing the equation of state in view of the present (NICER) and future (eXTP) X-ray astronomy missions as well as the observations of gravitational wave signals of two-neutron star mergers [Watts et al., 2016]. The future goal is to extend the analysis to the dynamical properties of matter in the dense interior of neutron stars, by determining the transport coefficients and their role in astrophysical processes, such as neutron star mergers. Also, the exploration of the high-temperature regime in highly energetic heavy ion collisions, such as at RHIC/Brookhaven or LHC/CERN, or the hot dense medium that will be created in the future CBM/FAIR experiment will be of major concern for effective field theories that include heavy degrees of freedom, such as charm or beauty.

Our challenge is to derive the observables in heavy ion collisions from QCD. The results of this research have an impact on particle and nuclear physics experiments at the RHIC, LHC, B-factories, neutrino and dark matter experiments, and experimental nuclear physics, as well as astrophysical observations of compact stars.

The methodology involved in solving QCD might have an impact on the developments of new algorithms with potential applications in many other fields.

Development of state-of-the-art algorithms and novel numerical methods

The LQCD community has made leading contributions to High-Performance Computing in terms of algorithms and code performance. They have been involved in the design of new hardware (e.g. IBM Blue Gene systems); they have led the development of widely-used algorithms (e.g. Hybrid Monte Carlo), and they have exploited new technologies in a timely manner (e.g GPUs for scientific computing). LQCD has been identified as a PRACE (Partnership for Advanced Computing in Europe) scientific challenge[3] for the upcoming decade.

New numerical algorithms will be needed to address the sign problem, with potential applications in numerous other branches of physics.

QCD a test bed for Quantum Information technologies

In the long term, the impact of quantum simulators on any computational problem in physics is in principle enormous. However, there is a long road ahead and QCD is probably one of the most difficult problems that can be addressed in this venture. The challenge at this point is more conceptual than practical. We need to develop the required concepts and tools that could lead to the treatment of complex strongly coupled systems such as QCD. In this respect, we are still on a primitive phase in which only simple systems of low dimensionality or reduced number of degrees of freedom can be tackled. Still, even at this level, synergies between LQCD and QI can have a strong impact. In particular, Quantum Field Theory studies have become a benchmark for the flourishing field of quantum computers and simulators, triggering a large amount of research activity in this field [Bañuls et al., 2019].

3 Accessible at: https://prace-ri.eu/about/scientific-case/

3. KEY CHALLENGES

In the following we describe the key challenges in which CSIC plans to contribute among the previously identified high-impact open problems.

Providing accurate predictions of flavour observables in the SM and beyond

The most obvious and yet complex prediction of QCD is the hadron spectrum. Hadron masses and decay properties are essential inputs to extract the fundamental parameters of the SM, such as quark masses and mixings, including the CP phases responsible for the asymmetry between matter and antimatter. Knowing the value of these parameters precisely enough is a prerequisite for solid, precise SM predictions that can be pitted against experimental results. This is then the basis to develop searches for new physics phenomena. In the past two decades, LQCD has made a tremendous progress to achieve the most precise determination of quark masses from the prediction of hadron masses.

The study of hadron decays, mediated by the weak interactions, is much more difficult. The strategy to predict these processes exploits the fact that the typical energy scales of the strong and weak interactions are vastly different — roughly in the range of hundreds of MeV and 100 GeV, respectively. It is then possible to use effective field theory techniques to express the relevant quantum transition amplitudes as a linear combination of so-called weak matrix elements, that encode the effective low-energy interactions of quarks and leptons subject to the color and electromagnetic dynamics only. A similar method can be used to study the effects of putative new physics (BSM), that could be contributing, besides the weak interactions, to these transitions.

The determination of these so-called weak matrix elements has long been a mainstay of LQCD studies. The results are nowadays regularly summarized in the Flavour Lattice Averaging Group (FLAG) report [Aoki et al., 2020], an initiative where IFT and IFIC researchers are playing a leading role.

The current scenario for weak matrix element computations bears witness to sub- percent and sub-to-few-percent uncertainties in the determination of the hadronic contribution to tree-level SM leptonic and semileptonic decay amplitudes of flavoured pseudoscalar mesons, respectively (see [Aoki et al., 2020] and references therein). A major challenge to reach the next level of precision will be to include electromagnetic effects in these processes. Much progress has been made recently to put QED corrections on solid grounds.

One of the most promising avenues has been proposed by a group that includes several CSIC re- searchers, but still more work is required for several key observables. These computations are also numerically very demanding, and they will probably require new algorithmic developments.

Other more complex hadronic decay modes, which are either loop-suppressed, involve vector particles, baryons, or multihadron final states are still very challenging for LQCD. For example, the very well-known example of non-leptonic kaon decays, which originally demonstrated the violation of the CP symmetry, remain still poorly predicted by LQCD.

The physics goals for the next generation of experimental efforts in flavour physics necessitate of several improvements and extensions of the current LQCD state-of-the- art. The main focus of future research in the short- and mid-term will be the physics of heavy quarks, especially in the quark bottom sector. The next decade will see much effort devoted to explore the existing tensions in various B-meson decay modes in the

LHCb and Belle-II experiments. First-principles computations of some of the involved amplitudes, such as the much-studied $B \to K^{(*)} l^+ l^-$ modes, are still missing, and should be subject to intensive LQCD studies. Effort will also be devoted to improving our control

of neutral $B(s)$-meson oscillation amplitudes, required for Cabibbo-Kobayashi-Maskawa (CKM) consistency checks. Increasing accuracy in core flavour observables will also require bringing several tree-level SM amplitudes to the sub-percent precision, which requires to include long-distance electromagnetic contributions. A well-defined goal is to reach the few-permille precision level for third-generation CKM matrix elements [Cerri et al., 2019]. Furthermore, full exploitation of experimental information, especially from baryon and multihadron modes provided by LHCb and its planned upgrades, will furthermore require a decisive effort to generalize the control over the relevant hadronic matrix elements[4].

The use of effective field theories in hadron physics has been extremely useful in the past and will continue to be so in the future. Exploiting the power of symmetries in correlating observables, effective field theories provide a very efficient parametrization of the relevant physics in terms of effective couplings that can be extracted from experiment or ideally determined from LQCD.

4 The reader is referred to Chapter 5 of [Catani et al., 2012] for a more detailed discussion of these topics.

Symmetries, such as chiral symmetry for the light-quark sector, or the spin-flavour symmetry for the heavy-quark ones are the relevant symmetries on which Chiral Perturbation Theory or Heavy quark effective theory rely. Thanks to them, one can systematically study hadronic systems, classify and describe recently discovered exotic states and predict the existence of new ones [Guo et al., 2018]. Effective field theories can also be extended to the study of hadronic matter in conditions of finite density and temperature [Tolos and Fabbietti, 2020].

Providing accurate predictions of HL-LHC observables in the SM and beyond

Due to confinement, quarks and gluons are bound into protons or other hadrons in the initial state of hadronic collisions. In the final state, they materialize as collimated jets of hadrons and leptons that inherit the emission directions of the underlying hard scattering process. Experiments at high-energy colliders are therefore sensitive to quark and gluon interactions at very short-distances, where QCD emerges as the most relativistic non-Abelian gauge theory that can be studied in the laboratory. In this regard, the factorization properties of QCD play a fundamental role in the theory interpretation of hadronic collisions at high energies,

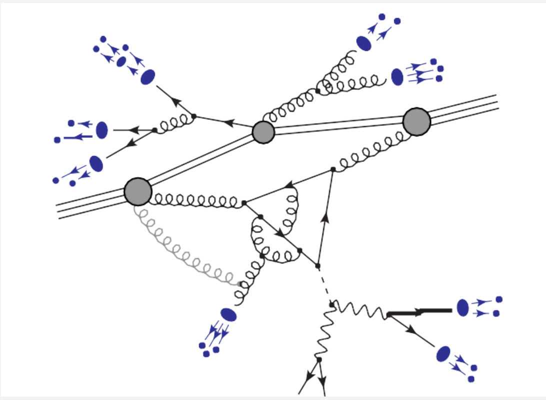


FIGURE 3.2—Theory interpretation of proton-proton collisions at the LHC

since they allow us to disentangle the description of the short-distance from the long-distance physics, as it is schematically shown in Fig. 3.2.

The quantum effects at the shortest distance scales are encoded in a systematic expansion in the strong coupling by scattering amplitudes [Dixon, 2011]. Scattering amplitudes are also objects with very interesting and striking mathematical properties. Scattering amplitudes at the lowest order in the strong coupling, or leading order (LO), which are represented by tree-type Feynman diagrams, provide however very rough theoretical predictions.

The first non-trivial contribution to theoretical predictions at colliders is next-to leading order (NLO), which involves Feyman diagrams with closed circuits, also called loops. Around ten to fifteen years ago, a number of breakthrough ideas led to efficient algorithms for the calculation of one-loop scattering amplitudes, and a number of tools to compute NLO cross sections for generic LHC processes in an automated way were developed. Current years have seen an incredibly fast progress at the next-to-next-to leading order (NNLO). Cross sections for processes with two hard partons in the final state, and more recently three, are known today thanks to the development of new methods to compute two-loop amplitudes. Beyond NNLO, results for Higgs boson production through gluon fusion have been published at N^3LO. The complexity of these calculations suggests that with the existing state-of-the-art theoretical methodologies it will be very hard to extend them to more intricate processes and higher orders as needed to match the target level of accuracy for the upcoming collider datasets.

Besides fixed-order predictions that provide the best theoretical description at high transverse momentum, also resumed calculations to all orders of contributions that become large in certain kinematical configurations have seen a leap in recent years, as well as their matching and merging with fixed-order predictions so as to have an improved accuracy in all regions of phase space. These enhanced contributions manifest as logarithms of the ratio of very different scales and are determined from asymptotic approximations to scattering amplitudes. Parton showers in Monte Carlo event generators are a key example of numerical resummation of large logarithms at small transverse momentum. In the analytic approach, effective field theories are the most powerful tool. The so-called soft collinear effective theory (SCET) isolates systematically the leading contributions to the amplitudes related to the emission or absorption of soft and collinear particles, which is useful for the detailed study of the substructure of collimated jets.

CSIC researchers have pioneered accurate theoretical predictions that have been instrumental for the success of several experimental analyses at high-energy colliders, particularly in the multijet, top quark and Higgs boson sectors, and have consolidated strong collaboration links with CSIC experimental groups and worldwide. The continuation of this theory-experiment collaboration will be even more essential in the future.

The main cause of the bottleneck for reaching better theoretical predictions at short- distances is intrinsic to the mathematical framework of quantum field theory. Scattering amplitudes in massless non-Abelian gauge theories like QCD always exhibit singularities due to the exchange of soft gluons and collinear splittings. These unphysical infinities, as well as those generated in the ultra-high-energy limit, also called ultraviolet, are conventionally regulated by modifying the dimensions of the space-time. This unphysical extension beyond the customary four space-time dimensions provides mathematical rigor to intermediate analytic expressions but increases the difficulties to evaluate scattering amplitudes. Researchers from IFIC-Valencia and Granada are the leading proposers of promising alternative methods aimed at overcoming these limitations by constructing representations of scattering amplitudes right in the four physical space-time dimensions.

The factorization properties of QCD are generally assumed in theoretical predictions, but there is still no complete proof of its universal validity. IFIC researchers, in collaboration with INFN and U. Buenos Aires researchers, have shown evidence that collinear factorization breaks down at higher quantum orders [Catani et al., 2012] due to the exchange of Glauber gluons. This breaking of factorization could lead to sizable and measurable effects at the HL-LHC. Processes at low transverse momentum also challenge the validity of factorization theorems. In this regime, an interesting configuration where factorization could be restored is the limit where particles are produced at large rapidity distances. This limit coincides with the kinematical region where Regge theory [Fadin et al., 1995] that introduces new multigluon states such as the Pomeron and the Oderon can be applied, and constitutes an active research line of IFT researchers which is also closely related to formal aspects of the supersymmetric versions of QCD. All these effects deserve further investigations and will be probed at the upcoming LHC's data taking phases.

Theoretical predictions at colliders depend on the fundamental parameters of QCD: the strong coupling, αs, and the heavy quark masses. Currently αs is known with a precision ~ 1% and is the least precisely known among the three gauge couplings of the SM. This uncertainty propagates to several key quantities. The value of the top quark and bottom quark masses are important inputs for the Higgs physics program at the LHC. LQCD can contribute decisively to improve the accuracy on some of these parameters. The state-of-the-art lattice determination of α_s [Bruno et al., 2017] has reached a better precision than other phenomenological determinations, see Fig. 3.3. This important result has been

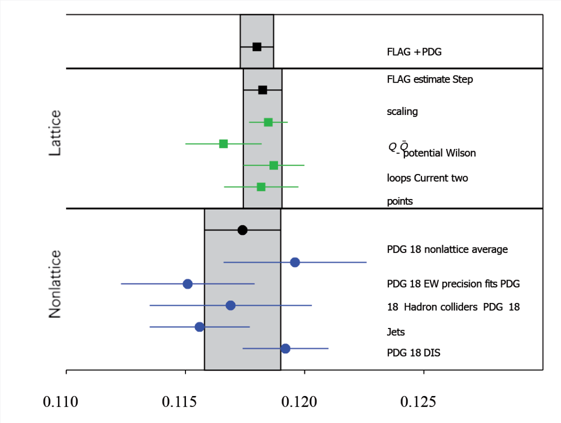


FIGURE 3.3—Comparison of different determinations of αs. Taken from [Aoki et al., 2020].

obtained by group involving CSIC researchers. The lattice has also provided the most precise calculation of the bottom quark mass. This is a mature, yet very active, area of research, facing the challenge of meeting the required precision for the HL-LHC, that will require a combination of new techniques, better algorithms and a substantial amount of computing power.

The genuine long-distance contribution to LHC scattering processes depends on how the fundamental quarks and gluons are tied together to form the proton. This effect is encoded in the so-called *parton distribution functions* or PDF's. They are universal and do not depend on the particular process under study. Traditionally, PDF's are directly determined phenomenologically from experimental data, and their perturbative evolution is currently known at third order. IFIC researchers have proven that the perturbative evolution of PDF's is able to generate a difference in the density distribution of quarks with respect to their antiparticles even in the absence of a non-perturbative asymmetry. Also, they have introduced PDF's for leptons and photons for the first time. Their effect, al- though tiny, will be probed in detail at the HL-LHC.

In recent years, LQCD has developed a first-principles theoretical computation of these quantities, with the potential to reach kinematic regions that are poorly sampled by the experimental data. This is an active area of research that is still in its infancy. There are still conceptual issues to be solved and even so it will be challenging to make LQCD extractions of PDF's competitive in precision with global phenomenological fits which, as of today, reach the few-percent level.

Understanding the proton structure

In the late 80's the results of the European Muon Collider (EMC) showed that only a small part of the proton spin comes from the constituent quarks. This result came as a surprise, and constitutes a flagrant contradiction with the constituent quark model. Further experimental results have confirmed that in fact only about 30% of the proton spin comes from its valence quarks. Understanding the origin of the proton spin is, as of today, still a theoretical puzzle. Moreover, the future eRHIC project will allow us to look inside the proton with an impressive precision. Not only better measurements of the proton spin, but also other questions, like the proton radius puzzle or the role of quarks and gluons in the structure of the atomic nucleus will be explored.

LQCD offers a unique opportunity to shed some light into these fundamental questions. Proton structure is a topic that has been studied with LQCD techniques for a long time. Nevertheless, some conceptual questions about renormalization and some technical issues, like the infamous signal-to-noise problem, prevent, as of today, to provide solid first-principles predictions. This is an area of research where much has still to be explored. With researchers at the IFT and IFIC having contributed to our current understanding of the non-perturbative renormalization of QCD, and their expertise in large-scale numerical simulations, CSIC is in a good position to contribute to our understanding of the structure of matter at this level.

Code and algorithmic development

The advent of exascale systems in EuroHPC will be key to the success of the LQCD goals mentioned in the previous sections, which require not just incremental improvements of the current state-of-the-art. It is difficult to overstate the importance of advances in code and algorithmic developments to fully exploit these new computational resources. Driven by research excellence, this is an opportunity to prepare the CSIC community for the new hardware, while fostering our links with other disciplines and industry.

The first pre-exascale systems will be made available in the first half of 2021. In order to be ready we need to develop a roadmap for the CSIC community to be able to exploit Exascale facilities, tying needs for code, algorithmic and hardware developments to physics targets, in cooperation with the rest of the Lattice community in Spain in the framework of the currently funded excellence network LATTICENET. In fact, we can use LQCD as a field test to develop and implement the system software evolution required to fully exploit Exascale level infrastructures involving the implementation of fault tolerance, computational strategies to handle massive I/O through the whole memory hierarchy or testing new architectures. As the cost of the simulations increases the need for an efficient and robust storage and distribution of large datasets becomes a major challenge that needs to be resolved. In particular, this will require a plan for data curation integrated in the European Open Science Cloud strategy of CSIC.

A key element will also be to expand the human capacity through a training program aimed at promoting the necessary skills in the CSIC community to become effective users of Exascale systems. We have an excellent track record at training a skilled workforce, which is often recruited by the private sector at the end of their studies.

Deriving the equation-of-state and transport properties of matter under extreme conditions

A first-principle study of the QCD phase diagram of Fig. 3.1 from LQCD is hampered by the so-called sign problem, and requires further algorithmic progress (see below). On the other hand, effective theories provide a consistent and reliable procedure for the analysis of matter under extreme conditions of density and/or temperature that incorporate the relevant scales and symmetries of QCD.

The physics of QED and QCD plasmas at high temperature T in the perturbative regime is extremely rich. In the early 90‘s it was found out that the soft scales in the plasma, where soft denotes a scale of order gT, and g is the gauge coupling constant, are properly described by the so-called hard thermal loop (HTL) effective field theory, and resummations of Feynman diagrams are needed [Litim and Manuel, 2002]. A member of the ICE institute has shown in the past that these diagrams, and more extensively, the physics of soft scales in the plasma, might be reproduced with the use of simple classical transport equations, which are amenable to study dynamical phenomena numerically. Transport equations also allow us to study the out-of-equilibrium phases of the plasmas, and are typically employed to compute the transport coefficients associated to these extreme regimes of QED and QCD.

The HTL effective field theory is nowadays the standard basis for any dynamical computation of the physics of QED and QCD plasmas at weak coupling within quantum field theory. However, it remains unknown how to push the HTL effective theory to higher orders in the soft momentum and coupling constant expansions, as the HTL are highly non-local.

In the last years a new effective field theory has been proposed by the group at ICE, the so called on-shell effective field theory (OSEFT) [Manuel and Torres-Rincon, 2014], describing only the on- shell degrees of freedom of the system. It has only been derived for the description of the (massless) fermions in a QED plasma to date, and the first quantum corrections to the classical transport equations have been obtained. In particular, and in the presence of fermion chiral imbalance, this allows one to reproduce the quantum chiral anomaly within transport theory. For the first time since the proposal of the HTL effective action, a reliable tool has been devised by a group at ICE to compute perturbative corrections to the soft scales in QED plasmas. The results will allow to address the challenge of computing perturbative corrections to different physical observables.

While initially developed to study only massless fermions in QED plasmas, the aim of the group at ICE is to develop similar techniques to deal with photons. The ultimate goal would be to generalize all these techniques to study also QCD, so as to have a clear strategy on how to improve the accuracy of the study of dynamical properties in weakly coupled plasmas, which cannot be addressed with lattice computations.

As for the non-perturbative regime of QCD, quarks are confined into hadrons, which exhibit highly nontrivial properties and interactions. Some of them, particularly for heavy quark systems, are being discovered in experimental facilities around the world.

The study of hadronic matter at extreme conditions, as those found in heavy ion collisions or neutron stars, requires the interplay between effective field theories and unitarization techniques. Recent successful efforts led by members of ICE have been achieved in the dense regime of neutron stars by addressing the equation of state of hadronic matter, in view of X-ray astronomy missions and the recent observations of gravitational wave signals of two-neutron star mergers [Watts et al., 2016]. This analysis will be extended in the near future by the experts in ICE to the dynamical properties of hadronic matter in the dense interior of neutron stars, by determining the transport coefficients and their role in astrophysical processes, such as neutron star mergers. Also, the exploration of the high-temperature regime in highly energetic heavy ion collisions (RHIC/Brookhaven or LHC/CERN), or the low-temperature and high-density regime created at CBM/FAIR experiment will be of major concern for the unitarized schemes based on effective field theories that include heavy hadronic degrees of freedom, opening new horizons in the physics of dense and hot matter under the CSIC umbrella.

Applications Beyond QCD: BSM and gravity

Many of the proposed BSM theories with the potential to solve several open problems in the SM, such as the hierarchy problem or the strong CP problem, rely on strongly- coupled gauge theories, similar to QCD. It will be possible to adapt the methodology to treat QCD to these extensions of the SM, such as for example composite Higgs models, providing quantitative predictions of their phenomenology that can then be confronted with experiment in future high-energy colliders or precision flavour experiments.

On the other hand, gravity seems to be a completely different to a gauge theory. It has a spin two force carrier instead of spin one, no color degrees of freedom, and a dimensionful coupling. Nevertheless, it appears to be intimately connected with QCD-like gauge theories. The AdS/CFT correspondence relates them holographically as a weak-strong duality, where the Pomeron of Regge theory could play the role of the dual of the graviton. The detailed machinery behind these dualities is still the subject of intense study. Recently, it was pointed out that the emergence of dynamical gravity from a strongly coupled gauge theory is encoded in the quantum entanglement structure of the gauge degrees of freedom. This realization opened the field to the inflow of new ideas from the theory of quantum information.

Feynman [Feynman, 1963] viewed massless Yang-Mills theory as an analog to address unsolved problems in gravity that later led to seminal advances toward perturbative QCD. Also in the high-energy regime, a weak-weak duality exists that relates scattering amplitudes for gravitons with *double copies* of gluon amplitudes [Bern, 2008]. It is quite likely that progress in QCD will also have an impact on our understanding of gravitational interactions.

The key challenge will be to discover and develop new mathematical structures and dualities in gauge field theories, deepening the links with string theory and gravity, both at the methodological as well as a more fundamental level. In the context of strong- weak coupling dualities, a pressing open problem is the precision numerical testing of the AdS/CFT correspondence for concrete gauge theories, an undertaking designed for LQCD techniques. In the context of weak-weak dualities, the challenge is focused in the completely new mathematical structures emerging in the study of scattering amplitudes at higher quantum perturbative orders.

Formal developments

Foundational questions regarding the mathematical status of quantum field theories cannot be disregarded, despite their remote character. In the long run, a key challenge will be to provide a rigorous proof of the mathematical consistency of Yang-Mills theory having as properties confinement and a non-zero mass gap. This can be studied first in certain simpler situations like the large *Nc* limit in the pure glue sector, modifying the topology of space-time or extending the theory to ensure additional symmetries like Supersymmetry. Such a proof will likely have a significant impact on other branches of physics and mathematics. Hints in this direction are already visible by the emergence of unexpected mathematical structures, even in relatively charted waters, such as the amplituhedron and other beasts of projective geometry governing planar scattering in Yang Mills theory, of infinite resummations of instanton corrections via resurgence theory.

Another challenge of great relevance is provided by the study of the topology of the moduli spaces of solutions to the equations defining the minima of the corresponding gauge functional. These include the moduli spaces of solutions to the self-dual Yang- Mills equations, Kapustin-Witten equations, Seiberg-Witten equations, and their dimensional reductions to three and two dimensions. Integration over these moduli spaces is a fundamental source of important mathematical invariants, corresponding to values of certain observables in the physical the theory.

QCD as a test bed for QI technologies

Two main roadmaps have been followed in the attempt to address gauge theories with quantum technologies. One of them is based on numerical simulation techniques and goes under the name of Tensor Network methods, providing a variational ansatz for quantum states which optimizes quantum entanglement. The other relies on the development of quantum hardware, the so-called quantum simulators, that in a given range of energies emulate the physical system of interest. When applied to gauge theories, both approaches rely on the Hamiltonian formulation of gauge theories in contrast to the Euclidean version that provides the basis for Monte Carlo techniques. Practical implementations of these two pathways, generically require the truncation of the infinite dimensional Hilbert space of the gauge theory. Successful approaches have been implemented for simple systems like the 2-dimensional Schwinger model or finite dimensional spin systems like those required in the field of condensed matter. The key challenge is to find ways of extending these techniques to higher dimensional systems, retrieving at the same time the gauge invariant, infinite dimensional, Hilbert space of the gauge theory.

CHAPTER 4

ABSTRACT

"How did the Universe begin?" is one of humanity's oldest questions. The approach to this question gets inputs from many fields, varying from the largest to the smallest scales, from cosmology to particle physics. The last century provided the opportunity to apply the scientific method to this question with results that have established a first comprehensive understanding of the origin, content, and fate of the Universe. The tremendous leap forward in technology experienced in this new century will allow us to challenge this model to extreme levels, establishing it robustly or leading to new and deeper open questions. This chapter is about the understanding of the content of the Universe, its history, and its ultimate fate.

KEYWORDS

Cosmic inflation | CMB | LSS | Gravitational waves | Dark Energy | Dark Matter | Baryogenesis

CHAPTER 4

ORIGIN AND FATE OF THE UNIVERSE

Coordinators
Rocío Vilar
(IFCA, CSIC – UC, Coordinator)
Martin Crocce
(ICE, CSIC – IEEC, Adjunct Coordinator)

Researchers and Centers
Gianluca Calcagni (IEM, CSIC)
Juan García-Bellido
(IFT, CSIC – UAM)
Olga Mena (IFIC, CSIC – UV)
Patricio Vielva (IFCA, CSIC -UC)

1. INTRODUCTION AND GENERAL DESCRIPTION

This challenge will address one of the most fundamental and important topic in sciences: "How did the Universe begin and how has it evolved to its current form?". Approaching these questions has required efforts from many fields, ranging from the largest to the smallest scales, from cosmology to particle physics, and will continue to do so. The last century provided the opportunity to apply the scientific method to these questions with results that have established a first comprehensive understanding of the origin, content, and fate of the Universe. The tremendous leap forward in technology of this new century will allow us to challenge this model to extreme levels, establishing it robustly or leading to new and deeper open questions.

Cosmology, astrophysics and astroparticle physics, on the one side, and particle physics, on the other, are deeply interrelated and try to understand the global properties of the Universe from its origin to the present. In order to address this challenge many topics, varying from cosmic inflation as the mechanism responsible for seeding the cosmic web, to the understanding of the origin of the baryons in the Universe, from the nature of dark matter (DM) as the main source of gravitational force to the one of dark energy (DE) encoding the accelerated cosmic expansion at late times.

In order to understand these questions, we specifically address the following issues from a theoretical and experimental point of view, reviewing present and future experiments. The use of gravitational waves as a new and revolutionary tool to approach to the precise determination of cosmological distances with both standard and dark sirens will be discussed in here.

Cosmic inflation

Current cosmological observations of the Cosmic Microwave Backgrounds (CMB) and the Large Scale Structure (LSS) point to cosmic inflation as the mechanism responsible for seeding the cosmic web. Ongoing and future experiments measuring the microwave sky and mapping the large-scale structure of the Universe could yield the final proof of the validity of cosmic inflation and, even more, could also provide details on the specific properties of the inflationary potential.

Baryogenesis

Visible matter in the Universe is made of fundamental pieces: the elementary particles, divided into fermions and bosons. Among fermions, we have quarks, which are the constituents of protons and neutrons; and charged leptons (electron, muon, and tau particles) or neutral leptons (neutrinos). For each elementary particle, there is an antiparticle that has exactly the same properties but opposite charge. Then, why do we not detect any significant amount of antimatter in the Universe?

Unravelling the origin of the matter-antimatter asymmetry in the Universe (Baryogenesis), which cannot be an initial condition in an inflationary stage of the early Universe, is one of the major unsolved challenges in physics.

Dark Matter

Understanding the nature of DM is one of the fundamental questions of physics in astrophysics, particle physics and cosmology physics. This is endorsed by committees such as Astronet, the European Strategy for Particle Physics 2020 (that includes CERN, DESY, and all the largest labs and institutions, including representations from other countries) and the DOE/NSF in their DM initiatives [Robson et al., 2013; European Strategy for Particle Physics Preparatory Group, 2019; Basic research needs for dark-matter small projects new initiatives, 2019]. Although gravitational effects provide us with overwhelming evidence of its existence, its nature still is unknown and the realm of possibilities for DM is still enormous from particles or waves to compact objects such as primordial black holes. The range of masses of potential DM candidate could cover 90 orders of magnitude in mass. The ideal case would be to be able to cover as much as possible, no stone left unturned in all fields involved.

Dark Energy

The discovery in 1998 that the expansion of the Universe is accelerating has turned into one huge mystery. By now, we have a better established cosmological model and overwhelming evidence of such expansion, but still little fundamental understanding of its reasons. The simplest explanation in terms of a cosmological constant is extremely fine-tuned and possibly not fully successful. Alternative models have been proposed, including new fundamental particle physics and/or fields or modifications to Einstein's general gravity at ultra-large scales. Ongoing and future galaxy surveys will be devoted to establish the nature of Dark Energy, distinguishing among these possibilities.

Gravitational Waves

The recent detection of Gravitational Waves (GWs) by LIGO-Virgo has opened a new observational window to the Universe. GWs from astrophysical sources carry a wealth of precious information about their production (via a waveform analysis) and propagation (via the luminosity distance) which, in turn, depend on the underlying gravitational theory. Thus, gravitational-wave astronomy is an opportunity to test general relativity as well as theories beyond Einstein gravity. The primordial background of GWs (PGWs) from inflation, as well as from reheating after inflation, is a fundamental prediction of any model, and both are potentially detectable by future gravitational wave experiments. The collapse to primordial black holes during the radiation era could also leave a signature on scales that will be probed by LISA and Einstein Telescope. Moreover, astrophysical phenomena like supernovae and kilonovae emit light, GW and neutrinos and their simultaneous detection has opened-up multi-messenger astronomy as a revolutionary approach that can shed light to many, if not all, our challenges.

2. IMPACT ON BASIC SCIENCE AND POSSIBLE APPLICATIONS

Understanding the origin and content of the Universe will have a crucial impact both on basic science and at a philosophical level. Getting advances, total or partial, in each of the key challenges addressed will have a clear impact on the fundamental physics that describe the laws of the Universe. It will help the advance in other areas, thus it will enlighten the path to get the necessary theoretical developments needed to complete the Standard Model of particle physics, for instances, or the Standard Cosmological model.

Basic science leads to discoveries of enormous economic and practical importance although very often there is a time-lag between the fundamental discoveries and their exploitation, but very generally it turns out to be highly profitable, examples such as WWW, the application in medicine, GPS or quantum computing or wireless technologies. These challenges will impact directly in the following:

- Important advances in handling and analysis of large amounts of data in the context of international experiments and Space Science (CERN, ESA, NASA or JAXA).
- Advancing in quantum computing, including the hardware, software and analysis techniques in this area.
- Development of applied science and technology because of the detectors and sensor needed to be developed for the different experiments: development of new materials, new communication technologies, the use of quantum sensors for imaging, and newer and faster electronics.

The scientific innovation and the technological background needed to meet these challenge will allow to generate technology-based companies. These companies strengthened the linkage between universities and industry, and also have a significant impact on regional economic and social development. These challenges provides an excellent training for PhD students, engineers, technicians, and so on in problem-solving, extremely useful skill to exploit while working in applied research or development in industry. Furthermore, this creates very valuable networks of links between researchers in different industries and in academia, which would not exist if all training took place in industry.

Technology transfer provides jobs for graduates and others within the local community, and provides feedback for entire scientific groups, and more rapid development of technological advances into useful products for the marketplace. Finally, the scientific community and the general public stand to benefit from improved communication of basic scientific research. Having a science-literate, or even sympathetic, public has major implications for both the health of our society and for the climate for public funding of research.

3. KEY CHALLENGES

The origin and fate of the Universe evolved through four different regimes: inflation, radiation, dark matter and dark energy dominated era. Despite remarkable success of the standard cosmological model (ΛCDM) in the last 25 years, determination of a number of cosmological parameters has gone from 100% to 1% precision; or the completeness of the standard model of particle physics and interactions (SM) with the discovery of the Higgs boson, the physical laws governing each of the aforementioned cosmological eras are still not understood. The knowledge we have so far comes from decades of theories and innovative experiments. The use of telescopes on the ground and in space has taught us what the Universe looked like when it was young and how it has been evolving. The use of particle accelerators have taught us the basic physics of the very high-energy environment of the early Universe.

We define the following key challenges for deeper research into all these open issues. It is worth recalling that there is a strong interaction among them, because only all of them combined can provide a genuine and unitary view on the origin and fate of the Universe. Addressing these key challenges is a paramount task that involves a very large fraction of the astronomical, cosmological and particle physics communities, in particular, within the CSIC framework. This objective requires detailed plans and the use of significant human, material and computational resources.

3.1. Understanding inflation

Cosmic inflation is the most studied and convincing mechanism of generation of primordial perturbations, which acted as seeds of the cosmic web. In particular, this is supported by the high degree of homogeneity and isotropy of the Universe at scales above a few hundreds of mega-parsecs, the spatially flat geometry of the Universe, and the nearly scale-invariant power spectrum of the primordial fluctuations, which correspond to scalar and adiabatic perturbations of the metric, following an almost Gaussian distribution, as established by the Planck collaboration [Planck Collaboration, 2018].

However, it is commonly accepted that the detection of PGWs generated during inflation will probably confirm this mechanism. The direct detection of PWGs is a tremendous challenge. A stochastic background given by the superposition of gravitational waves of different amplitudes and phases coming from all directions in the Universe has not been observed yet but future missions such as LISA have the capability of detection.

Nevertheless, it is believed that the most promising way to probe the PGWs produced by standard inflation, is through the imprint left on the B-mode polarization of the CMB. The amplitude on these anisotropies is directly related to the energy scale of inflation. This amplitude is commonly defined as the ratio with respect to the scalar perturbations (r), being current constraints $r \lesssim 0.06$ at 95%, when combining information from Planck, WMAP, BICEP2 Keck Array and Baryonic Acoustic Oscillations (BAOs). Future CMB experiments such as LiteBIRD [Lee et al, 2019], Simons Observatory, CM-Stage IV or the European Low Frequency Survey aim to detect primordial gravitational waves associated with $r > 0.001$, which corresponds to inflation energy scales $\gtrsim 5.3\times 10^{15}$ GeV.

It is worth recalling that, for inflationary models beyond the standard single-field slow-roll, precise information about them could be obtained from the direct detection of the stochastic PWGs. In particular, specific features of the inflation potential may enhance the amplitude of the primordial power spectrum of curvature fluctuations on intermediate scales. Such scenario would leave very specific signatures in both the binary black hole coalescence events seen by LIGO and future 3G detectors like the Einstein Telescope, and a stochastic GWs background on the nanoHertz to mHz range of frequencies, detectable with Pulsar Timing Arrays on ground and LISA in space.

Besides the PGWs science, further constraints on inflation can be obtained from the CMB and LSS data. First, the CMB can be used to test the statistical isotropy of the Universe, one of the fundamental consequences of standard inflation. Current observations from Planck have provided strong evidences for it, although some *anomalies* have been found at very large scales. Cosmic variance-dominated polarization maps as the one expected from LiteBIRD could help to shed light on the nature of the large-scale *anomalies*. In addition, galaxy surveys covering large volumes (e.g, Euclid, LSST or JPAS) are also fundamental to study the homogeneity and isotropy of the Universe. Second, observations of both, the CMB and the LSS, are also very powerful to probe the probability density distribution of primordial fluctuations and their Gaussian nature, as predicted by the simplest cosmic inflationary models. The CMB has already placed strong constraints on primordial non-Gaussianity, but the most important improvement in the sensitivity to detect any primordial non-Gaussianity will come from different LSS tracers.

3.2. Baryogenesis

In 1967 Sakharov formulated the required conditions to generate a baryon asymmetry in the early Universe: (a) baryon number should not be conserved, (b) Charge symmetry and Charge-Parity symmetry must be broken, and (c) there should be a deviation from thermal equilibrium in the Universe's expansion. Charge-conjugation parity-reversal (CP) symmetry implies that the physical laws should not change in an antimatter mirror world and it is a key ingredient to explain why there is an excess of matter over antimatter in the Universe. Even if the SM has all the three ingredients required to produce a baryon asymmetry in an initially baryon-symmetric Universe, it is not large enough to match the observations. Baryogenesis may therefore require new sources of CP violation and far from equilibrium conditions that may be attained in extensions of the SM. A very appealing possibility is to link this asymmetry to the neutrino sector (leptogenesis). Searches for leptonic CP violation and also for lepton number violation processes are crucial to test the theoretical leptogenesis framework. Very recent results from the T2K Collaboration exclude leptonic CP conservation (suggesting therefore that leptonic CP violation has occurred) at a 95% confidence level, providing the first hint to the origin of the matter-antimatter asymmetry in our Universe [Abe et al., 2020]. Regardless the soundness of this result, a much higher statistical significance in the measurement of leptonic CP violation is needed before making strong claims. Future neutrino experiments are being designed to find a definite and conclusive answer to this question. These highly multidisciplinary experiments will also test other properties of neutrinos, such as their mass ordering, will serve as neutrino telescopes for supernova neutrino searches and they will also look for neutrinos and antineutrinos from DM decays or annihilations in the galactic halo, in the Sun or in the Earth's interior. The T2HK experiment in Japan and the Deep Underground Neutrino Experiment (DUNE) based at the Sanford Underground Research Facility (SURF) in South Dakota (US), could provide a definitive answer in the quest for leptonic CP violation in the following 15 years.

3.3. Dark Matter

Understanding the nature of DM is one of the most important goals in modern science. Revealing the nature of dark matter —its cosmological origin, its constituents (if any), its interactions among and other properties— connects such disparate scientific areas as the formation of stars and galaxies, the earliest moments of our Universe, and particle physics. Understanding DM also involves synergies across multiple levels between experimentalists and theorists and between the areas above and other disciplines, such as nuclear, atomic, and condensed matter physics. So far, evidence for DM comes from astrophysical observations. BAOs derived from LSS tracers and the CMB, among other cosmological probes, require about 85% of the mass to be in the form of DM. Also, these observations confirm that DM is mostly *cold*, i.e., they behave as non-relativistic particles. Many new DM candidates have emerged lately together with the "traditional" ones such as Weakly Interactive Massive Particles (WIMPs) or axions, because none of those could be found experimentally so far. These new candidates are highly motivated either from theoretical considerations or driven by experimental data, but are qualitatively different in their experimental implications. Theoretical developments have highlighted the importance of searching for DM particles in a range from the lightest mass, consistent with galactic structure, possibly as light as 10−22 eV, to candidates as massive as primordial black holes of tens of solar masses. There exists also the alternative interpretation that DM does not exist, which would require a modification of the theory of gravity.

In order to achieve this range, a diverse and innovative set of experimental proposals have been suggested, including potential game changers; Potential new DM particle candidates could be discovered. Due to special interactions with the detector, this requires the development of new materials and technologies. This implies strong synergies with the technological challenges described in chapter 8.

Solving the mystery of the nature of DM will allow to understand its role in the structure formation of the Universe as well as the role of DM in starting the higher-density "seeds" that led to the formation of galaxies. And since many galaxies have large halos made of dark matter, a question is: How does this affect their interactions with one another?

DM is, certainly, a very important topic for several other challenges addressed in his White Book. Nevertheless, the "Origin and fate of the Universe" provides a unique umbrella to address the understanding of DM from an ambitious perspective. This is due to the synergic essence of this challenge, that, on the one side, contemplates the astronomical, cosmological and particle physics point of views of the role played by DM on the origin and evolution of the Universe, and on its specific properties related to its interaction, constituents, and nature. Bringing together these different disciplines represents by itself already an important step forward, because these communities have only recently started to discuss a common approach to this question. In addition, this challenge also covers instrumental developments, observational strategies, data analysis techniques, and theoretical interpretations. For all these reasons, the solution to this key challenge requires a strong collaboration among different national and international groups, going beyond the typical collaborative networks.

3.4. Dark Energy

The beginning of the last century saw the establishment of General Relativity (GR) and soon after the recognition that the Universe was expanding, supported by direct observations by E. Hubble in 1929 and later by the detection of a cosmic microwave background temperature, consistent with an early dense and hot phase of the Universe. GR predicts that the cosmic expansion, in a homogeneous Universe filled with matter or radiation, will slow down over time. However, towards the end of the century different observations, as well as the simplest models of Inflation, started to point towards the existence of an unknown matter component (cold dark matter) and to the Universe being spatially flat and filled with less than critical mass, implying the need for an energy component with very particular physical properties.

The pivotal change came in the late 1990's when two independent studies of distant supernovae showed that their distance-redshift relation provided direct evidence that the expansion of the Universe has accelerated over the last five billion years. Since then, cosmic acceleration has become one of the most profound puzzles in contemporary physics, with explanations that go from energy components which are not solutions and raise key questions, to abandoning General Relativity as the correct theory on large scales.

Mathematically, the simplest solution to the cosmic acceleration is a cosmological constant, that is, the presence of a fluid whose density is constant in time and with negative pressure. In that context, the last two decades have seen the establishment a concordance cosmological model, flat CDM, that describes several observations together with the accelerated expansion of the Universe. The model relies on two "dark components": DM, unseen directly but essential for the formation of structure, and Dark Energy (DE), the mysterious mechanism set to explain the acceleration.

But the nature of the accelerated cosmic expansion (or "dark energy" as it is commonly referred to) remains a mystery. From a physical point of view this could be viewed as the gravitational contribution of quantum fluctuations, but the required density is small, too small, compared to predictions by particle physics. Other alternatives include fields with time-evolving equations of state (the relation between pressure and density), or non- trivial modifications of GR affecting only large scales (e.g. gravity leaking into extra-dimensions or changing the gravitational action). Any of these theoretical alternatives is in rather desperate need of observational input.

Therefore, the profound implications of cosmic acceleration have motivated the development of wide and deep massive galaxy surveys that map a good fraction of the observable Universe collecting information for hundreds of millions of galaxies, with the aim of measuring the history of the expansion and the growth of structure with percent level precision. The four most well established methods for making such measurements are: the study of the correlated shapes of galaxies (Weak lensing) and their correlated positions (BAO) on large-scales, the abundance of rare objects (Clusters) and the distance to exploding stars as provided by their standardized luminosities (SNIa).

These observations use different astrophysical objects, all tracing the large-scale structure, and hence dealing with completely different observational and theoretical systematics with different reduction pipelines and calibrations. They are all complementary to the high redshift window to a completely different Universe provided by the CMB. For the moment, some mild tensions have surfaced, e.g. on the dark matter density, the amplitude of fluctuations and the value of the local expansion rate H0 [4-8], which makes this quest even more exciting and important. Next generation surveys will dominate the field of DE and challenge the cosmological model as never before, hopefully settling our understanding of dark energy.

3.5. A transverse key challenge: observation of gravitational waves

The recent detection of Gravitational Waves by LIGO-Virgo [Abbott et al., 2016] has opened a new observational window on the Universe. Binary systems of compact objects such as black holes and neutron stars emit tiny ripples in spacetime when they inspiral and merge into a new object. The ground-based interferometric system LIGO-Virgo is measuring several such events per year with astounding precision and a new generation of experiments, both ground- based and space-borne, will come to life in the next two decades, among which KAGRA has just begun operations, LISA [Amaro-Seoane et al., 2017] is unfolding an ambitious science program, and projects such as Einstein Telescope [Maggiore et al., 2020] and DECIGO [Seto et al., 2001; Kawamura et al., 2011] are on the table.

Some astrophysical phenomena like supernovae and kilonovae emit not only light but also neutrinos and gravitational waves, which can be used to understand both the source and the propagation to us of those “messengers”. In this context, *multi-messenger astronomy* is a revolutionary approach, with completely different systematics, to the precise determination of cosmological distances with both standard and dark sirens. Future interferometers will allow us to determine the contribution of primordial black holes to DM, and also open the possibility of detecting the stochastic background of gravitational waves (see section 4.3.1) from the early Universe, be it from inflation, from cosmic phase transitions or from alternative scenarios.

GW astronomy constitutes a challenge on its own due to the science involved, and, indeed, it is largely covered in other chapters of this White Book, but here it is just considered as an additional source of information that opens a new window for unravelling the key challenges previously described.

3.6. Strategy

The general strategy for solving the key challenges goes through a series of different mission, telescopes and experiments that are already operational or will be on short to intermediate time-scales and that will provide enough sensitivity to address the challenges.

A summary of the most important projects is given in Table 4.1.

On the CMB side, the major expertise is founded on the Spanish leadership role in the ESA Planck mission. Future plans run on two different and complementary roads: measurements from the ground and the exploration from space. The major scientific objective is to probe inflation via the detection of primordial gravitational waves.

The QUIJOTE experiment is one of the fundamental seeds for future plans to map the CMB polarization from the ground. In particular, the European CMB community (ECMB) has recognized the need to explore the lowest range of the microwave spectrum (from 10 to 120 GHz). This initiative, called European Low Frequency Survey (ELFS), aims to provide a survey of the CMB anisotropies with instrumental characteristics allowing to probe cosmic inflation. There are several reasons for pursuing this kind of experiment.

Firstly, Europe in general, and Spain in particular has a strong leadership in detector technologies particularly suitable for this frequency range, namely, radiometers and Kinetic Inductance Devices (KIDs). This will achieve polarization sensitivities of around a few microkelvins and a large frequency range to exploit the spectral diversity of Galactic foregrounds, which would allow to measure the imprint of primordial gravitational waves with an amplitude $r \sim 10^{-3}$, compatible with the level predicted by the Higgs or the Starobinsky inflationary models. Secondly, this frequency range is complementary to higher frequencies already planned to be observed by US-led programs (e.g, CMB Stage-IV). Thirdly, the lowest frequencies of the ELFS programme (from 10 to 40 GHz) are also a very interesting complement to future space missions such as LiteBIRD, offering a high-sensitivity mapping of radio foregrounds. In addition, CMB technological developments also account for exploring novel interferometric approaches to map the relic fluctuations, in which the microwave detection is afterwards correlated on the optical/infrared range.

Finally, the JAXA LiteBIRD satellite (with collaborations from NASA, Canada and Europe): currently is the only approved space mission (to be launched in 2027) to map the CMB polarization, represents a unique opportunity for the CMB science. LiteBIRD allows further studies of cosmic inflation, not only through the detection of the primordial B-mode, but also, by probing statistical isotropy and Gaussianity of the E-mode a new opportunity to explore the so-called CMB large-scale anomalies.

The origin of the observed matter-antimatter asymmetry could be related to the process of baryogenesis in the early Universe, via out of equilibrium phenomena occurring immediately after inflation, with new sources of CP violation, or via leptogenesis, with a subsequent conversion into baryons thanks to high-energy sphaleron transition. These new sources of CP violation are searched for in high energy colliders, e.g, LHC-b and Belle2; also ATLAS and CMS could complement these studies.

For testing whether or not baryogenesis took place via leptogenesis processes in the early Universe, leptonic CP violation searches are crucial, as mentioned above. The future DUNE detector will use several thousand tonnes of liquid argon to detect neutrinos from a man-made accelerator neutrino beam at Fermilab in Batavia, Illinois (US), about 1300 km away from the SURF facility. This experiment is therefore complementary to the Japanese one, T2HK, as they use different detector technologies and different baselines.

ProtoDUNE-SP and ProtoDUNE-DP, which have demonstrated the feasibility of the large-scale DUNE detector. With important responsibilities in the design and construction of the first two far detector modules at SURF, particularly in the areas of photon detection and cryogenic instrumentation.

Experiments at LHC, such LHCb and ATLAS/CMS, search for new sources of CP violation, recent results show that there is a light deviation but compatible with SM, more data will be need it to clear this out.

The nature of Dark Matter and its interactions can be probed by several experiments. These different types of searches are complementary. Each come with very different sensitivities in relation to the details of the SM–DM interaction, the detection techniques used, and the assumptions in terms of astrophysics that help to infer the true nature of DM. The search in accelerators attempts to create DM particles from collisions of standard model particles. They must be produced together with SM particles in order to be detected.

The search for the particles mediating between Standard-Model ones open a full spectrum of new particles, for instance the ones associated to the Dark Sectors such as dark photons, or for long-lived particles(LLP)[1] are becoming a hot topic. These particles are important not only for their contribution to DM but also to study baryogenesis processes. Alternatively, it could be that dark sector communicates with the Standard Model, via different portals, such as Higgs, vector, axion, neutrinos portals. Colliders and detectors are the main tools for this. LHC enters a period of hibernation until spring 2021, during which the machine and the experiments will be upgraded, then data will be taken for 3 years until next stop to prepare for HL-LHC. This program will offer huge opportunities to do significant advances relative to today's landscape knowledge. The developments in theoretical models that provides different signatures are crucial for this[2].

For instance, searches for DM candidates associated to top quark or Higgs boson and the more traditional SUSY searches will be improved with better analysis tools, Machine Learning techniques or others. Improved trackers, calorimeters or trigger systems together with new subdetectors, such timing detectors, will be an asset to reach a full DM program. LLPs searches have already appeared as benchmark channels in the detector Upgrade Technical Design Reports and of the Yellow Reports. DM candidates can be searched via their interaction in the detector medium. Measuring these processes would provide information on the candidates and their interaction probability with

1 Accessible at: https://arxiv.org/abs/1903.04497

2 Accessible at: https://lpcc.web.cern.ch/content/lhc-dm-wg-dark-matter-searches-lhc

ordinary matter that complements the information from other sources. There are many initiatives to try to cover a wide range of candidates exploiting different characteristics of the possible interaction with ordinary matter and they are technologically very challenging.

One the technologies that probed to be adequate is the one used at the DAMIC experiment based on the CCDs originally developed for DES. It has been proven that the technology works for mass sensitivity below 1 GeV. Recently they got the best sensitivity to light DM particles interacting with electrons. Soon a new experiment made of 1kg mass of CCDs with single electron resolution, called DAMIC-M, will be ready to take data by the end of 2023. It will search for low-mass in range from 1 eV to few GeV with unprecedented sensitivity to DM-electron scattering and hidden-photon. The CCDs program continues with a future project called OSCURA that it is in the R&D phase funding for the DOE.

Other technologies could be the ones used in axions search, like the haloscope technique in a mass range of 10-100 μeV using an array of small microwave cavities connected by rectangular rises, in an arrangement commonly used in radio-frequency filters. There is a prototype in CAST but they are working in the upgrade exploiting the knowledge of KIDs devices from the CMB, capable of energy-resolving single photons. A future goal will be to installed an instrument in the magnet of the future axion helioscope IAXO.

Neutrino experiments, used as well for Baryogenesis studies, can be used to search for DM. Antares has successfully shown the feasibility of the undersea water Cherenkov technique with a rich harvest of scientific results. KM3NeT is being deployed and will mean a tremendous step forward in neutrino astronomy. The search for DM is one of main goals of the experiment. Neutrino telescopes, and KM3NeT in particular, offer the possibility of looking at several kinds of sources, not all of them available to other indirect searches. DM can be searched for in the anomalous components in cosmic rays due to the annihilation of DM pairs in the galactic halo, on the top of the standard astrophysical production, as well as in γ-rays and, more generally, in multi-wavelength photons. The decay products can be detected by means of ground-based telescopes, balloon-borne detectors and space-based experiments. Fermi-LAT used the data to search for gamma rays coming from WIMP annihilation over the gamma-ray background from different sources.

Analyze for WIMPs and ALPs candidates were explored using LAT data, this provides an alternative and complementary information to those axions searches performed by experiments at the lab. The Cherenkov Telescope Array(CTA's), ground-based gamma-ray instrument, could extend the energy range from some tens of GeV to about 300 TeV, in the search for DM. Current efforts focus on providing realistic predictions in the search of DM in galactic (both dark and visible) satellites and in galaxy clusters. In the future, the combination of Fermi-LAT and CTA data should allow to robustly test the full WIMP miracle [Basic research needs for dark-matter small projects new initiatives, 2019], by testing the thermal relic cross section for WIMP masses between few GeV up to dozens of TeV.

On smaller scales, strong gravitational lensing can map directly the distribution of dark matter in galaxies and galaxy clusters. The latter have provided some of the most direct proofs of DM, especially in colliding clusters, for which the bulk of the baryonic mass physically separates from the DM, offering a pristine view of the DM. Observations of the strong gravitational lensing effect, combined with X-ray observations of the Bullet cluster have been used to set upper limits on the self-interacting cross section. These limits will be improved in the near future with more detailed observations of this, and other clusters.

On even smaller scales, microlensing is being used as well to set limits on the abundance of yet another candidate to dark matter, primordial black holes, that could also be responsible for some of the gravitational waves observed by LIGO/Virgo.

In turn, the Square Kilometre Array (SKA) will be a multi-purpose radio-interferometer with a collecting area of 1 km2, distributed over a distance of at least 3000 km, collocated in Africa and Australia. Qualified as landmark project in the European Strategy Forum on Research Infrastructures (ESFRI) and high-priority in the ASTRONET roadmaps, SKA is a unique instrument, with up to 10 times more sensitivity, and hundred times faster survey capabilities than current radio-interferometers, that will provide leading edge science in the 21st century involving multiple science disciplines. The physics cases of this future multidisciplinary powerful telescope includes Cosmology and Large Scale Structure, Epoch of Reionization (EoR).

Concerning cosmology, SKA will be able to probe DM properties (interactions, velocities and nature) through the detection of the redshifted 21 cm line in neutral hydrogen, during the so-called Dark Ages, before the period of reionization. SKA will also be able to test the Dark Energy properties and the difference between some modified gravity and Dark energy scenarios through HI galaxy number counts, by detecting the 21cm emission line of neutral hydrogen (HI) from around a billion galaxies over 3/4 of the sky, out to a redshift of z 2.

The study of DE and cosmic acceleration revolves around the deployment of very large mappings of the LSS of the Universe. The largest 3D survey of the next decade is DESI, a 14000 deg2 spectroscopic redshift survey that will use a multifiber spectrograph to collect the 3D position of 20M galaxies out to redshift 3.5. It will measure the growth of structure using the anisotropic galaxy clustering pattern. By 2023, the Large Synoptic Survey Telescope (LSST) will start imaging 18000 deg2 of the Southern sky multiple times with an unprecedented sampling rate and to an accumulated depth of 27 mag_{AB}.

It science objectives are several and include the study of dark energy using weak gravitational lensing. Clearly its database will be an invaluable source for astronomical and cosmological research. In turn, Euclid is the biggest European astronomical consortium, selected in 2011 to be one of the next M-class ESA missions. After its launch, scheduled for 2022, it will dominate the field of dark-energy studies by measuring weak lensing and baryon acoustic oscillations over 15000 deg2.

In particular, it will measure the shapes of 1.5 billion galaxies down to magnitude 24.5 and the precise redshifts of over 50 million Hα galaxie in the redshift range 0.9 to 2.1. The development of smaller but very complementary experiments and instruments fully designed, constructed and operated at the national level, such as PAU and JPAS. These experiment provide galaxies surveys using low-resolution spectra from multi narrow band photometry, combining the pros and cons of the larger surveys. The study of gravitational waves can be used as an instrument to understand the origin and the fate of the Universe.

We might be able to link at least some aspects of the key challenges (Dark Matter and Dark Energy) through the study of Gravitational Waves. We can clearly argue that GW-events provide (dark) standard sirens which can be used to study the distance-redshift relation and the expansion of the Universe, putting stringent constraints on Dark Energy Models. In the near future we might the able to study the spatial distribution of GW sources. On the other hand, GW from binary coalescence can shed light on the issue of primordial black holes as a significant constituent of Dark Matter.

TABLE 4.1: Instruments and collaborations with current direct CSIC involvement at the infrastructure level and their contribution in terms of our key challenges. Red tick-marks indicate their main scientific goal, and black ones their transversality into other challenges.

a (Instrumentation) - b (Data & Software) - c (Simulations) - d (Theory) - e (Management)

INSTRUMENT	KC1	KC2	KC3	KC4	TKC	INSTITUTES
CURRENT						
DESI	✓		✓	✓ (red)		ICE[a,c,d], IFT[c,d]
PAU				✓ (red)		ICE[a,b,c,d,e], IFT[c,d,e]
JPAS	✓		✓	✓ (red)		IFCA[b], IAA[b,c]
QUIJOTE	✓ (red)				✓	IFCA[a,b,c,e]
DAMIC/DAMIC-M			✓ (red)			IFCA[a,b,c]
LHC		✓	✓			(IFCA,IFIC)[a,b,c,d,e], IFT[d,e,c]
Fermi-LAT			✓			IFT[b,e]
Belle2		✓	✓			IFIC[a]
ANTARES			✓			IFIC[a,b,c,e]
KM3NET			✓		✓	IFIC[a,b,c,d,e]
T2K		✓ (red)	✓			IFIC[b]
FUTURE						
EUCLID	✓		✓	✓ (red)		ICE[a,b,c,d,e], IFCA[b], IFT[c,d,e]
LSST	✓		✓	✓ (red)		ICE, IFT (negotiating participation)
SKA	✓			✓		IFCA[b], IFIC[d], IAA[a,b,c,e], , ICE[d]
JWT			✓			IFCA[b]
ET	✓		✓	✓	✓ (red)	ICE[a,d,e], IEM (negotiating part.), IFT[b,c,d,e]
ELGAR			✓	✓	✓ (red)	ICE[a,d,e]
LiteBIRD	✓ (red)				✓	IFCA[a,b,c,e]
ELFS	✓ (red)				✓	IFCA[a,b,c,e], CAB[a]
LISA	✓		✓	✓	✓ (red)	ICE[a,b,c,d,e], IFT[c,d], IEM[d], IFIC[d]
DAMIC-M/OSCURA			✓ (red)			IFCA[a,b,c]
HL-LHC		✓	✓			(IFCA,IFIC)[a,b,c,d,e], IFT[e,d]
RADES			✓ (red)			IFCA[a,b],CAB[a]
CTA			✓			IFT[b,c,e,d]
future accelerators		✓	✓			IFT[b,c,e,d]
DUNE		✓ (red)	✓			IFIC[a,b,c,e,d], IFT[d]

KC1 (Inflation) - KC2 (Baryogenesis) - KC3 (Dark Matter) - KC4 (Dark Energy) - TKC (Grav. Waves)

CHAPTER 5

ABSTRACT

This challenge aims at understanding how the large–scale structure of the Universe originated after the Big Bang, starting with the first stars that ignited at around 300 Myr and leading to the diversity of galaxies in the present time. We need to shed light on the mechanisms that regulate the connection between the dark matter halos originated from primordial fluctuations and the growth and evolution of galaxies, filaments, and (super)clusters. The assembly of galaxies depends on non-linear processes such as star formation, feed- back and quenching, black hole accretion, interaction with the environment, and galaxy merging. Key questions are: How did the Universe emerge from the dark ages? How did the structure of the cosmic web evolve? How did the present-day Hubble sequence of galaxies form? What was the formation history of the Milky Way?

KEYWORDS

Structure formation and evolution

Galaxies assembly and evolution

Galaxies, clusters and dark matter haloes

Massive black holes and active galactic nuclei

Gas inflows and outflows

The Milky Way Laboratory

CHAPTER 5

FORMATION AND EVOLUTION OF GALAXIES AND LARGE STRUCTURES

Coordinators

José Manuel Vílchez
(IAA, CSIC, Coordinator)

J. Miguel Mas-Hesse
(CAB, CSIC – INTA, Adjunct Coordinator)

Researchers and Centers

Almudena Alonso-Herrero
(CAB, CSIC – INTA)

Santiago Arribas (CAB, CSIC – INTA)

Francisco J. Carrera (IFCA, CSIC – UC)

José María Diego (IFCA, CSIC – UC)

Miriam García (CAB, CSIC – INTA)

Rosa González-Delgado (IAA, CSIC)

Jordi Isern (ICE, CSIC – IEEC)

Jorge Iglesias-Páramo (IAA, CSIC)

Jesús Maíz-Apellániz (CAB, CSIC – INTA)

Isabel Márquez (IAA, CSIC)

Mar Mezcua (ICE, CSIC – IEEC)

Miguel Ángel Pérez-Torres (IAA, CSIC)

Miguel Ángel Sánchez Conde
(IFT, CSIC – UAM)

1. INTRODUCTION AND GENERAL DESCRIPTION

This challenge aims at understanding the present day large–scale structure of the Universe. To properly understand this fundamental question we need to start from the study of the primordial structures and galaxies, traced by the light of the first stars that ignited after the dark ages, around 300 Myr years after the Big Bang. These galaxies are distributed along a web-like filamentary structure, as has been found by the many galaxy surveys performed since the pioneering works in the 1980s. This cosmic web outlines the large–scale picture of the Cosmos and is constituted by dark matter and gas, upon which galaxies are built, with dark matter making up the most important fraction of the Universe's gravitating mass. It is gravity that binds the large structure of filaments, galaxies, clusters and superclusters of galaxies together, along bridges of dark matter and gas which is way too dim for an easy detection. This large–scale structure is pervaded by the dark energy repulsive field, still far from being properly understood in the present days.

We aim to shed light on the mechanisms that regulate the connection between the dark matter halos originated from primordial fluctuations and the growth and evolution of (super)clusters, filaments and galaxies (see also Chapter 4). The growth of these structures and the assembly of galaxies depend on non-linear processes such as star formation, feedback and black hole accretion, or galaxy merging, among others, having led to the diversity of galaxies in the present–day Universe, as reflected in the Hubble sequence of galaxies. Understanding galaxy formation and evolution constitutes one of the paramount challenges for present day astrophysics and space science, in order to obtain a coherent picture of the history of the Universe after the Big Bang, since the early dark cosmic age to our local cosmic backyard. In particular, and more specifically the following questions need to be answered, as different pieces of a big puzzle that only together will provide a coherent view:

1. What are the masses and sizes of dark matter (DM) halos and galaxy cluster structures at all scales?
2. What were the properties of the first luminous objects (stars and accreting black holes) and how did they form? What was the stellar initial mass function at the Cosmic Dawn?
3. How did chemical evolution proceed and influence structure formation?
4. How did galaxies assemble and evolve?
5. What was the origin and growth history of massive black holes?
6. What were the role and properties of gas inflows and outflows (mass, momentum, radiation and feedback)?
7. How did structure formation depend on and influence its environment?
8. What can we learn from "laboratories" in the local Universe: the Milky Way and nearby galaxies?

To advance in the understanding of these key questions a new generation of ground- and space-based observatories and instruments is being developed all over the world.

As we discuss below, the challenges and specific objectives discussed in this chapter are very closely linked to other topics within the Theme "Understanding the basic components of the Universe, its structure and evolution", creating relevant synergies between researchers working on these questions:

- "Origin and fate of the Universe"
- "Gravity"
- "Understanding the cycle of matter in the Universe"
- "New instrumentation and techniques"
- "Understanding matter and radiation under extreme conditions"

2. IMPACT ON BASIC SCIENCE AND POTENTIAL APPLICATIONS

The questions to be studied in the next decade will have a strong impact on our conception of the Universe as a whole and our place in it, from physics to philosophy. The questions that will be addressed in this challenge lie at the heart of modern astrophysics. We want to stress that the concept of an expanding Universe, originated from a singularity, emerged from the study of the location and velocity of large samples of galaxies almost one century ago. Furthermore, a deeper insight into the large structures of the Universe led to the discovery of both dark matter and dark energy, two elements that revolutionized our knowledge of the Universe and whose nature we don't know yet. Research findings from this challenge will surely contribute to a better understanding of the properties of both dark matter and dark energy, also of central interest for challenge 4 "*Origin and fate of the Universe*", providing a cosmological model able to reproduce the observational data. Furthermore, the development of the challenge needs a broad and ambitious observational approach that will require to build new observational facilities and advanced astronomical instrumentation, both on Earth and in Space, which will lead to state of the art technological developments. The analysis of the data to be produced will also require the development of new techniques in the fields of Machine Learning, Artificial Intelligence, and Big Data. This means that the development of this challenge will provide in addition a significant feedback in the technical capabilities of our society, that can extend far beyond astronomy (e.g. image processing and pattern recognition techniques in medicine).

3. KEY CHALLENGES

3.1. Precise properties of dark matter haloes and galaxy clustering structures on all scales

In the standard cosmological model (the so–called ΛCDM framework), nearly 85% of the matter in the Universe is composed of a yet unknown form of non-baryonic matter. This so–called dark matter (DM), that is supposed to be collisionless, neutral and weakly interacting with the ordinary, visible matter, is nevertheless instrumental for the formation and evolution of structures across cosmic history. Indeed, ΛCDM predicts that small structures entirely constituted by DM, known as *halos*, were the first ones to form by gravitational collapse and then, by accretion and merging, gave rise to larger structures.

These halos acted as the gravitational seeds in which the baryonic matter would fall and ignite galaxy formation in later times, thus shaping the Universe as we observe it today.

Much progress has been made in the past decades to elucidate the statistical and structural properties of DM halos and their substructures, or subhalos. Nowadays, state-of- the-art hydrodynamical simulations including both DM and baryons, and involving up to hundreds of billions of particles, are capable of forming galaxies whose properties match reasonably well those of actual galaxies. Yet, despite the undeniable advancement, key questions remain.

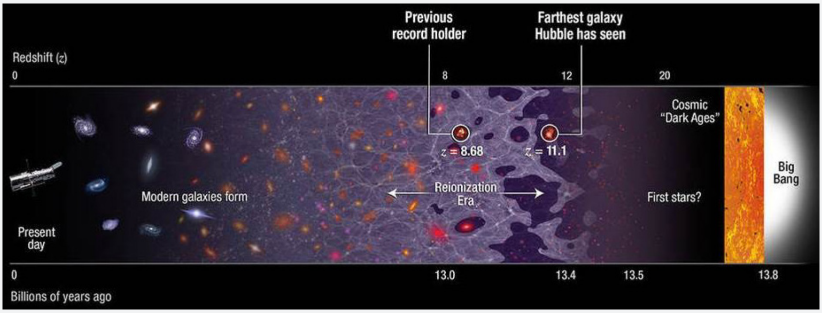
Simulations are not able to resolve the whole halo hierarchy. The current resolution limit of $\sim 10^5\, M_\odot$ (being $M_\odot$ one solar mass) at redshift $z = 0$ is orders of magnitude above the minimum predicted halo mass, of $\sim 10^{-6}\, M_\odot$. ΛCDM predicts the existence of dark, Earth-mass structures while we cannot observe dwarf satellites below $\sim 10^8\, M_\odot$. Yet, different techniques have been proposed to unveil these *dark satellites*, whose discovery would be a definitive support for ΛCDM: study of gaps in stellar streams, strong lensing, and gamma ray emission in case the DM be made of WIMPs (weakly interacting massive particles). With the advent of future observatories like Euclid and CTA, in which CSIC researchers are deeply involved, new data will become available that will allow to probe smaller masses.

A mismatch between simulations and observations occurs at the scale of galaxy clusters as well, the former systematically providing less concentrated DM density profiles with respect to that inferred from actual cluster mass estimates using X-ray measurements, Sunyaev–Zeldovich effect and strong lensing. Being galaxy clusters the largest bound structures being formed in today's Universe, recent assembly history and merging shocks, infall/outflow of material at the outskirts, violent baryonic-related episodes, etc. are all capable of modifying not only subhalo abundances at these extreme scales but also the inner and outer slopes of the cluster's DM density profile.

3.2. The Cosmic Dawn: formation and properties of the first luminous objects

According to the current paradigm the first stars in the history of the Universe ignited around 300 Myr after the Big Bang, when the clouds of primordial hydrogen and helium fragmented and collapsed down to protostellar scales. Since these first generations of stars were composed of only H and He (with some traces of other elements), their properties should have been quite different to the stars in

FIGURE 5.1—GN-z11 is currently the oldest and most distant known galaxy in the observable universe, detected when the Universe was just ~ 400 Myr old. It is 25 times smaller than the Milky Way and has just one percent of our galaxy's mass in stars. However, it is forming stars at a rate about 20 times greater than our galaxy does today (Credit: NASA/ESA).



the present Universe. While the Initial Mass Function (IMF) of these extremely low metallicity Pop III stars is not known, many theoretical studies indicate that stars with unusually large masses (up to 1,000 $M_\odot$) may have formed, even preferentially. On the lower mass end, it is presently assumed that low mass stars ($M < 1\ M_\odot$) could not have formed from the primordial pristine gas.

The evolution to supernovae of these Pop III stars depended critically on their initial mass. Evolutionary models of Pop III stars indicate that in the mass range 130 to 260 $M_\odot$ they exploded in pair-instability supernovae (SNe), releasing about half their mass in metals, but not leaving any black hole as remnant, while those with masses above ~ 300 $M_\odot$ essentially collapsed into a black hole without producing metals. If such large initial masses were confirmed, these stars would have been extremely short-lived, merely 1–2 Myr, almost a sudden flash in the history of the galaxies.

- Determining the IMF of these Pop III stellar populations is a key challenge for the next decade, since this is essential to understand the formation of the first stars and galaxies, as well as the formation and evolution of the first black holes and the production of metals in the first hundreds Myr.

Only recently we have been able to get hints of these first generations of stars. Combining Hubble and Spitzer photometry and spectroscopy, a galaxy at redshift $z \sim 11.1$ was discovered in 2016, merely 400 Myr after the Big Bang. Unfortunately, present observatories allow to just detect these primordial galaxies, but it is still not possible to characterize their stars.

- Observatories starting operations in the next 5–10 years, to whose development CSIC researchers have contributed and will have thus privileged access, such as the James Webb Space Telescope or the Extremely Large Telescope, will provide for the first time the capability to perform spectroscopy of these first generations of stars, entering a new phase of characterization rather than just identification.

The formation and rapid evolution of very massive Pop III stars would have produced a much stronger relative ionizing flux than for stars below $\sim 120\, M_{\odot}$. This strong ionizing flux would have produced extremely large values of the Lyman α and $HeII\lambda1640$ emission lines equivalent widths. Moreover, a significant fraction of it escaped and led to the complete reionisation of the Intergalactic Medium (IGM). The reionisation started in bubbles and superbubbles around the galaxies with the highest star formation rates, and through percolation they completed the reionisation of the Universe by $z \sim 6$, within the first 1 Gyr of evolution of the Universe, as shown in Figure 5.1.

While Lyman α would be very difficult to detect, since the mostly neutral IGM would have multiply scattered and in practice destroyed these photons, the $HeII\lambda1640$ emission line and the associated stellar rest-UV continuum should be detectable with the JWST up to redshifts $z \sim 10-12$. A detailed study of star-forming galaxies above $z \sim 7$ with JWST/NIRSPEC will allow to give clues on the properties of the IMF of primordial Pop III stars. CSIC researchers are heavily involved in the development of the JWST, being part of the Scientific Teams of the instruments and having access to guaranteed time.

- Redshifted 21 cm radio emission will reveal the response of the IGM to these sources, allowing to perform a 3-D map of HI during the Epoch of Reionisation. SKA pathfinder facilities such as LOFAR, HERA and MWA will play a key role in the next years collecting data with the goal of measuring this signal.

While galaxies dominated by star formation have been detected very close in time to the onset of the first stars, the X-ray background demonstrates that Activ Galactic Nuclei (AGN) emit one order of magnitude fewer UV photons at $z \sim 6$ than required to reionize the IGM. The relative number of AGN at high redshift is very small, constraining the details on the mechanism that led to the formation of supermassive black holes in the center of galaxies. In any case, the rapid evolution of very massive Pop III stars should have produced a significant population of intermediate mass black holes, that could have merged at the centers of galaxies within the first 1 Gyr. They would have become so the seeds of the supermassive black holes that developed later, as discussed in Sect. 5.3.5.

- A proper identification of accretion dominated vs. star-formation galaxies in the Primordial Universe could be performed by detecting and analyzing the profiles of their Lyman α and *He II*λ1640 emission lines. A detailed analysis of the evolution of the fraction of both mechanisms with time will be essential to understand the process of massive black hole formation in parallel to massive star formation.

3.3. Chemical evolution and structure formation along the history of the Universe

The first metals appeared after the first (massive) stars exploded as supernovae, ejecting the elements synthesized in their nuclei to the interstellar medium. As mentioned above, the short life of Pop III stars implied that the contamination of the interstellar and even the intergalactic media proceeded very rapidly. Since then, the content of metals in galaxies has evolved with cosmic time, being determined by the history of star formation and the massive events of gas accretion and outflows inherent to galaxy formation and evolution. Chemical evolution is then directly linked to galaxy formation and evolution, as exemplified by theoretical simulations. A tight relation between metallicity and total stellar mass of galaxies, the mass-metallicity relation, has been observed to hold locally and at high redshift. In addition, the scatter of this global relation correlates with the rate of star formation in galaxies, as a secondary parameter, resulting in what is dubbed the Fundamental Metallicity relation. It has been claimed that this relation holds since early epochs of galaxy evolution – as expected from simulations – when massive accretion of (near) pristine gas from the cosmic web fed star formation, diluting the interstellar medium metallicity. Little is known about the evolution of these processes of massive accretion over cosmic time, or on how the Fundamental Relation evolved from Cosmic Dawn to Cosmic Noon epoch, when star formation rates peaked.

- How do these processes impact on the early chemical evolution of galaxies remains a challenge to be met; the access to the future Extremely Large Telescope ELT and the new space (e.g. JWST) facilities would give a competitive advantage to our CSIC groups to study this unsolved question.

There are several open questions in nucleosynthesis that constitute relevant challenges for the next 10 years:

- An improved understanding of the nucleosynthesis of heavy elements should decipher the role played by the environment and metallicity of stellar progenitors, their rotation, and the derivation of precise yields. All these factors are essential for interpreting observed abundances in galaxies and can be successfully tackled by CSIC researchers. Critical open questions here include: i) the evolution of binary systems, needed to understand the evolution of Supernovae (SN) Ia progenitors, and ii) the unknown origin of the heaviest neutron rich r-process elements that cannot be created by Core Collapse SNe, which do not explain either the observed anomalies in the abundances of elements beyond $Z = 29$, as Se, As and Ge.

3.4. Assembly and evolution of galaxies

As described in Sect.5.3.1 the evolution of cold dark matter is the current basis for the cosmological picture that explains the large-scale structure and evolution of the Universe, from the first generation of dark matter halos and sub-halos to present day galaxy clusters and superclusters. It is also believed that the stellar components formed at $z \geq 2$ likely evolved into elliptical galaxies and bulges, probably through mergers of the *primordial* star forming disks. However, this framework falls short to explain the properties of the galaxy population at $z \sim 1$ and how the present–day Hubble sequence of galaxies was assembled. The reason is that the growth of galaxies is not related in a simple way to the growth of dark matter structures. The interplay of energy and matter exchange (between the process of gas accretion and cooling and star formation) is essential to understand the growth of the gaseous and stellar components in galaxies.

There are already several observational fundamental results that are key to trace the cosmic evolution of the star formation in the Universe. The star formation rate density in the Universe peaked at redshift z ~ 2, and declined thereafter. At any z, star forming galaxies show a correlation between Star Formation Rate and the total mass in stars, known as the Main Sequence of star formation. A list of important issues yet to understand include the effect of the environment in these relations, the link between total stellar mass and evolutionary state, how much of the stellar mass is already in place at each epoch, and how much of the star formation is related to environmental effects. Further, we do not know which are the evolutionary tracks that gas rich, disk, star forming galaxies formed at very high redshift followed to evolve to the massive-quiescent galaxies observed at $z \sim 2$. Panchromatic large sky surveys led by CSIC researchers will pursue these questions, in particular:

- SKA, the Square Kilometer Array, will play a key role to fully understand the growth of mass, gas and stars in the Universe and their cosmic evolution. SKA will allow radio continuum surveys, where dust obscuration is no longer an issue as is the case of the rest-frame ultraviolet emission. It will trace the star formation rate of obscured and unobscured galaxies of un-biased samples from the high redshift Universe to the present. The dominant emitter population are core–collapse supernovae from evolved stars of 8 $M_\odot$ or higher, that will trace the star formation rate of the obscured and unobscured star forming galaxies that are very abundant at high redshifts. CSIC researchers play a significant role in the development of SKA and will have guaranteed access to its data.
- JWST and J-PAS will be crucial to understand the formation and evolution of massive galaxies (similar to or more massive than the Milky-Way), that are the major contributors to the peak of the star formation activity. Ultra-deep imaging and spectroscopy of galaxies at $z > 2$ from the JWST (MIRI, NIRSpec), and J-PAS narrow band imaging covering from 3500-9500 Å, equivalent to spectra with resolution $R \sim 50$, for galaxies at $z \leq 1$ will be crucial to distinguish between the different scenarios proposed by galaxy formation models for Milky-Way like galaxies.

However, galaxies are not simple structures. During the last decade, we have carried out a significant effort to understand how the present-day Hubble sequence of galaxies was assembled, as traced by their dynamical mass, stellar and gas content, and metals. Integral Field Spectroscopy (IFS) of nearby galaxies has enabled a leap to constrain the spatially-resolved galaxy formation models, providing 2D spatial + 1D spectral information for each galaxy. However, a wide and deep view is not yet available due to the limitations of the number of galaxies in the samples studied. Future IFS surveys led by CSIC researchers will improve significantly these restrictions:

- J-PAS will do IFS-like observations with a spectral resolution equivalent to R ~ 50, mapping the stellar population properties of ~100,000 galaxies at $z \leq 0.1$ with a sampling of 0.23 arcsec/pixel to research the effect of the environment in the formation and evolution of the Hubble sequence.
- An Integral Field Spectrograph at the Calar Alto Observatory and WEAVE@WHT will target galaxies of the local Universe to provide constraints to the sub-grid physics for the formation of disks in M31 and galaxies similar to the Milky-Way.

- HARMONI and MOSAIC at ELT will be crucial to get 2D kinematics information and 2D maps of stellar populations properties of galaxies and ionized gas distribution of galaxies at the peak of the star-formation activity of the Universe, and study the progenitors of Milky-Way like galaxies.
- SKA will provide HI galaxy mapping of nearby galaxies, providing 2D spatial +1D kinematic information of the neutral gas and constraining the interplay between gas inflows, outflows, radial gas flows, and the star formation efficiency in galaxies. In combination with IFS surveys, we will study the role that the environment plays to transform galaxies along the Hubble sequence.

3.5. Origin of galactic black holes and evolution of nuclear activity in galaxies

A result emerging from the last decade of observations is that there are apparently supermassive black holes (SMBH) with masses > $10^6\ M_{\odot}$ in the centers of most, if not all, major galaxies. These SMBH grow by the gradual accumulation of matter (accretion) which releases its potential energy in the form of radiative and kinetic energy, shining over the full electromagnetic spectrum (giving rise to the Active Galactic Nucleus phenomenon) and generating particle jets and matter outflows with the potential to affect the entire galaxy. This fact, together with the tight relationships of the masses of the SMBH with some properties of their host galaxies and the parallel evolution of the growth of galaxies by star formation and of SMBH by accretion, strongly suggests that AGN and galaxy formation and evolution are strongly linked. AGN are thus fundamental ingredients to understand galaxy evolution, as well as being bright beacons up to the Cosmic Dawn and fascinating objects in their own right.

The discovery of AGN with SMBH when the Universe was less than 1 Gyr old ($z > 6$) poses pressing questions about the assembly and early growth of galaxies and SMBHs. These are thought to grow from seed BHs of $10^2 - 10^5\ M_{\odot}$ formed either from Pop III stars or direct collapse of primordial gas, as discussed in Sect. 5.3.2. Besides, hundreds of low-mass SMBH (< $10^6\ M_{\odot}$) are being found in dwarf galaxies. These are thought to be the relics of the early Universe seed BHs and can shed further light onto this hot topic.

- Moving beyond discovery towards statistical population studies and detailed characterization of high redshift luminous AGN and seed BHs, are two of the main objectives of several observational facilities. Working in multiple electromagnetic bands, CSIC scientists are involved and have significant presence in many of them: JWST will perform deep surveys and detect thermal and line emission of the AGN, ALMA is already and will continue studying dust and molecular gas around the AGN and in the host galaxy, SKA will explore star formation rate and AGN incidence from the mid-2020s on, and ELT will probe rest-frame optical/UV emission of the host galaxy and AGN from 2025 on.

Closer to the current epoch, the emission of radiation by star formation and AGN shows a broad peak around $z \sim 1-3$, but most of the emission produced by accretion in the Universe is obscured. The obscuring medium is responsible for hiding the AGN nature of many sources in different bands and it is also thought to be related to the bewildering variety of AGN types. In addition, some of these obscuring circumnuclear structures (from about a few to ~100 parsec from the SMBH) may be part of the launching pads of some of the galaxy-scale outflows that shape the host galaxy. Understanding the distribution and nature of the obscuring medium and completing the census of the AGN hidden by it is mandatory to fully assess the prevalence of AGN among galaxies and their mutual dependence.

Again, this problem requires a multi-wavelength approach to study different manifestations of the obscuring medium and to uncover AGN signatures. Observations with ALMA are already revealing the geometry and kinematics of the obscuring material and incipient outflows in the nearest AGN, and can study host galaxy properties at higher redshifts. The advent of JWST in the next few years will revolutionize this field by uncovering in the mid-infrared reprocessed continuum radiation from the obscuring material, and detecting direct spectral evidence for AGN presence to larger distances and earlier epochs than currently possible. In the mid-term, ELT spectroscopy in the optical – near infrared will allow disentangling spectroscopically AGN and host galaxy emission in the rest-frame optical/UV in much fainter and farther sources.

- Providing and selecting targets for the studies described above will require sensitive surveys over large areas: J-PAS will provide in the mid-term (≥ 2026) a sensitive multi-filter and multi-band survey of a significant part of the northern hemisphere. Concurrently, the EUCLID space mission (~ 2022 – 2028) will map a third of the sky in the visible and near-infrared bands. By the second half of this decade SKA surveys in the radio band will uncover yet more AGN via their radio excess emission.

An important parameter to deal with in the study of AGN is their intrinsic power, with low luminosity AGN most probably being related to smaller accretion rates onto the SMBH. Athena, at the beginning of the next decade, will detect typical AGN up to $z \sim 7$ and characterize AGN with extreme obscuration up to $z \sim 3$. SPICA/future far IR facilities will detect even less luminous extremely obscured AGN at the same redshifts or equally luminous ones at even younger epochs of the Universe, and characterize low luminosity ones up to $z \sim 2$. This will be done through extensive surveys and targeted observations. The launch of LISA around 2035 will extend the range of AGN studies to the gravitational wave domain.

- The direct detection with LISA of SMBH mergers over the entire history of the Universe will provide precious information on the incidence of this alternative mode of SMBH growth.

3.6. Role and properties of gas inflows and outflows

Galaxies are open systems that evolve into a quasi-stationary state, where inflows and outflows of gas balance the star formation rate and AGN activity, and therefore the global properties of their stellar populations and central black holes. Gas accretion from the cosmic web maintains galaxies forming stars during long periods of time. Star formation (SF) and/or black hole activity due to this gas accretion, in turn, expel the surrounding interstellar medium by the mechanical and radiative energy liberated, generating outflows of gas. The role of outflows governing the subsequent galaxy evolution is believed to be crucial, as they can regulate and quench both star formation and black hole activity, being also the primary mechanism redistributing dust and metals over large scales within the galaxy, or even expelled into the circumgalactic and intergalactic media. Therefore, the galaxy mass function, the mass-metallicity relation, and the black hole-spheroid mass correlation, are thought to be shaped by outflows.

Although this general scenario is relatively well established, it requires the self-regulated balance among gas accretion, gas outflow and SF/AGN activity which depends on the coupling between physical processes that involve multi-phase gas at very different physical scales, from cosmic web structures to molecular clouds. This implies a significant challenge for both theoreticians and observers, and it is therefore clear that it will remain a hot topic of research for the next decade and beyond.

A key element for our advancement in this field will be the advent of new first-class observing facilities over the coming years, like the James Webb Space Telescope (JWST), the Extremely Large Telescope (ELT), and the Square Kilometer Array (SKA), among others. These facilities, which have been (or are being) developed with the active participation of CSIC scientists, will provide unprecedented some uniquely suited capabilities to address the most challenging questions in this field of research like:

What was the role of inflows and outflows in the Early Universe?
Theory predicts that cold gas accretion is particularly important at high redshift, when dark matter haloes have low mass and the accreted cosmic gas remains cold and ready to form stars at high rates. Similarly, strong outflows are expected to play a pivotal role in shaping primeval galaxies, and their subsequent evolution, as well as in redistributing dust and metals into the circum- and the intergalactic medium. The rather limited observational constraints provided by current facilities suggest that galactic winds are ubiquitous at intermediate redshifts (i.e. $z \sim 2$), and may be extremely powerful at very early epochs.

The JWST, with its orders of magnitude improvement in sensitivity, will allow us to explore the epoch of early galaxy evolution with unprecedented detail. Thanks to its multi-object capability, it will obtain a complete census of the properties of galactic outflows across cosmic time. This statistical approach is key to understanding their role in the emergence of the different galaxy populations. Moreover, the integral field unit will obtain spatially resolved spectroscopy (at sub-kpc scales for $z = 2 - 10$) of the extended structures associated to the outflows, and therefore dig into the physics behind the global integrated properties. On the other hand, observations with the ELT will permit to study high-z galaxies at spatial resolutions similar to the ones we now achieve with local samples.

What is the overall feedback impact of the outflow phenomenon?
Local samples will continue being ideal laboratories to study in detail the physics of outflows, and their effects. Key observational inputs, like precise outflow mass, energy, and momentum rates for the different gas phases will be obtained in nearby objects to evaluate feedback effects on the star formation and AGN activity in the host. Not only *negative* feedback (i.e. reduction of star formation as a consequence of gas ejection) will be a matter of detailed study, but also *positive* feedback (i.e. outflowing gas may foster star formation). Main related questions that need well-grounded answers include: what is the fraction of outflowing gas that escapes the galaxy and is able to enrich with metals the intergalactic medium? are mass-loading and mass-loss rates in low-mass galaxies as high as predicted?

The combination of sensitivity and high angular resolution at optical and near infrared wavelengths provided by the ELT, will provide a detail far greater than never before. Its integral field spectrograph (HARMONI) will explore the vicinity of the AGN and the regions of intense star formation, where powerful galaxy-wide outflows are generated, with resolutions of up to 10 milliarcseconds (equivalent to ~ 0.15 parsecs at the distance of M82), tracing the ionised, neutral, and warm-molecular gas phases. The cold molecular gas is currently observed with ALMA, while the "MHONGOOSE" Large Program at MeerKAT will be unique to unambiguously detect cold accretion in the next 5 years. It will probe a factor ~ 50 deeper in column density. Finally, SKA will extend this study to larger samples providing higher angular resolution and access to deeper column densities.

3.7. Interdependence between large structures and galaxy environment

The primeval uniform distribution of matter evolved (through gravitational interaction) into an intricate foam-like structure, with nearly empty voids surrounded by sheets, filaments, and clusters containing most of the mass and galaxies. The void regions account for most of the volume of the Universe but include only a small fraction of all galaxies. Around the voids, sheets of galaxies curve and form filaments in the intersection regions between two sheets. Filaments themselves merge in the nodules of this network forming the densest and massive structures, galaxy clusters, that can harbor up to hundreds of galaxies per Mpc^3.

At the center of these clusters, and as a result of cannibalizing unfortunate neighbors, we find the most massive galaxies, the brightest cluster galaxies, or BCGs (also cD galaxies), with masses well in excess of 10^{12} solar masses. Often supermassive black holes (SMBHs), with masses as high as several times 10^{10} solar masses, lurk at the center of these very massive galaxies. These SMBHs can affect the dynamics and structure of the BCGs (and its surroundings), not only during their active phase and through feedback phenomena from the material falling in them, but also through their quiescent phase through gravitational effects that can eject material, including stars, from the central regions. The centers of these massive galaxies are useful probes of dark matter, in particular if dark matter has a small probability of interaction with itself (parameterized by its cross- section). In addition, the large potential wells of galaxy clusters can heat up the gas in the central region up to several keV. This hot intracluster plasma radiates away energy, mostly through X-ray emission (Bremsstrahlung).

Moreover, galaxy clusters can also cluster, forming vast overdense regions known as superclusters. These superclusters can dominate the dynamics around them, up to scales of hundreds of Mpc. Thus, environment is a key factor in determining the type of evolution of a galaxy. Low and high density environments appear different in terms of the types of galaxies they host. In the local Universe, low density regions contain a large fraction of spiral galaxies. On the contrary, in dense environments like galaxy clusters, elliptical galaxies are the dominant type. The evolution with cosmic time and environment of the early/late type mix in cluster galaxies is a matter of study. Galaxy clusters frequently show also material stripped from galaxies by ram pressure as they cruise through the intracluster medium. Star formation can be triggered in this stripped material which adopts a jellyfish structure, with star forming regions trailing the galaxy.

Answers to all these questions relative to structure formation, galaxy evolution, and its dependence on environment will inevitably require also the theoretical input from numerical simulations.

3. 8. Laboratories in the local Universe: The Milky Way and nearby galaxies

Massive star formation.

Massive stars are the great galactic disruptors, as their influence is found across large distances due to their ionizing radiation, kinetic energy output (winds and supernovae), and chemical enrichment. How efficient these effects are depends, to a large extent, on their proximity in space and time: one supernova alone cannot puncture the interstellar medium of a galaxy but hundreds of them in a stellar cluster exploding within a few million years could be able to do so. Therefore, it is crucial to understand how massive stars aggregate around each other. The standard picture for some time was that "all stars form in clusters" and that the Initial Mass Function was quasi-universal; requiring that hundreds of low-mass stars be formed for each massive star. In the last two decades that has changed and it is now accepted that massive stars can also form in unbound stellar associations, following hierarchical patterns in space and time, and it is possible to find massive stars under conditions very different to those in clusters.

- The challenge for the next decade is to identify how massive stars are distributed within a few kpc of the Sun. This will be accomplished through a variety of surveys: astrometric (Gaia), photometric (e.g. GALANTE), and spectroscopic (e.g. WEAVE). Combining them we should be able to disentangle the impediments posed by extinction, crowding, multiplicity, and distance indetermination and obtain the full 6-D (spatial + velocity) distribution of massive stars in our neighborhood.

On the other hand, massive stars signal on-going star formation in galaxies. They log the latest history of assembly of the bulge of the Milky Way, and can help to understand the processes shaping dwarf irregulars (dIrr), –alleged ancestors of dwarf ellipticals and possible building blocks of massive galaxies–. In the Local Group we can assess the significance of gas inflows/outflows, interactions and other processes such as stripping, for these galaxies.

CSIC is leading the efforts to locate and characterize massive stars in Local Group galaxies but, as of today, the largest telescopes of the world are only scratching the surface. We are promoting the construction of a multi-object spectrograph for the ELT, MOSAIC, and instrumentation for future large space telescopes such as LUVOIR, that will reach individual, faint massive stars. With this study:

- We will identify the processes shaping star-formation in dIrr's, and whether they impinge a specific IMF. Updated star formation rates, complemented with measurements of the elusive molecular gas content with the SKA, will allow a stress-test of the Kennicut- Schmidt law (empirical relation between the gas density and star formation rate in a given region), helping to establish their true star formation efficiency. On the other hand, our recent results indicate that extinction is not negligible in dIrr, implying that the mass stored in dust and stars is underestimated. At CSIC we are also making a fundamental contribution to future FIR missions and we will contribute to:
- Accounting for the dust content that escaped Herschel's sensitivity limit, deriving the total mass locked in gas, stars and dust, fundamental to compute the dark matter budget of dwarf galaxies.

Galactic Archaeology.

The goal of Galactic Archaeology is to reconstruct the history of the Galaxy using the fossil stars present in its structures. This fossil content is made of low mass stars and white dwarfs. This ambitious goal can be achieved thanks to the first maps with information of the chemical composition, positions and kinematic properties provided by Gaia for a statistically significant number of stars. The first results have shown that the evolution of the Galactic structures is probably at odds with what was traditionally assumed (e.g. that the halo is composed by stars of Galactic origin and stars accreted via collisions like the Gaia-Enceladus event).

The chemical and kinematic properties are not enough to extract the historical record contained in the fossil stars and precise dating of the sample is critical. For low-mass stars this quantity can be obtained fitting the color-magnitude diagram (CMD) of a given population if the distance is known. Gaia has provided accurate distances to stars within an important volume ($\approx$ 2 kpc around the Sun) so it has been possible to apply this method. However, to avoid the degeneracy of the CMD method between age and metallicity of red giants it is necessary to go down the turning point, where stars are dimmer; but in the CMD they can overlap within the observational errors for a wide range of masses and evolutionary stages.

With the advent of asteroseismological capabilities from space it is now possible to determine accurate properties from these stars including their age, provided that the metallicity is known (e.g. recent determination of age of νInd). CSIC researchers have a leading role in the M3 ESA mission PLATO (at co-PI level), which will characterize by asteroseismology hundreds of thousands of stars from 2026 on, providing an invaluable database to reconstruct the history of the Galaxy, and complementing Gaia.

An additional, complementary way to obtain the age of a population is through the cooling of white dwarfs. They are the final evolutionary stage of $\leq 8\, M_\odot$ stars, and their cooling age and luminosity are related allowing to use them as chronometers. Main uncertainty comes from metallicity that strongly affects the lifetime of their progenitors, but massive ones have relatively massive progenitors with short lifetimes that can be neglected. A drawback is that white dwarfs are dim objects and for the moment they can provide only local information. This will be overcome in the future when LSST will provide information of an unprecedented deep sample containing $\sim$ 50 million white dwarfs.

The emerging history is that our Galaxy has been growing at the expense of its vicinity and that, at the same time, has been strongly perturbed by its neighbors:

- The challenge is to determine this history answering the associated questions: i) Which is the contribution of external galaxies to the evolution of the Milky Way? ii) When and how the bar and the bulge formed and evolved? iii) Which is the relationship between the thin and thick disks and the origin of their population? iv) How important is the radial migration of disk stars? v) How was the halo assembled? In all cases the estimate of the age of stars with a precision better than the 10% is a must and constitutes one of the key challenges for the next 10 years.

CHAPTER 6

ABSTRACT

This challenge addresses the processes that form and transform nuclei, atoms, molecules, and dust into stars and planets across the history of the Universe. This baryonic matter follows a life-cycle that is intimately related to that of stars (from birth to death). The enriched chemical products of stellar evolution replenish the surrounding interstellar medium through stellar winds and supernova explosions, which, in turn, trigger new star formation and maintain the cycle alive. This challenge requires to understand the formation, physical properties, chemical composition, and evolution of stars, protoplanetary systems, and planets, as well as the interplay with their natal interstellar medium. It involves the application of several techniques: from astronomical observations and development of new instrumentation, to powerful computational models and laboratory experiments.

KEYWORDS

- Stars: formation and evolution
- Galaxies assembly and evolution
- Planetary systems: formation and evolution
- Interstellar medium
- Element abundances and stellar nucleosynthesis
- Astrochemistry

CHAPTER 6

UNDERSTANDING THE CYCLE OF MATTER IN THE UNIVERSE

Coordinators
Francisco Najarro
(CAB, CSIC – INTA, Coordinator)
Javier R. Goicoechea
(IFF, CSIC, Adjunct Coordinator)

Researchers and Centers
Guillem Anglada (IAA, CSIC)
Aldo Serenelli (ICE, CSIC – IEEC)
Pedro Sarriguren (IEM, CSIC)
Jose Luis Taín (IFIC, CSIC – UV)
Miriam García (CAB, CSIC – INTA);
Izaskun Jiménez-Serra
(CAB, CSIC – INTA)
Ignacio Mendigutía (CAB, CSIC – INTA)
Jose Francisco Gómez (IAA, CSIC)
Tomás González-Lezana (IFF, CSIC);
Martín A. Guerrero (IAA, CSIC)
Luis F. Miranda (IAA, CSIC)
Carmen Sánchez-Contreras
(CAB, CSIC – INTA)

1. INTRODUCTION

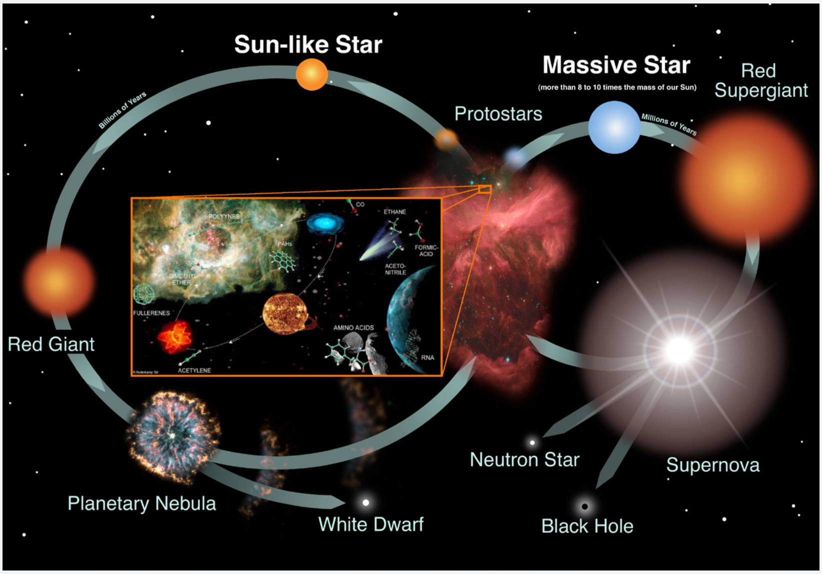
This fundamental science challenge aims at addressing the processes that form and transform nuclei, atoms, molecules and dust grains into stars and planets, and vice versa, the enrichment and return of matter to the interstellar medium (ISM, the space between stars). This baryonic matter, the stuff we are made of, follows a life-cycle that is intimately related to that of stars, from their birth to death (Fig. 6.1). The formation and fate of a star basically depends on its mass and composition. Depending on its initial mass, the star will follow two well differentiated evolutionary paths: a slow track for low-mass stars (the vast majority of stars in the universe) and a faster track for high-mass stars (at least 8 times the Sun's mass). The latter are very scarce and their evolution is far less understood. Although massive stars "live fast, die young", they have a profound impact on the interstellar environment and on the life-cycle of matter. Further, they also "leave beautiful corpses" in the form of neutron stars or black holes which, if formed in binary systems, may end up merging and sending us a final post-card through gravitational waves.

The first step of the cycle occurs inside stars, where nuclear fusion converts light nuclei into heavier elements at the stellar core. Some of the nuclear reactions taking place in stellar interiors, or at the stellar dead-throes and final explosions, are still not well characterized. This limits the predictability of nucleosynthesis models (see Fig. 6.2). In any case, observational evidence shows that the enriched chemical products of stellar evolution replenish the surrounding ISM through stellar winds and supernova explosions which, in turn agitate the ISM, trigger new star formation, and maintain the cycle alive. The ISM is the reservoir of baryonic matter and, as galaxies evolve, its constituents are gradually converted again into stars by processes that are not yet fully understood. Long thought to be a scaled up version of low-mass star-formation, our knowledge of how massive stars form is still very incomplete. A critical difference is that massive stars reach maturity (emitting mighty ionizing radiation and powerful winds) when they are still embedded in their natal cloud. Thus, their impact on the surroundings is more dramatic. At the other extreme of the stellar mass distribution, the lack of appropriate observations has rendered very difficult to assess the formation of very low-mass stars and brown dwarfs.

In many objects, molecules are the most usual form of baryonic matter. Molecules exist in a very large variety of environments: from interstellar clouds, to planet-forming disks, to the atmospheres of (exo)planets and cool stars. Around 200 species, including prebiotic molecules connected to the origin of life, have been detected in space, more than 20% first discovered by CSIC researchers. Unfortunately, we still do not understand how these molecules form and lead to the observed diversity and complexity. Even though the dust-to-gas mass ratio in the ISM is only about 1%, dust plays many fundamental roles. However, we do not know the processes that lead to the formation of dust in evolved stars. Yet we do see rocky planets. Thus, it is crucial to understand how dust grains grow and how they are ultimately incorporated into planet-forming disks, to bridge this enormous gap in size (microns to km). Indeed, the detection of proto-planets is a great observational challenge and a mandatory step to constrain the existing theories of (exo)planet formation.

All in all, this challenge translates into understanding key aspects of the formation, physical properties, chemical composition and evolution of stars, protoplanetary disks and planets, as well as their interplay with the ISM. We advance here that this is a monumental task that includes the study of objects and processes with energy-, time-, and spatial-scales that vary by tens of orders of magnitude. Indeed, understanding the life-cycle of baryonic matter, in our Galaxy and beyond, requires a very multi-disciplinary approach, and involves the simultaneous participation of several disciplines: from stellar to interstellar

FIGURE 6.1–Simplified view of the life-cycle of matter in a galaxy



physics, from nuclear to molecular processes; and the use of very diverse techniques: from observations at nearly all wavelengths, to powerful computational models and laboratory experiments (able to reproduce the relevant astrophysical processes), to developing new instrumentation for future telescopes.

The main questions that we would like to answer in the next years can be summarized in the following challenges (that we develop in more detail in Section 6.3):

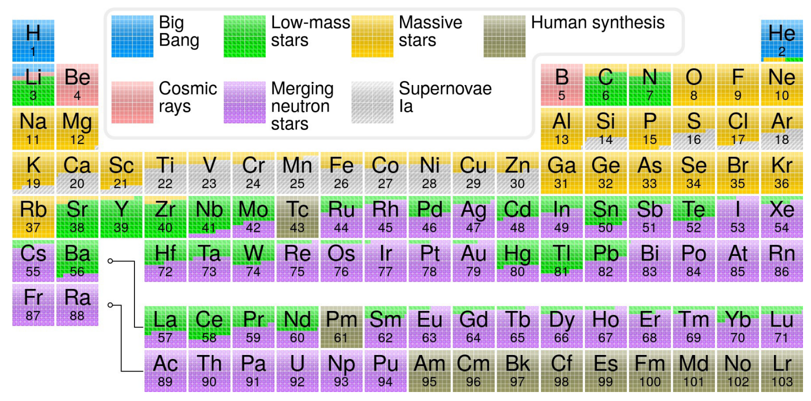
1. How do stars and planets form? What mechanisms transform molecular clouds into young stars surrounded by planetary systems?

6.3.1) Understanding high- and low-mass star formation. Studying the evolution from interstellar clouds to protoplanetary disks.

6.3.2) Understanding planet formation: Protoplanetary disks and planet formation mechanisms. Bridging the gap between dust grains and already formed planets.

6.3.3) Evolution of chemical complexity: from simple to prebiotic molecules.

FIGURE 6.2—Proposed origin of different elements currently found in the Solar System and neighborhood, still a puzzle under debate in nucleosynthesis.



2. How do stars evolve from maturity to death across cosmic history?

6.3.4) Understanding massive stars and their evolution as cosmic agents. Enrichment and feedback with the galactic environment. Supernovae explosions and formation of black holes and neutron stars. Gravitational wave sources.

6.3.5) The life cycle of low-mass stars. Asteroseismology. Evolved stars and their circumstellar envelopes as present-day dust and heavy elements factories. The end points of stellar evolution. Nucleosynthesis in cataclysmic events.

3. Which are the nuclear, atomic, and molecular processes that drive and enrich the cycle of baryonic matter?

6.3.6) Understanding nuclear reactions and nucleosynthesis in stellar evolution. 6.3.7) Characterizing the key physical and chemical processes that drive the evolution of the interstellar and circumstellar media.

2. IMPACT ON BASIC SCIENCE AND POTENTIAL APPLICATIONS

This challenge is at the crossroads of astrophysics, astrochemistry and nuclear physics, with relevant implications for particle physics, cosmology, planetology and theories of the origin of life. It deals with the immense task of understanding the formation of elements, molecules and dust, how they aggregate into the basic building blocks of the observable sky: stars and planets, and how this matter is returned back to the ISM. Planets are a *by-product* of the poorly known mechanisms that convert a fraction of the ISM into new stars. The detection and characterization of mature exoplanets is a very active field or research. Here we propose to go one step back and try to understand how these planets actually form. This includes studying planet-forming systems and detecting individual proto-planets. In at least one of these planets, ours, the type of complex molecules we currently detect in the ISM and around protostars conspire and lead to the emergence of life. Understanding how planets form and how molecular complexity evolves from interstellar clouds to planets will have a deep impact, both scientifically and for society.

The Milky Way is the best laboratory to study the many aspects of the life-cycle of baryonic matter in great detail (at high spatial resolution), providing useful templates to better understand the spatially unresolved emission from more distant galaxies. One example is the *feedback* of massive stars (the impact of their winds, ionizing radiation, and supernova explosions) at disrupting interstellar clouds and star formation. Recent studies of nearby regions by CSIC astronomers suggest that the relevant *feedback* processes act on much smaller spatial scales (0.2-2 parsec) than are resolved by current cosmological simulations that simulate the evolution of our universe (more than 50 parsec). Future observations of representative samples of massive star clusters using next-generation instrumentation will provide validation for these theoretical models used in broader contexts.

It will be impossible, however, to understand the cycle matter from the perspective of a single area of research. Instead, and this is a major change that will impact the way we do science, it will require the simultaneous participation of areas of expertise able to encompass the tremendous range of scales and processes we need to investigate. Understanding processes as “simple” as the agglomeration of a submillimeter-size dust grain in a protoplanetary cirumstellar disk or the photo-dissociation of a prebiotic molecule by an ultraviolet photon requires that astronomers, quantum chemists, surface scientists, and laboratory experts sit together and discuss at the same table. These are just two examples. Developing networks or platforms that promote these collaborations will imply that CSIC researchers may be able to have a leading role in solving some of these challenges.

Finally, parallel societal and applied benefits often result from the development of our instrumentation, which finds application in medicine for nuclear astrophysics, or in communications, electronics, and cryogenics for other branches of the research described in this chapter. Recent examples of our groups are the development of a proton scanning device and a neutron dosimeter for proton therapy treatments and a gamma-ray imaging device for diagnosis in oncology, or the use of radio receivers in laboratory experiments.

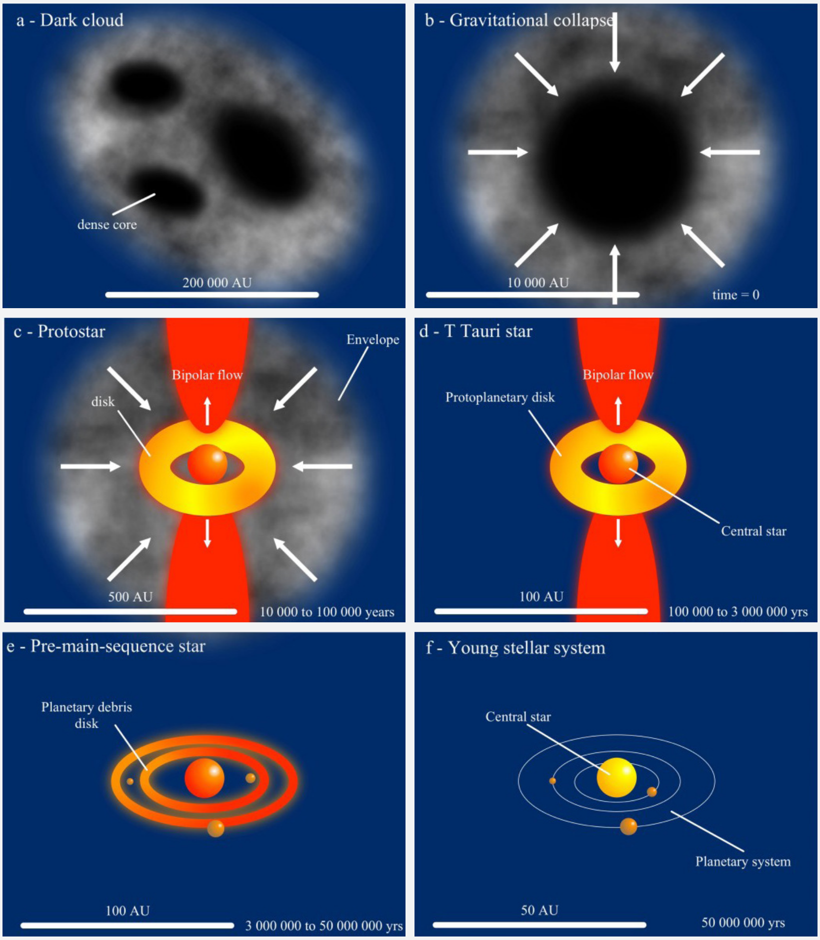
3. KEY CHALLENGES

3.1. Understanding high- and low-mass star formation.

The formation of stars is one of the fundamental processes in nature. It is an extremely complex process, which transforms the huge (tens of pc in size), diffuse (just a few hundred particles per cm^3) and cold (about 10 K) interstellar clouds of molecular gas, into hot and dense objects such as the stars. During this process the clouds of interstellar matter will shrink in size by about seven orders of magnitude, their density will increase by about 20 orders of magnitude and their temperature by six orders of magnitude. Therefore, a precise tracking of all the phases of this process, in which the physical conditions undergo such drastic changes, is a real challenge, which requires powerful observational facilities, computing tools and a deep knowledge of the physical (and chemical) properties of matter in conditions far beyond the usual ones in terrestrial laboratories.

Nowadays it is well established that individual stars form as a consequence of the gravitational collapse of dense cores of molecular gas and dust (~0.1 pc scale) resulting from the fragmentation of interstellar (pc-scale) molecular clouds (Fig. 6.3a-b), frequently after the development of filamentary structures. Because of the rotation of the core, matter does not fall directly onto the central (proto)star but through a circumstellar accretion disk that is developed at scales of ~10-100 au. A fraction of the infalling matter is ejected in a direction perpendicular to the disk, in the form of a collimated jet that removes the excess of mass and angular momentum (Fig. 6.3c-d), thus allowing the formation of the star. In turn, a planetary system can be formed as a result of the evolution of the accretion disk (Fig. 6.3e-f). Thus, the investigation of the fragmentation and collapse of molecular clouds, and the development, evolution and properties of the protostar-disk-jet systems are essential steps to better understand the process of formation of stars and planets. Both gravity and magnetic fields are essential ingredients in the whole process. Radical progress has been made on these topics since the pioneering simulations of a collapsing cloud by Larson, fifty years ago, and we anticipate even more advances in the coming decades. Specifically, we identify the following key challenging points:

FIGURE 6.3—Sketch of the different stages in the overall process that transforms interstellar molecular clouds (a-b) into stars and planets (e-f). The development of a disk-jet system (c-d) plays a central role in the whole process [Johnstone, 2018].



How is star formation triggered in molecular clouds? The connection with small-scale cores. The role of interstellar filaments and the factors that determine the process of cloud fragmentation must be investigated. The relationship between the initial mass function of the cores and that of the resulting stars, as well as the whole process of clustered star formation must be better understood.

How do massive stars (> 10 $M_\odot$) form? Because of their fast evolution and large distances, the early stages (prior to the HII phase) in the formation of massive stars are still poorly known. Can the process be a scaled-up version of low-mass star formation (e.g., monolithic collapse), or is a totally different process (e.g., coalescence of low-mass stars) required? Recent results (observation of disks and jets in massive stars) suggest a similar process, but there are no good examples for the most massive stars. What is the maximum mass that a star can reach?

How do young stars grow? Understanding star formation requires knowing how material from the surrounding circumstellar disk accretes onto the stellar surface. For low-mass young stellar objects (YSOs) there is consensus that accretion is magnetically driven, while for more massive stars without magnetic fields it is believed that accretion may proceed directly from the disk to the star through a "boundary layer", but it is still necessary to understand how young stars of all masses grow through disk-to-star accretion. Intermediate mass Herbig Ae/Be stars represent a fundamental regime that bridges the gap between the accretion properties of low and high-mass young stars. Understanding stellar accretion has implications on the way that "macroscopic" parameters like the "star formation rate" are estimated, on the disk lifetime and dissipation processes, on modeling protoplanetary disks, and on outflows driving the angular momentum transfer.

How are protostellar jets ejected and collimated? Young stars are associated with ionized jets characterized by thermal emission which provides a means of determining their physical properties, such as density, mass-loss rate, and temperature. It is believed that these jets are ejected through a magneto-centrifugal mechanism, but the magnetic field, which must play a fundamental role in this mechanism, is still a major unknown in YSOs. Our detection of non-thermal synchrotron radio emission in a protostellar jet may represent a major milestone to understand the ejection/collimation mechanisms in YSO jets, since synchrotron emission allows to measure and map the magnetic field. There are synchrotron jets in many other astrophysical scenarios (relativistic jets in X-ray binaries or active galactic nuclei), whose magnetic fields have been studied for decades, but lacking thermal emission to determine other physical parameters. A systematic study of thermal and non-thermal emission in YSO jets could decipher their ejection/collimation mechanism, making it possible to extrapolate it to all astrophysical jets (in synergy with Challenge 9.10). The strong shocks in the non-thermal lobes of the YSOs' jets seem to be equally key to understanding the generation of cosmic rays in the Galaxy. On a broader scope, understanding how magnetic fields collimate the ionized astrophysical jets is of interest for the design of Tokamak fusion reactors, whose purpose is to confine by a powerful magnetic field a hot ionized plasma to produce energy in a controlled thermonuclear fusion (synergy with Theme 8).

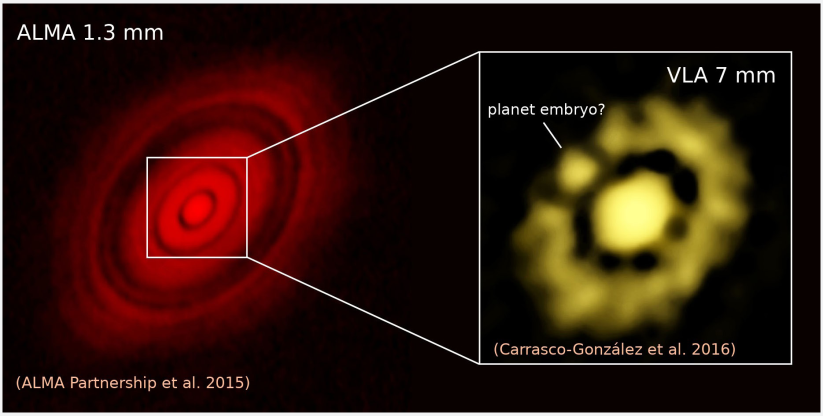
3.2. Understanding planet formation.

It is well established that star and planet formation are intimately related processes that occur during the first ~10 Myr of stellar evolution. Also, there is solid evidence that the formation of planets occurs in disk-like structures of gas and dust, called "protoplanetary disks", that surround the young stars. The first indications of such structures came from the analysis of the spectral energy distributions (SEDs) of young stars, which show infrared excess above the stellar emission due to the dust of the disk. Later-on, direct imaging confirmed that circumstellar disks are indeed common structures during all the stages of the star formation process. However, the exact planet formation mechanism is still unknown. Also, it is unknown at which stage in the star-forming process planet embryos do start to appear. Two main paradigms are considered. The "core-accretion" model is a bottom-up view where micron-sized dust particles in disks coagulate and grow until massive enough solid cores experience runaway accretion to finally form planets. The "disk-instability" scenario is a top-down perspective where planets are directly formed in collapsing regions within the disks that are cold and massive enough to experience gravitational instabilities. An important difference between these two views is the timescale of planet formation: 3-10 Myr according to core-accretion, and 1 Myr according to the disk instability model. Thus, a better understanding of how planets form requires to constrain when the very first steps of the planet formation process start to take place.

Most of the observational information supporting our planet formation theories is based on small (micron to submm sized) dust particles at early stages of the process, and on already formed planets (>1000 km) at the latest stages. Thus, it is of crucial importance to bridge this gap by directly detecting planets that are actually forming. Four decades ago, the quest for protostars was considered by Gareth Wynn-Williams the "Holy Grail" of infrared astronomy. In the present decade, the direct detection of a protoplanet has become an equivalent key challenge. The compilation of a large enough sample of protoplanets in different evolutionary stages, the possible study of their environment and associated physical processes, the detection of circumplanetary material and accretion onto the planet (and possibly outflow) would clarify the whole process of planet formation. When planet formation is better understood, then even the quest for proto-moons can become a very active area of research in the coming decades.

The problem of understanding planet formation is identified by the major space and astrophysical agencies as one of the most important challenges for

FIGURE 6.4–(Left) ALMA and VLA images of the HL Tau protoplanetary disk, revealing a sequence of bright and dark rings [ALMA Partnership et al., 2015]. The VLA image, at a longer wavelength and lower opacity, shows details of the innermost ring substructure that remain hidden in the ALMA image, suggesting the fragmentation of the ring to form a planet embryo [Carrasco-González et al., 2016].



the next decades. Only by understanding the processes that lead to planet formation we could put our own solar system in perspective. This has an obvious, direct impact on society's self-perception in a universal context. Some key challenging points are described in what follows.

The quest for protoplanets. Our current knowledge about planet formation is fundamentally limited by the scarcity of detections of forming planets. Detecting a true planet embryo (i.e., a protoplanet) is a challenging task since the presence of circumstellar material and the activity of young stars make it difficult to apply the methods generally used for mature stars. Our current observational efforts rely on high-spatial resolution techniques, either aimed at the direct detection of accretion signatures of circumplanetary material falling onto the planets, through imaging and spectroscopy at optical and IR wavelengths, or the detection of the circumplanetary material itself, through imaging with large cm/mm interferometers. The future development of appropriate observational techniques, along with the use of forthcoming facilities promise a strong improvement on our detection rate of forming planets around young stars. This will allow us to obtain reliable statistics for many types of stars and different evolutionary stages, in order to get robust answers on the timescale of planet formation and, thus, on the planet formation mechanism itself.

Protoplanetary disks. It is necessary to determine the physical/chemical properties of protoplanetary disks, their evolution and comparison with the properties of mature exoplanetary systems. Also, it is necessary to find ways to accurately measure disk masses, e.g. based on stellar accretion rates, different gas probes, and detailed disk modeling. It is not well understood how microscopic dust grains can grow up to planetary scales since a number of mechanisms in the disk work as barriers to their growth. This study is hampered by the fact that some key processes occur in regions of the disk that are opaque even at mm wavelengths, and only can be explored at longer wavelengths (near 1 cm) where the emission is optically thinner and can penetrate these regions (Fig. 6.4). However, since dust emission decreases steeply with increasing wavelength, very sensitive observations are required. This kind of observations are currently a big challenge but they should be easier to perform with the forthcoming large facilities such as the ngVLA and the SKA. Protoplantary disks are made of gas and dust. However, gas is almost absent in the interplanetary medium. Thus, the last stages of disk evolution should be characterized by the dispersal of its gas. Photoionization and photoevaporation by high energy radiation (mainly extreme UV and/or X-ray radiation) are favored mechanisms for gas removal. However, it is still unclear how these processes work and which is the dominant one.

What determines the final mass and composition of a planet? A major question in the formation of a planetary system is to discern what determines the formation of a rocky planet, like our Earth, versus a giant gaseous planet, like Jupiter. The study of the so-called snowline (the frontier between the outer parts of the disk where molecules are frozen onto dust grain mantles and the inner parts where they are released to the gaseous phase) can provide the clue to solve this question, but the role of planet migration and the relationship with the metallicity of the host star must also be better understood. Other pending questions include: What is the minimum mass for a disk to form a planetary system? What is the maximum mass of a star to host a planetary system? Which is the role of magnetic fields? How do planets form around binary stars?

Debris disks as a tool to study exoplanetary systems. The outer part of our solar system is populated by a large number of rocky bodies and dust distributed in a ring called the Kuiper Belt. This belt constitutes a relic of the accretion disk that surrounded the young Sun. Imaging similar dust belts around close-to-Earth stars with known exoplanets can inform us on the architecture and dynamical history of their exoplanetary systems, as we have done in an exploratory work with ALMA of Proxima Centauri, the star closest to the Sun. This kind of observations inform us on the dust content of the interplanetary medium in these systems, which is related with meteorite impacts and possible mass extinctions, and therefore with planet habitability. Also, it is of great interest for trajectory planning and optimization of projects like the Breakthrough Starshot mission, which aims to send ultra-light spaceprobes to Proxima Centauri at a fraction of the speed of light, since collisions with dust particles could pose a fatal threat to any of these high-speed probes.

3.3. Evolution of the chemical complexity: from simple to prebiotic molecules

The number of molecular species discovered in space gives an idea of its high level of chemical complexity. Understanding how this degree of complexity is attained in different environments (from interstellar clouds and planet-forming disks to the surface/atmospheres of comets, moons or planets) represents a great challenge that involves different disciplines (astrochemistry, laboratory astrophysics, planetary geology, and astrobiology).

Among these molecules, there is a sub-set termed “complex organic molecules” (or COMs) defined as carbon-bearing compounds with more than 5 atoms. Some of these COMs are of prebiotic interest since they represent key ingredients in theories of the origin of life. Examples are formamide (NH2CHO), glycolaldehyde (CH2(OH)COOH) or urea (NH2CONH2), which could have been precursors of amino acids, sugars and nucleobases on a young Earth and perhaps in other planets. Despite the increasing number of detected prebiotic molecules in the ISM, we still do not understand when and how they form.

We currently know that chemical complexity starts in dense and cold (about 10 K) starless/pre-stellar cores, the precursors of Solar-type systems. Several COMs have been detected in these objects, revealing an unexpectedly rich chemistry. Once a protostar is formed and heats its environment, chemical organic complexity reaches its maximum during the stage called “hot core” (around massive protostars) and “hot corino” (around low-mass protostars). The first detections of the prebiotic molecules urea or glycolonitrile (HOCH2CN) have been reported, respectively, toward these two environments. At a later stage, a few COMs such as methanol (CH3OH) have been recently detected in proto-planetary gas disks. The organic reservoir in disks, however, is likely locked as ice mantles around grains. JWST will probe this solid material. CSIC researchers are frequent users of ALMA and VLA (detecting gas-phase COMs) and have guaranteed time in JWST (to image the ice emission in protostars and protoplanetary disks).

A major question to answer in the next years is whether all this organic material can be transferred to, and retained by, small Solar-system bodies. And also whether the prebiotic content can be delivered to young Earth-like planets, triggering the biochemical processes involved in the origin of life. In this context, the main goals are: *i*) to understand how complex the chemistry can be along the process of star formation (i.e., whether even larger prebiotic COMs such as amino acids, complex sugars or nucleobases can be found in space); and *ii*) to determine whether COMs can be delivered to young planets that present similar conditions to those of an early Earth. CSIC is very well positioned to reach these goals because several groups are very active in this field. Both observationally (using ALMA, VLA, IRAM, and JWST and SKA in the future) and developing dedicated astrochemical models (IFF and CAB). These groups should strengthen their collaborations with theorists and experimentalists (at IEM, CAB, and other institutions) who are experts in the synthesis and spectroscopy of COMs (see Sect. 6.3.7).

3.4. From the First Stars to the present-day Universe: Massive stars as cosmic agents

Cosmic History has witnessed the lives and deaths of multiple generations of massive stars. In life they are extreme sources of UV-radiation that create ionized bubbles and inject kinetic energy into the interstellar medium (ISM). Their extremely disruptive death as supernovae (SNe) and/or γ-ray bursts (GRBs) is also a source of energy and life: it releases fresh atoms produced by the fusion nuclear reactions that fueled the star along its evolution. This is the origin of most of Oxygen and other elements crucial for life (P, Si, S, Na, K, Ca, Mg) that were inherited by planetary systems like our own later on.

All this interaction with the environment, referred to as *feedback*, enters small and large-scale processes spanning the age of the Universe, including triggering or inhibiting subsequent generations of stars and planets, dust creation and destruction, and galactic- scale gas flows in star-forming galaxies. Many astrophysics fields ingest models of the formation and evolution of massive stars as a function of chemical composition, which is a proxy for the varying environment from the Big Bang to the present day.

Massive stars have been extensively studied in the Milky Way (MW) and the nearby Large and Small Magellanic Clouds (LMC, SMC). Their chemical metallicity (Z), ranges from Solar-like ($Z_\odot$) to 1/5 $Z_\odot$. In terms of cosmic history, 1/5 $Z_\odot$ means that we are not looking back enough in time to cover the characteristic metallicity during the tperiod of maximum star formation in the Universe (~10,000 million years ago). *We also lack information at the extremes:* the first stars of the Universe (Z~0), crucial to model the first 200,000 years after the Big Bang, and the massive stars currently forming in the nuclei of massive spiral galaxies, such as the Center of the Milky Way (2 $Z_\odot$).

What is the impact of metallicity in massive stars winds? The momentum of the stellar UV photons will be transferred to the metal ions in the stellar atmosphere, launching strong stellar winds which will peel off the outer layers, reducing the amount of available stellar fuel. Thus, mass-loss will modulate evolution, feedback, and the final fate of the star. Recipes for the winds and their metallicity dependence are implemented in all codes of massive stellar evolution. However, theory has been shown to fail at both luminosity and metallicity extremes, including the effect of wind inhomogeneities. Therefore, *quantifying the true mass loss of massive stars during their evolution is a fundamental challenge in the field. CSIC researchers are leading joined international efforts.*

What is the impact of metallicity in massive stars evolution? Observations provide important constraints to the evolutionary models. At low metallicity and high rotation regimes, massive stars are expected to follow chemically homogeneous evolution (CHE) leading to a completely different evolutionary path. Very hot and luminous CHE stars could be responsible for extreme HeII emission in star-forming galaxies and also provide a channel to form ~30 $M_\odot$ double black holes, progenitors of gravitational wave events. However, no CHE star has been observed to date, reflecting the *ill-constrained evolutionary sequence of massive stars beyond the MW-SMC metallicity range.*

What is the role of binaries? A major but necessary challenge is to convolve the evolution of massive stars with their belonging to a binary system. Interaction with a nearby companion can induce mass loss/gain, even stellar mergers, with a deep impact on the evolution of the star. Massive binaries can explain long-standing problems such as the UV-excess detected in star-forming galaxies and the origin of short-GRBs, but new questions arise. Already the first gravitational wave detected, GW150914, evinced the existence of ~30 $M_\odot$ double BHs larger than any model could form. *Much work is left to do both on the observational and the theoretical side to understand the real impact of binarity on the evolution of massive stars.*

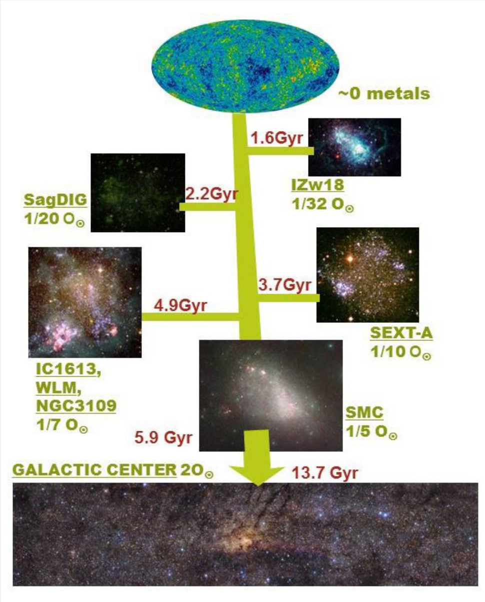


FIGURE 6.5–Nearby star-forming dwarf galaxies and the Milky Way provide a sequence of increasing metallicity that simulates the Universe's growth in chemical complexity with time. We study massive stars in these environments emulating analogs from the early Universe to the present-day.

Strategy: a metallicity ladder to study massive stars at all cosmic epochs

The MW and nearby galaxies make a sequence of metallicity that emulates the chemical evolution of the Universe (see Fig. 6.5). By studying local analogs in multiple metallicity-points we can construct a paradigm that holds at all cosmic epochs. The challenge requires exceptional-quality, multi-epoch ultraviolet, optical, and infrared spectroscopy of larges samples in different galaxies. *On-going and near-future massive spectroscopic surveys with formidable collecting power and multiplexing capabilities will enable, for the first time, multi-epoch access to of entire massive stars populations in the Local Group.* Further, exquisite spatial resolution will break down nearby multiple-systems providing new insight into the hierarchical spatial arrangement of massive stars.

Spectroscopic analysis techniques, perfected for decades now, will yield accurate stellar parameters. However, such an exceptionally large database will need an automatic approach and in the next few years we will embrace *machine learning*. The results of these studies will enable defining the evolutionary sequence of single and binary massive stars. Establishing true mass loss rates requires parallel theoretical developments. Increasing computing power will enable implementing wind models computed from first principles, and relaxing the simplifications historically adopted to save computing time. Nonetheless these advances must be verified by observations. *A multi-wavelength approach, combining data from present and future facilities such as LUVOIR, GTC, VLT, ELT, JWST, SPICA and ALMA, will finally constrain the true mass loss rates of massive stars.*

3.5. The life cycle of low-mass stars

Asteroseismology of low mass stars

In the last decade, asteroseismology has opened a window into stellar interiors and we can, for the first time, poke into their detailed internal structure. Results from the Kepler mission have made ever more evident the inadequacy of state-of-the-art stellar models for providing an accurate description of the dynamical evolution of stellar interiors. What is the size of convective cores during H-core and He-core burning, that determine the lifetime of stars? What are the mechanisms driving transport of angular momentum and chemical species? Dynamical processes are fundamental in the evolution of stars because they determine their chemical structure, the transport of nucleosynthesis products to the surface, and the underlying structure of stars in thermonuclear explosions. Yet, the physical descriptions used in stellar evolution theory today are toy models at best.

It is not computationally possible to model simultaneously the long-term evolution of stars and the dynamical processes in stellar interiors. Instead, the fundamental challenge is to: 1) identify most critical phases of stellar evolution and stellar domains in which multi-dimensional effects play a major role, and most relevant dynamical processes, 2) develop (magneto-)hydrodynamic multidimensional simulations that capture in detail these processes, 3) use those simulations to construct phenomenological models to be used in stellar evolution codes to model their long-term impact in stellar evolution.

Some test cases of this paradigm have been done in the past, and results have had a large impact in stellar physics. Current computational power allows for a much wider range of possibilities. Asteroseismic data from Kepler offers a large constraining and testing power of models and this will only improve in the next decade with PLATO.

Evolved stars and their circumstellar envelopes as present-day dust and heavy elements factories; enrichment of the ISM

During the asymptotic giant branch (AGB) phase, late in their evolution, low-mass stars inject enriched elements into the ISM and are the primary production factories of cosmic dust in the Milky Way. Understanding what cosmic dust is made of, and its cycle since its formation until it is incorporated into planets is an ambitious goal. AGB photospheres gradually expel huge circumstellar envelopes of molecular gas and dust. Grains are thought to form in the innermost layers of these envelopes. Which are the gas-phase precursors, how do condensation nuclei grow, and how these particles do evolve toward the grain compositions and size distribution observed in the ISM is not known.

Observations that spatially resolve the dust formation layers of AGB stars are very challenging. ALMA observations of the dust formation zone in CW Leonis, the closest and richest (more than 60 molecules have been detected, many of them by CSIC researchers) carbon-rich AGB star reveal hundreds of unidentified molecular lines. The spectrum of many molecules containing C, Si, Mg, Fe, Al, Ti, and Ca is not known. Assigning these lines may take years, but they hide the inventory of the molecular seeds of grain formation. A related problem is the formation of polycyclic aromatic hydrocarbons (PAHs). Recent laboratory experiments carried by CSIC groups suggest that their formation in AGB stars, as believed before, may not be efficient. It is still a mystery how these widespread aromatic species are formed. Shedding light on the dust formation processes and on the life-cycle requires input from chemistry, mineralogy, plasma and surface physics. Multi-wavelength observations, models, and lab experiments resembling the conditions in AGB envelopes are needed to investigate the many processes involved.

Planetary Nebulae

Low mass stars will undergo the planetary nebula (PN) phase, before becoming naked white dwarfs. PNe are glowing shells of gas and dust around stars that have just left the AGB phase, and show a spectacular variety of morphologies. During the post-AGB phase, the central star shrinks and raises its temperature until it is hot enough (about 25,000 K) to photoionize the circumstellar envelope formed earlier by the slow wind during the AGB phase. There is broad consensus that non-spherical PNe are created by binary (or multiple) stellar systems, but there is a plethora of models to explain how the PN is shaped. In the case of close binaries, this could happen through the formation of toroidal structures in a common envelope phase during the AGB phase. In wider binaries, jets could break the spherical symmetry of the circumstellar envelope, opening cavities over which the ionization front will proceed when the PN is created. During the PN phase, also, the heating of the central star and the shocks produced by interaction of the fast post-AGB and slow AGB winds lead to the alteration of the chemical composition of the nebula, and the molecular yields to the ISM. Moreover, PNe show spectral lines of elements such as neon and helium, which are otherwise not observable in cooler phases of stellar evolution, and are therefore unique tracers of late phases of AGB evolution and chemical evolution of the Milky Way and galaxies of the Local Group.

Some of the open questions are: (a) When does jet launching start and which is the exact mechanism of jet launching and collimation? (b) Where are the binary central stars? (c) What is the chemical elements and molecular inventory of matter in PNe?

Multifrequency observations with high angular and spectral resolution can determine the morphology, kinematics, energetics of the ejected material. In conjunction with state- of-the-art hydrodynamical simulations and an active search for binary companions and magnetic fields in the central stars of PNe, this is mandatory to assess the shaping mechanism of PNe, and their contribution to the chemical enrichment of the ISM.

Understanding thermonuclear supernovae

Low mass stars end up their lives as white dwarfs and host spectacular thermonuclear explosions if they accrete matter from or even merge with a companion. Explosions can involve only the outer layers of the white dwarf (novae), or can completely obliterate the star producing a thermonuclear supernovae (SNe), or Type Ia SNe. Ashes of these explosions return to the ISM and are major components of galactic chemical evolution.

The apparent homogeneity of Type Ia SNe explosions initially paved the way for the discovery of the accelerated expansion of the Universe. Observations have later on revealed a large degree of heterogeneity, and the most fundamental question remains unanswered. Which is the dominant channel leading to Type Ia explosions? The so-called single degenerate (SD) channel in which a white dwarf accretes matter from a normal companion until reaching the critical mass for explosion, or the double degenerate (DD) in which two white dwarfs merge, or even collide, and then explode?

The characterization of the environment surrounding the star before the explosion can yield very valuable information regarding the SD or DD channel. Radio interferometric observations of Type Ia limited to 25 Mpc from us have been used to constrain stellar mass loss before the explosion, which should be absent in the DD channel, but not in the SD channel. On the other hand, Type Ia produce several radioactive isotopes. Ni in particular drives the evolution of the lightcurve, and emits γ-rays when it decays. Observations in γ-rays have been obtained only for one Type Ia by INTEGRAL, and have allowed a direct determination for the first time of the amount of 56Ni produced in a Type Ia explosion. While thousands of Type Ia lightcurves have been observed, very little is known with certainty about the inner workings of the explosion mechanisms.

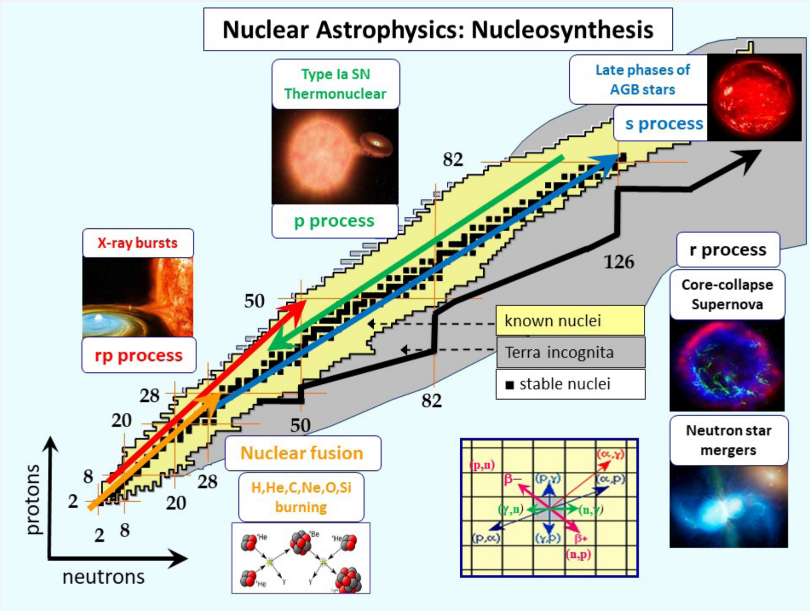
In the future, with the Square Kilometer Array (SKA) in radio, Athena in X-rays, and a possible γ-rays mission (similar to eASTROGAM, previously proposed to ESA), it will be possible to make qualitative progress in our understanding of Type Ia explosions.

3.6. Understanding the role of nuclear reactions and nucleosynthesis in stellar evolution

Nuclear fusion in stars converts light elements into heavier nuclei up to the iron region, where nucleons are maximally bound. Nuclei beyond iron are basically produced by a variety of neutron-capture processes. The particular sequence of nuclear reactions and their rates depend on the mass of the star. the end-products of stellar evolution are typically compact objects, white-dwarfs in the case of low-mass stars, black-holes or neutron stars in the case of massive stars. Additional contributions to the chemical evolution arise from nuclear processes taking place at explosive events on compact stars (Supernovae Ia, Novae) or neutron stars mergers. Nuclear physics is a crucial ingredient for understanding of the evolution and explosion of stars and of the chemical evolution of the Universe.

The aim is to understand the origin of the chemical elements and to answer the fundamental question of where and how the rich variety of present nuclear species has been created from the original composition of hydrogen and helium after the Big Bang. The broad picture exists but our current knowledge of the underlying nuclear physics involved is still incomplete in many basic respects and the unambiguous identification of the astrophysical sites for all of the nucleosynthesis processes is still an open question. Two

FIGURE 6.6–Schematic chart of nuclides showing the various astrophysical processes of relevance for nucleosynthesis, as well as the most likely astrophysical sites where they occur.



outstanding examples are 1) understanding the puzzle of element formation from iron to uranium through neutron reactions and 2) determining the oxygen/carbon ratio at the end of He burning phase in stars that determines later stages of stellar evolution and involves the carbon producing triple alpha process ("the reaction of life") and other light-nuclei reactions [National Research Council, 2003; Nuclear Physics European Collaboration Committee, 2017]. CSIC researchers are well positioned to contribute to both.

Many properties of the nuclei involved in nucleosynthesis processes (s, r, p, rp, and other less common processes, see Fig. 6.6), such as masses, weak-interaction rates (beta- decay, electron capture and neutrino interactions), and nuclear reaction rates (capture of neutrons, protons and alpha particles on a variety of targets), have not yet been determined with enough precision (up to two-orders-of magnitude uncertainty in some cases). These properties also determine to a large extent the different stages of the stellar evolution. These uncertainties are mainly due to the extremely small reaction rates involved in some processes, as well as the difficulty of measuring the properties of very unstable nuclei of relevance in others. A significant push in this direction is expected in the coming years with a new generation of accelerator facilities, generating a wealth of new data from high intensity radioactive-beam and neutron-beam facilities, but also from underground laboratories, stable-beam, and neutrino facilities [Arcones et al., 2017]. Nuclear theory complements the experimental information providing predictions of the properties of key nuclei that have not yet been produced in the lab and connecting experimental measurements under terrestrial conditions with relevant astrophysical quantities at stellar densities and temperatures.

The specific challenges that we will address in the coming 10-20 years, based on their involvement at new/future facilities (like n_TOF/CERN, ISOLDE/CERN and FAIR), are:

Nucleosynthesis in massive stars (see also 6.3.4)

- Understand the rapid neutron-capture process (r process) in core-collapse supernovae and neutron-star mergers and the weak-s process during the core He-burning phase to assess their relative contribution to heavy element formation throughout the history of the Universe. This will require better knowledge of masses, beta-decay half-lives, beta-delayed neutron emission and neutron-capture rates of very neutron-rich unstable nuclei, which are expected to be accessible at the new experimental facilities.

Nucleosynthesis in low-mass stars (see also 6.3.5).

- Understanding light ion fusion reactions at the low energies typical of stars contributing to light element formation. This involves extremely low cross-sections and unstable targets that will require radioactive beams and very low background underground measurements.
- Improving our knowledge on neutron-producing and neutron-capture reactions of nuclei close to stability in particular on some key unstable isotopes, acting on the slow neutron- capture process (s-process) that occurs in AGB stars and contributes to heavy element formation. Both powerful neutron beams and underground labs are required.

Thermonuclear explosions. Novae, supernovae Ia, and X-ray bursts.

- SNe Ia account for 2/3 of cosmic iron and are dominant for many iron-peak elements.

- SNe Ia are the current dominant sources of iron-peak elements (see Fig. 6.2). Novae and X-ray bursts contribute to the synthesis of rare isotopes. Nuclear masses, weak-decay, proton/neutron capture rates of both unstable proton- and neutron-rich nuclei are needed.

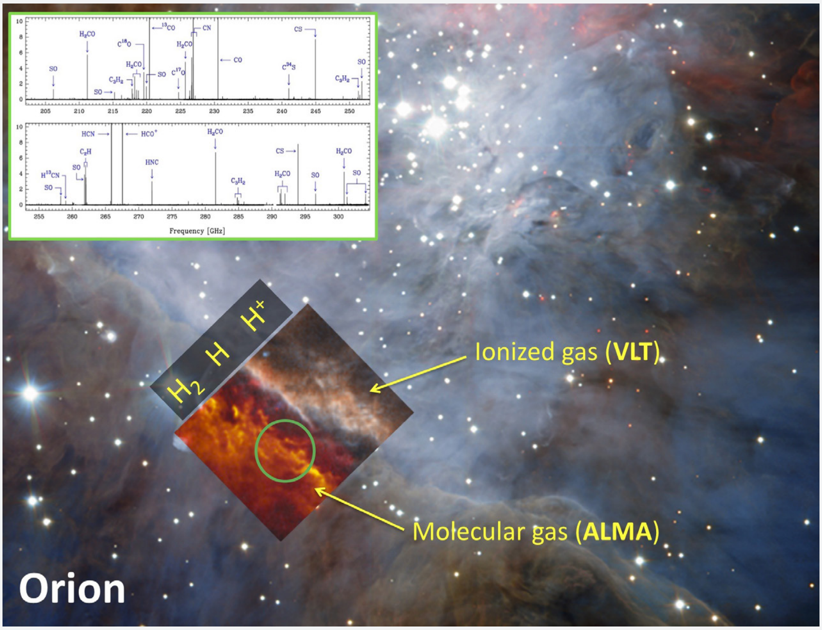
3.7. Characterizing the key physical and chemical processes that drive the evolution of the interstellar and circumstellar media

The multi-wavelength emission from the interstellar and circumstellar media results from a plethora of subtle physical and chemical processes that occur at microscopic level, mostly molecular. The molecular content of many objects of the sky is indeed surprisingly rich and changes with time. Hence, molecular abundances serve as proxies for their evolution. Molecular emission (infrared to radio) probes many dust-obscured environments in which UV to visible light is heavily attenuated (star-formation, galactic nuclei, etc.). Molecules are not only *exotic* species, chemically interesting by their own, they are also powerful diagnostic tools in astrophysics: of physical conditions, magnetic fields, gas kinematics, or ionization rates. They also play a critical role in protostellar gas cooling.

High-precision understanding of the Molecular Universe. The unprecedented detailed information, both spectroscopic and spatial, that the next generation telescopes will provide of so many environments will only be usefully interpreted through implementing new astrophysical models combining magneto-hydrodynamics, thermodynamics, chemical evolution and radiative transfer. These models in turn critically require as input accurate spectroscopic information, cross-sections, and rates for all micro-processes that form, destroy, and excite molecules and atoms in space (chemical reactions and collisions).

Physical conditions change drastically in the evolution from clouds to stars and planets. Ultraviolet photons from massive stars penetrate inside molecular clouds (Fig. 6.7) heating the gas to about 1,000 K, photo-dissociating molecules, exciting polycyclic aromatic hydrocarbons (PAH), and processing dust grains. The details of these microscopic processes are not well understood. This affects our ability to quantify many astrophysical parameters; for example, the lifetime of star-forming clouds exposed to ultraviolet radiation or the origin of the strong PAH emission from very distant starburst galaxies. In the coldest objects inside molecular clouds (10 K in

FIGURE 6.7—Central parsec of the Orion Nebula, illuminated by UV radiation from massive stars in the Trapezium cluster [Goicoechea et. al., 2016]. The inset shows the millimeter-wave spectrum of the molecular cloud edge (taken with the ICTS IRAM 30 m telescope at Pico Veleta, Spain).



prestellar cores) chemical reactions on the icy surfaces of dust grains drive the formation of complex molecules. Surface reactions at such cold temperatures and low pressures, however, are very hard to study.

This challenge aims at articulating a platform between different communities working in low-energy (astro)physics, capable of dealing with the many open problems in molecular astrophysics from an *integral* point of view. It will be mandatory to combine state-of-the-art astronomical observations with *coordinated* computational simulations of molecular processes and laboratory experiments, both able to determine the cross-sections and precise rates of as many as possible gas-phase and surface processes of astrophysical relevance. From the laboratory point of view, the use of synchrotrons and ultra-vacuum chambers for realistic experiments will be needed. However, the tremendous range of conditions probed make these rates hard, sometimes impossible, to measure them in the lab. In these cases, they can only be computed theoretically (by quantum simulations for simple molecules, by quasi-classical methods for complex organic molecules). Such simulations will have to make use of the most advanced high-performance computational methods, simulation techniques, increased CPU time, and big-data analysis tools.

Experimentalists should also improve their spectroscopic techniques to characterize the spectrum of new molecules that could potentially be detected in space. There is still an embarrassing number of unidentified features that populate the spectra of many astronomical objects (hundreds of unassigned “diffuse” interstellar bands in the visible, solid-state bands in the infrared, and many unidentified rotational lines in the radio). These features hide the secret of how carbon chemistry evolves from small hydrocarbon molecules to PAHs and fullerenes, and likely hide the signature of many more prebiotic molecules.

Assigning these features is a great challenge because it involves computing the structure of increasingly complex molecules, minerals, and ice mixtures. The detection of a new molecule in space requires such a line frequency accuracy that only spectroscopy laboratories able to synthesize the species and record its spectrum using the latest technologies (from radio to laser) can provide such precision. Spectroscopists will develop machine-learning tools to assign carriers in very crowed laboratory and astronomical spectra.

Following this roadmap, CSIC researchers in molecular astrophysics will be able to maintain their leading role in the detection of new molecules in space, and also in exploiting their diagnostic power in different environments of the life-cycle of matter. Succeeding in this challenge will allow us to accurately understand, not only how galaxies transform interstellar matter into stars and planets, but how they acquire their physical conditions, abundances, and degree of chemical complexity.

CHAPTER 7

ABSTRACT

Gravity is the fundamental interaction that rules the Universe at large scales but also defines the space-time microstructure. Its very successful current description, Einstein's theory of General Relativity, brings with it many fundamental theoretical and observational problems at all scales: the problem of quantization of matter fields in presence of Gravity (possibly related to another issue, the cosmological constant problem) and of Gravity itself (Quantum Gravity), the prediction of singularities and event horizons in classical solutions, the black-hole information puzzle, the apparent need to include unknown forms of matter and energy to reconcile predictions with observations (dark matter, dark energy), and so on. All this suggests that General Relativity may not be the final theory of Gravity and that we do not yet understand this interaction correctly. Our goal will be to improve our understanding of Gravity from the Planck scale to cosmological scales through the search for answers to those problems using observational, computational and theoretical methods. This is a very timely project: our technology can now provide the required observational tools, and theoretical research on Gravity is flourishing with new ideas and approaches.

KEYWORDS

Gravity | Cosmology | Quantum information | General Relativity | Gravitational waves | Alternatives to General Relativity

CHAPTER 7

GRAVITY

Coordinators
Tomás Ortín
(IFT, CSIC – UAM, Coordinator)
Gianluca Calcagni
(IEM, CSIC, Adjunct Coordinator)

Researchers and Centers
Guillermo Ballesteros (IFT, CSIC – UAM)
Carlos Barceló (IAA, CSIC)
José Luis Fernández Barbón
(IFT, CSIC – UAM)
Óscar García Prada
(ICMAT, CSIC – UAM -UCM – UC3M)
José L. Gómez (IAA, CSIC)
Fernando Marchesano (IFT, CSIC – UAM)
Guillermo Mena (IEM, CSIC)
Gonzalo J. Olmo (IFIC, CSIC – UV)
Carlos F. Sopuerta (ICE, CSIC – IEEC)

1. INTRODUCTION

Gravity is one of the four fundamental interactions of Nature, but its current description (Einstein's General Relativity) is completely different from that of the other three (Quantum Gauge Field Theories). General Relativity provides the space-time framework for the other interactions but we do not yet know how to quantize the latter on curved spaces (the cosmological constant and vacuum energy problems) nor do we know yet how Gravity interacts quantum mechanically with them. On top of this, conventional General Relativity predicts the existence of singularities (mainly within black holes and at the origin and perhaps at the end of the Universe); a prediction that raises many important questions that remain unanswered.

Thus, even though many (but not all) of the predictions of General Relativity have been successfully checked through experiments and observations [Will, 2014], the general consensus is that General Relativity may not be the final description of Gravity. There is a growing body of theoretical and observational evidence showing that we do not yet understand the true nature of Gravity at all scales. It is necessary to keep testing General Relativity and searching for viable alternatives and solutions to the above problems.

This is an exceptional moment in the history of Physics in which our technology can provide us with the tools that we need to make the observations and experiments that will allow us to deepen our knowledge of Gravity and test the current paradigm (General Relativity) and the alternatives. The direct detection of gravitational waves (LIGO, Virgo, Pulsar Timing Array, LISA, Einstein Telescope) and the direct observation of black holes (Event Horizon Telescope) open the door to the study of Gravity in astronomical settings we have never observed before, with very strong gravitational fields or possible cumulative effects beyond Einstein's theory. On the other hand, the advance in quantum control technologies will allow us to test the Equivalence Principle with an unprecedented precision. At intermediate and cosmological scales, the data provided by gravitational wave astronomy will be supplemented by instruments such as the Square Kilometer Array or the Cherenkov Telescope Array and the very large galaxy surveys EUCLID, the Wide Field Infrared Survey Telescope and the Large Synoptic Survey Telescope.

At the same time, there are many new theoretical ideas to study and test, like the surprising connections between Gravity and seemingly unrelated areas of Physics such as thermodynamics, fluid dynamics, Quantum Information theory and Yang-Mills gauge theories.

We will undertake the challenge of understanding Gravity at all scales using all the observational, experimental, numerical and theoretical tools available. In what follows, we concisely review the problems that arise in the study of Gravity at different scales.

The central role played by Gravity in Nature and the broad scope of this enterprise makes it akin to several of the challenges included in Strategic Topic 9. The reader will find many synergies with more than half of them, but we would like to mention specially the challenges 4 "Origin and fate of the Universe", on the theoretical, phenomenological and experimental side, and 8, "New instrumentation and techniques for understanding the Universe, its structure and evolution" on the instrumental side.

1.1. Gravity at the Planck scale

The main problem at this scale is how to combine Gravity and Quantum Mechanics in a consistent way. If Gravity must be quantized, what would the theory of Quantum Gravity be like? A consistent field-theory quantization of General Relativity using the conventional methods has long been known to be at least extremely troublesome and thought to be outright impossible. The quantization of Gravity is a long-standing problem that is fair to call the *Holy Grail of Theoretical Physics*. It has been addressed in many ways. We are going to review below those on which CSIC groups are working. Some of what is known about a hypothetical theory of Quantum Gravity has been learned through the study of quantized fields on classical curved spacetimes and gives rise to many questions and new problems.

On the experimental side, gravitational wave astronomy can test the non-linear and dynamic radiative regime of Gravity, giving us access to the degrees of freedom of the theory, which cannot be studied in any other way. Combining observations in several bands of the spectrum we can also test the very foundations of the current Physics paradigm: Is Lorentz symmetry violated? What is the propagation speed of Gravity? Does the graviton have mass? What is the microscopic structure of spacetime?

Some approaches to a theory of Quantum Gravity

Quantum Field Theory and Gravity A first step towards the reconciliation of Gravity and Quantum Mechanics, simpler than constructing a full-fledged theory of Quantum Gravity, is to study the quantization of fields in a classical curved spacetime [Wald, 1995]. This approach led to the discovery that particle creation occurs in the very early Universe, with very important implications in cosmology, and in the vicinity of black holes (the famous Hawking radiation). The theoretical discovery of Hawking radiation gave rise to the field of black-hole thermodynamics and gave meaning to the Bekenstein–Hawking entropy of black holes, but it also led to the idea of the "evaporation" of black holes with the apparent loss of all its information content (the so-called *black-hole information paradox*) and to the problem of the identification of the statistical (or microscopic) interpretation of the black-hole entropy. The relation between the area of the black hole horizon and its entropy has also inspired the holographic paradigm that we will review below. All these problems and ideas seem to be ultimately related to Quantum Information [Harlow, 2016] with quantum entanglement of states playing the role of the *éminence grise* behind it all.

Another very important problem that arises in this approach is related to the interpretation of the cosmological constant (and dark energy) as the vacuum energy of its matter field content, which, when computed using the Standard Model, gives a number which is off by dozens of orders of magnitude [Weinberg, 1989; Martin, 2012] with respect to its recently measured experimental value, which is very small, but drives the accelerated expansion of the Universe. Either the calculation is affected by some yet unknown mechanism that makes this number small or we do not understand how the vacuum energy couples to Gravity: it is assumed that it obeys the Equivalence Principle but nobody has ever weighted it and checked this assumption (see Section 7.3).

A full-fledged theory of Quantum Gravity is expected to solve all these problems, but, still, important non-perturbative aspects of Quantum Field Theory induced by Gravity remain to be explored and related to current and near-future observations in cosmology and gravitational waves.

Analog Gravity and analog Hawking Radiation. The Hawking radiation of astrophysical mass black holes is too weak to be detectable. However, in 1981 Unruh showed that this prediction of Quantum Field Theory in curved spacetime could be tested in the lab, opening a new field of indirect research on Quantum Gravity.

Steinhauer and his group provided the first experimental evidence of (quantum) analog Hawking radiation from acoustic black holes in Bose–Einstein condensates, toy models for curved spacetime, using a method (density correlator) first proposed by collaboration including CSIC researchers. They could also study the time dependence of the evaporation process. A theoretical understanding of such results requires the study of the backreaction of the Hawking flux on the condensate itself. In the context of Gravity, such an analysis is relevant to address the information loss paradox.

The propagation of surface waves in water provides another interesting toy model. Unruh's and Rousseaux's groups have provided interesting experimental results on stimulated Hawking radiation from white-hole-like flows. The first scattering experiment of surface waves on an analogue black hole flow made by Rousseaux's group opens up the possibility of observing the analogue of the Hawking effect from a black hole horizon in such a system, complementing Steinhauer's experiments on Bose–Einstein condensates.

Modified Gravity and extensions of General Relativity There are several important motivations for searching and studying modifications of General Relativity:

1. If General Relativity cannot be quantized by conventional methods, perhaps a modification of this theory can. It might also happen that the modified Gravity theory solves some of the problems of General Relativity mentioned at the beginning.
2. One should consider the possibility that the anomalous (from the General Relativity point of view) motion of matter in galaxies and of the galaxies in the Universe which are usually attributed to the presence of unknown (*dark*) forms of matter an energy are, instead, a sign that

General Relativity has to be modified. Observe that the dark-matter and dark-energy issues arise at intermediate and cosmological scales.

3. Some modified Gravity theories (especially those with terms of higher order in curvature) could also be regarded as the effective theory of a yet unknown full-fledged theory of Quantum Gravity beyond General Relativity. From a phenomenological perspective it might describe both the phenomena typically associated to the quantum gravitational regime and those proper of astrophysical and cosmological scenarios.
4. A modification of General Relativity (such as *unimodular Gravity*) in which vacuum energy does not couple directly to Gravity may help to solve the cosmological constant problem (page 137).
5. Modifications of Gravity have also been suggested to provide mechanisms for the large energy density fluctuations which are needed to produce primordial black holes (see page 143).

Although General Relativity is manifestly difficult to modify without spoiling its internal consistency and beauty, this is an extremely active field of research. In order to make progress in it, it is essential to determine the number and type of degrees of freedom that describe gravitational interactions within the foreseeable experimentally/observationally accessible regimes. The construction of specific models beyond General Relativity, the analysis of their theoretical consistency, and their confrontation with observations will provide useful information in this direction.

Loop Quantum Gravity If General Relativity cannot be quantized consistently using the conventional methods of perturbative Quantum Field Theory, perhaps other approaches can give a deeper insight. Loop Quantum Gravity is a non-perturbative and background independent program for the quantization of General Relativity based on the Hamiltonian formalism. Its application to cosmological spacetimes has opened a new area of research known as Loop Quantum Cosmology. A relevant part of its foundations have been established by CSIC researchers, who developed the real SU(2) connection formalism used in the formulation of the theory, quantized for the first time a gravitational system with an infinite number of degrees of freedom in the framework of Loop Quantum Cosmology, and introduced the hybrid quantization scheme used to study cosmological perturbations.

Non-local Quantum Gravity Another alternative proposal of Quantum Gravity theory is based on a perturbative field theory of Gravity whose dynamics is governed by weakly non-local operators, kinetic terms with infinitely many derivatives. This promising theory, proposed in the 1980s but developed systematically and rigorously in the 2010s, has non-singular solutions, is unitary and renormalizable and, on a cosmological background, provides a robust justification of successful inflationary models such as Starobinsky's. The phenomenology of non-local Quantum Gravity is under very active study.

Multi-scale Gravity In the great majority of theories of Quantum Gravity, the geometry of spacetime changes with the probed scale. This phenomenon, called dimensional flow, and its observational consequences have been studied both as a universal feature and within an independent proposal, called multi-fractional spacetimes, emphasizing its wide observational impact. From cosmic to particle-physics scales, constraints on the fundamental scales of the geometry have been extracted.

Superstring theory Superstring theory is an internally consistent theory of Quantum Gravity which is not based on standard Quantum Field Theory, where quantum fields get excited and decay by absorption or emission of point-particles. The quanta of Superstring Theory are vibrating one-dimensional objects ("strings"), whose interactions (the joining and splitting of the strings) do not cause the ultraviolet divergences that ruin the quantization of General Relativity (a standard field theory). This single interaction describes in a unified way the four fundamental interactions; all the observed particles being seen as different states of a single object (the string).

Superstring Theory has provided the first microscopic interpretations (including precise calculations) of the black-hole entropy proposed by Bekenstein and Hawking and has also inspired the paradigm of *holography* which can be described as the art of deriving local quantum physics (the framework of both Gravity and Quantum Field Theory) from quantum systems living on boundaries (see page 143). Further- more, its internal consistency has inspired the *swampland program* (see below) on which CSIC researchers are at the forefront of current research.

At low energies, when the small strings can be approximately seen as point-particles, Superstring Theory is effectively described by a modified Gravity theory of the kind described above. In most cases, the leading term is a theory of *supergravity*.

Supergravity Supergravity theories are theories of General Relativity in which Gravity is coupled in a very subtle way to bosonic and fermionic matter fields so that the theory has *supersymmetry*. Supersymmetry is the largest space-time symmetry consistent with the foundations of Quantum Field Theory (and, therefore, of the Standard Model) and it can incorporate in its structure all fields and symmetries. Therefore, it is not surprising that some of them arise as low-energy effective field theories of Superstring Theories and provide the main tools to study (apart from theories beyond the Standard Model) black holes from the superstring point of view.

Some of these theories are, however, very interesting by themselves. For instance, it is possible that the theory known as $N = 8$ supergravity can be quantized consistently in the conventional fashion. The absence of ultraviolet divergences has been proven to five loops in [Bern et al., 2018].

The swampland program A quantum description of Gravity is also essential to have a proper understanding of fundamental physics because the implications of a theory of Quantum Gravity go well beyond gravitational systems. There is a wide consensus that the features of Quantum Gravity also affect other domains in High Energy Physics, and that this could even lead to predictions to be tested in future experiments. The *swampland program* aims to characterize the constraints that a consistent theory of Quantum Gravity such as Superstring Theory places on Quantum Field Theories like those describing the Standard Model of Particle Physics and Cosmology. "The swampland" is the set of Quantum Field Theories which are perfectly consistent by themselves but inconsistent when seen as effective field theories of the type stemming from a theory of Quantum Gravity such as Superstring Theory. In the last decade, there has been a sustained effort in trying to find the structure of the swampland, mapping its boundaries. During this time, the program has quickly gained command of the fundamental understanding of open questions in particle physics and cosmology, ranging from the hierarchy of fundamental scales in Nature, to the origin and final fate of the universe.

A clear example of development within the swampland program is the so-called *Weak Gravity Conjecture*, which states that Gravity must be the weakest fundamental force in any consistent theory of Quantum Gravity. Albeit motivated by the physics of black holes, the Weak Gravity Conjecture has direct implications for many early-universe cosmological models, which are formulated in terms of effective field theories. The same approach can be applied to enhance our understanding of the properties of the Standard Model of Particle Physics. For instance, detailed constraints on the number and character of low-energy species of neutrinos have been derived in this approach.

Holography Gravitational Holography has been a dominating paradigm in the quest for a theory of Quantum Gravity for more than two decades now. It was inspired by the celebrated Bekenstein–Hawking formula for black holes, Entropy=Area/4, which establishes a deep connection between geometry and information theory. The basic idea of holography is to take this formula literally as a measure of the fundamental degrees of freedom in Quantum Gravity. It follows that they must be associated to boundaries rather than volumes of spacetime and that one should be able to deduce local quantum physics from the boundary description.

The most popular and best studied realization of this idea, called *AdS/CFT*, is a string-theoretical duality between a gravitational theory on negatively curved space- times and a non-gravitational Conformal Field Theory living on the boundary at infinity. AdS/CFT is a well-defined mathematical laboratory for holography. In recent years, important progress has been made by importing ideas from Quantum Information Theory. In broad terms, the holographic emergence of the interior space is realized via specific entanglement patterns in the boundary theory. Concepts such as quantum error correction codes and computational complexity are becoming routine notions for the Gravity theorists working in this realm, but the quantum information community is also sensitive to these results as well.

These fascinating ideas represent main research lines in the top Quantum Gravity groups around the world, but one should not forget that, ultimately, we do not live in a spacetime with negative cosmological constant. The extension of these ideas to the flat-space and De Sitter settings are long-term challenges in the field.

1.2. Gravity at astronomical scales

Black holes are quickly becoming central objects of study in astrophysics and cosmology but, are the objects that we observe those predicted by General Relativity? Do they have the same properties?

This fundamental question may be answered by multi-band gravitational wave astronomy. For instance, the gravitational wave spectrum in the ring down phase of the collision of neutron stars or the coalescence of massive black holes can be compared directly with the predictions of General Relativity. The coalescence of binary systems with intermediate or extreme mass ratios will be detectable by LISA in the low-frequency band and in the Hz band in future detectors (the Einstein Telescope project). The gravitational waves emitted and detected will contain a map of the geometry of supermassive black holes and it will be possible to compare the multipole structure that will be extracted from it with, again, the predictions of General Relativity.

The Event Horizon Telescope will also be able to study General Relativity in supermassive black holes through repeated observations of increasing precision. In the near future we will have better ground-based antennae testing General Relativity with increased precision in a neighborhood of the supermassive black hole at the center of the Milky Way, whose mass and position we have measured very precisely. In the long term, antennae in orbit around the Earth (the Millimetron project, for instance) in combination with the Event Horizon Telescope will give us angular resolutions below the micro-arc-second level.

In summary, all these instruments will allow us to test the current General Relativity paradigm with unprecedented precision and in unexplored situations, both from the phenomenological and fundamental points of view. The construction and study of alternative, consistent theories is an absolute necessity in order to make real progress.

We should recall now the question of the existence and nature of dark matter and dark energy discussed in the previous section as a motivation for modified Gravity theories (page 138). It has also been suggested that the existence of primordial black holes could account for some of the dark matter. In any case, the understanding of dark matter requires the understanding of Gravity at these scales.

Black hole alternatives, theory and phenomenology Ultracompact configurations with no horizons, which the current observations cannot discern from General Relativity's black holes, have been proposed as an interesting alternative to the latter. This kind of proposals stimulate progress in the field since they demand improvements in the observations and tests of General Relativity to distinguish between different candidates. For these alternative models to be taken seriously, the formation and stabilization mechanisms of these ultracompact objects have to be studied, all the while paying special attention to the phenomenological characteristics that could make these models verifiable or falsifiable by near-future observations.

Primordial black holes and gravitational waves The old idea that the dark matter of the Universe could be composed by black holes formed in the early Universe by a non-astrophysical mechanism has regained momentum in the last five years, becoming a hot topic in the context of Gravity and cosmology. There are several interesting issues, directly related to our understanding of Gravity, that arise in this context. The following is an incomplete list:

- Hawking and others proposed in the 1970's that black holes may originate from the collapse of large overdense regions of the Universe such as those caused by density fluctuations in the early Universe above a certain thresh- old. Black holes produced by this mechanism would be called *primordial*. Roughly speaking, the correlation length and the size above the threshold of these fluctuations would determine their mass and abundance, but a detailed understanding of the collapse process and the physical and statistical properties of the ensuing black holes is missing. A proper understanding of the formation of primordial black holes would clarify the relevance of these objects for cosmology and it would undoubtedly augment our understanding of Gravity. This challenge requires different expertise, from theoretical cosmology to numerical simulations of Gravity in the deeply non-linear regime.
- A possible source of the required density fluctuations could be primordial inflation. Although the spectrum of inflationary fluctuations inferred from the Cosmic Microwave Background and Large Scale Structure data has a much lower amplitude than required for abundant primordial black hole formation, these observations only allow us to study a small fraction of the inflation needed to solve the horizon and flatness problems. Therefore, in the period of inflation that we cannot observe by those means, much larger fluctuations may have been produced. Several mechanisms have been proposed for the generation of such large perturbations, some of which amount to modifications of Gravity in the early Universe. Thus, primordial black holes could become a unique probe of the inner workings of Gravity at high energies.
- A signature of the formation of primordial black holes from inflation is the generation of an associated stochastic background of gravitational waves. The detection of such a background with LISA and its proper identification as a relic from the formation of primordial black holes after inflation, on top of proving their existence, would have momentous implications for our knowledge of Gravity and, possibly, of dark matter.
- Black hole evaporation through Hawking radiation (page 137) is negligible for astrophysical-mass black holes, but relevant for those with small masses. Primordial black holes lighter than $\sim 10^{15}$ g would have

mostly evaporated already, but the detection of Hawking radiation from heavier primordial black holes evaporating today would open a new window to study Gravity and (perhaps even the information loss paradox) in a currently inaccessible regime.

- Since 2015, LIGO and Virgo have identified gravitational waves emitted in the merging of several binary black holes. There is a chance that some of those that will be detected will be of primordial origin, with huge implications for our exploration of Gravity. The statistics of such mergers could provide us with yet another window for learning about Gravity in the very early Universe.

Visualizing black holes: the Event Horizon Telescope The Event Horizon Telescope is the virtual Earth-size telescope that captured the first image of the "shadow" of a supermassive black hole. The expansion of the Event Horizon Telescope in the next decade (the so-called *next-generation Event Horizon Telescope*) will be capable of making the first real-time movies of supermassive black holes and their emanating jets. These movies will resolve the complex structure and dynamics at the event horizon, bringing into focus not just the persistent strong-field Gravity features predicted by General Relativity, but also details of active accretion and relativistic jet launching that drive galaxy evolution and may even affect large scale structures in the Universe. The next-generation Event Horizon Telescope will turn the extreme environment of their event horizons into laboratories where astronomers, physicists and mathematicians can actively study the black hole boundary in real-time, and with a sensitivity and angular resolution that will allow them to attack longstanding fundamental questions of how Gravity works from completely new directions.

1.3. Gravity at cosmic scales

Two of the main problems in our understanding of the Universe bear a direct relation with Gravity: the existence of dark matter and dark energy. Could a Modified Gravity theory (see page 138) account for them? Is dark energy just the cosmological constant? Are they the vacuum energy? (See page 138). The very large galaxy surveys and the study of the 21 cm absorption lines will in a near future provide us with information about these questions that can be used to test alternative theories. Another, more recent, problem that might be solved by modifications of General Relativity is that of the tension in the value of the Hubble constant H_0, described below (see page 146).

A big unknown in our current theory of the origin of structure in the Universe is that of the value of the ratio between scalar and tensor modes in the primordial fluctuations that gave rise to those structures. Tensor modes are associated with the presence of a gravitational-wave background and its value could validate or falsify the (typically inflationary) models proposed to explain the origin of the inhomogeneities. The imprints of these waves on the Cosmic Microwave Background are being actively searched for by the experiments QUIJOTE and BICEP2, but they may be directly detected by LISA (in which CSIC participates) if that gravitational wave background is large enough (see below). LISA will also be able to discriminate between General Relativity and other models.

Search and characterization of cosmological backgrounds in LISA Since the official approval of the LISA mission in 2017, a big effort has been put in predicting and characterizing stochastic gravitational wave backgrounds from the primordial universe and check its detectability. In particular, the shape of the energy density spectrum of gravitational wave backgrounds from 1.- cosmic inflation, 2.- the non- disturbing processes of particle production (such as post-inflationary overheating), 3.- possible first-order phase transitions in physics beyond the Standard Model, 4.- the formation and evolution of cosmological defects, and in particular of cosmic strings, 5.- several models stemming from Quantum Gravity theories or scenarios, and 6.-the formation of primordial black holes, is being studied by numerical and analytical techniques, as well as the non-Gaussianity and chirality of these backgrounds.

Beyond Einstein's theory in LISA The production and the propagation of astrophysical and primordial gravitational waves are affected by the underlying theory of Gravity. Through the observation of standard sirens (sources of gravitational waves with an optical counterpart) and their luminosity distance, LISA will allow us to detect these effects in models where the cosmic expansion and the dispersion relations do not agree with the predictions of General Relativity. Many models of dark energy or High Energy Physics are currently under scrutiny, stimulating questions of wider scope such as the types of parametrization of the luminosity distance one can use to interpret data correctly.

The Hubble constant tension and modifications of Gravity The measurements of the Hubble parameter H_0 (describing the expansion of the Universe today), that are inferred from the Cosmic Microwave Background and from the study of supernovae and other observations at "small" redshift are very different (in statistical terms, they are separated by approximately "4.5 sigmas"). This is very uncommon in present-day cosmology and may point to the need of a profound revision of our description of the Universe at large spacetime scales. No satisfactory theoretical explanation has emerged so far, but the possibility exists that the problem might be fixed by an adequate modification of General Relativity at the time of the CMB formation. Although it is perfectly conceivable that so far overlooked systematics in the observations could also resolve the current situation, it is fair to assume that in the next few years there will be a surge of activity around this topic, with potentially revolutionary consequences. Understanding if and how a modification of Gravity could address the issue is an ideal playground for testing Gravity at large scales. This line of research could be nicely woven with observational efforts (LISA and Euclid, to give just two examples).

1.4. The mathematics of Gravity

The search for alternative theories of Gravity has stimulated the exploration of new mathematical tools and theories of increasing complexity. The twistor theory, introduced in order to solve Einstein equations has had a great impact in mathematics relating to fundamental problems in complex algebraic geometry and differential geometry, including the study of moduli spaces of bundles, Yang-Mills instantons, monopoles, hyper-Kähler manifolds, etc. The Superstring/Supergravity theory approach to Gravity has also motivated the study of mirror symmetry, Calabi-Yau manifolds, Langlands duality, and created a new mathematical field, sometimes referred as quantum geometry. Conformal Field Theory holographic theorems, complex systems, fractal geometry constrained Hamiltonian systems, gauge theories, non-commutative algebras and non-commutative geometry, generalized functions and the theory of partial differential equations are just some among the many fields touched upon during our quest.

General Relativity, linked from the very beginning with Riemannian geometry, has also motivated a tremendous body of mathematical work, including the recent fundamental work on Kähler-Einstein manifolds and the relation to stability criteria in algebraic geometry. However, General Relativity itself constitutes a challenge at the numerical level when new computing techniques are required to determine, for instance, the wave form of a gravitational-wave source at the precision achieved by near-future interferometers.

2. IMPACT ON BASIC SCIENCE AND POTENTIAL APPLICATIONS

Out of the four fundamental interactions it can be argued that Gravity is the most fundamental one insofar as it governs space and time, the playground for the rest of Physics. Any improvement in our understanding of gravity and any change of paradigm in its description will have a huge impact not just in Physics (Cosmology, Particle Physics etc.) but in all fundamental disciplines (including Philosophy). As we have explained in the Introduction (page 137), there seems to be a deep connection between Gravity and Quantum Information theory which is pushing research in both fields. The consequences of new discoveries concerning this connection simply cannot be foreseen.

On the other hand, it is clear that the most immediate applications of any improvement in our understanding of Gravity will take place in the study of environments in which extreme gravity plays an important role: the early Universe, active galactic nuclei, black- hole mergers, and so on. However, we have mentioned a number of problems in which Gravity plays or may play a crucial role even in non-extreme conditions: the clarification of the nature of dark matter and dark energy (the cosmological constant and the vacuum energy), the nature of event horizons, the luminosity distance of gravitational waves, etc. The technological impact will probably remain mostly indirect for a long time.[3] However, it has to be taken into account that all the experiments in this field (just as in Particle Physics), spatially LIGO-Virgo and LISA, demand technology (software, hardware, mathematical tools) that, often, does not yet exist and has to be developed *ex professo*, acting as powerful drivers of innovation. These cutting-edge technologies will undoubtedly have many spin-offs in its due time.

3. KEY CHALLENGES

Within the overall challenge of achieving a better understanding of Gravity at all scales there are several intermediate challenges which are keys to reach the main goal. We have formulated them as problems and questions in the introduction: the cosmological constant problem, the dark matter and dark energy problems, the black-hole information paradox, the microscopic interpretation of the black-hole entropy, the real existence of event horizons and/or singularities in nature, the existence of primordial black holes, the Hubble parameter tension, the measurement of the tensor-to-scalar ratio, etc. Far from being mere intermediate steps in our progress to the main goal, many of these challenges are extremely hard questions that have occupied for a long time the scientific community working on Gravity. They are, indeed, our long-term key challenges as well.

Nevertheless, it is convenient to formulate challenges that the research groups working on Gravity will face in shorter terms. Among many, we have chosen the following:

- Develop and put to the test a specific model of black hole alternative based on semi- classical and emergent Gravity ideas, as well as on any of the abovementioned proposals for Quantum Gravity or theories beyond General Relativity. Is semiclassical physics able to produce and stabilize ultracompact objects, or does one need additional contributions? Is the singularity resolved? What wave-form would produce the inspiralling and merger of two such objects, and could it be tested with gravitational-wave interferometers?
- To prove or disprove the viability of an emergent theory of Gravity based on condensed- matter systems. It has been shown that many laboratory systems exhibit curved spacetime properties. Is it possible to design a system in which those spacetimes satisfy the Einstein equations?
- Scrutinize the effects that new gravitational physics beyond Einstein's theory could have on self-gravitating systems with well understood composition as a way to minimize the uncertainties that affect compact objects.
- Understanding in detail if and how primordial black holes can form from the col- lapse of large overdense regions in the early Universe and if they can do so abundantly enough to account for the dark matter. Developing the tools for calculating precisely their abundance. Determining the electromagnetic and gravitational tests that may reveal their existence unequivocally or rule them out.
- Clarifying whether the current so-called Hubble tension has anything to do with the properties of Gravity at different distance, time or curvature scales. Determining if gravitational wave observations may help to unravel this tension.
- To determine if metastable De Sitter vacua belong to the swampland or not. In other words, to determine if they are compatible with Superstring Theory's type of Quantum Gravity.

- In the last ten years it has become clear that massive quantum entanglement is key to the emergence of Einstein Gravity in AdS/CFT models. A crucial question is whether other ingredients are needed as well, such as concepts derived from the theory of computational complexity, a notion that so far has remained elusive, despite considerable effort in its elucidation. In other words, is entanglement enough?
- AdS/CFT is the only mathematically well-defined working model, but the real world is definitely not AdS. Rather, at large scale, it is better approximated by a De Sitter spacetime with positive dark energy. The generalization of holography to such cosmological situations is the most important open problem.
- To develop new techniques in Numerical Relativity to access the full non-linear and dynamical regime of Gravity, in General Relativity and in other alternative theories. The goal would be to provide tools to tackle a number of problems: Modeling of sources of gravitational waves in the different bands of the spectrum; dynamics of black holes in different theories and scenarios, applications to physical problems where the AdS/CFT correspondence can play a major role etc.
- To produce the first real-time movies of black holes with an angular resolution capable of discerning any deviation from the Kerr metric near the horizon. In particular, the next-generation Event Horizon Telescope main science objectives for the next decade are aimed to answer some of the most fundamental questions related to Gravity and accretion onto supermassive black holes, namely: Are supermassive black holes described by the Kerr metric? Does General Relativity break down near the event horizon? How do supermassive black holes form and evolve? What is the black hole mass function across cosmic time? Are there horizon-less compact objects? What drives black hole accretion? How do black holes form and power relativistic jets?
- Starting from well-developed formalisms of Quantum Gravity such as String Theory, Supergravity, Loop Quantum Gravity, Non-local Quantum Gravity, Multi-scale Gravity, or any other proposal described above, extract falsifiable predictions for cosmology and for gravitational-wave Physics which can be confronted with observations of ongoing or planned missions that explore the behavior of Gravity at astronomical and cosmic scales.

CHAPTER 8

ABSTRACT

The key science questions to understand the Universe necessitate extremely ambitious research projects, spanning over very long periods of time and requiring large collaborations. Instrumental development in this area enables tool-driven revolutions that can open the door to future discoveries. This is only possible with the appropriate support to innovation. Interdisciplinarity and coordination are fundamental ingredients, as are foresight and planning. Technological advances in instrumentation have found quite often applications in other areas of science and everyday life. All in all, technological activities in this area result in a significant stimulation of R&D and industrial activity. A number of key challenging points are identified, that will pave the way to the implementation of the instruments of the future.

KEYWORDS

Particle accelerators | Telescopes | Radiation detectors | High resolution sensors | Realtime Data handling | Cryogenics | Detector mechanics and services

CHAPTER 8

NEW INSTRUMENTATION AND TECHNIQUES FOR UNDERSTANDING THE UNIVERSE, ITS STRUCTURE AND EVOLUTION

Coordinators
Carlos Lacasta (IFIC, CSIC – UV, Coordinator)
Lourdes Fábrega (ICMAB, CSIC, Adjunct Coordinator)

Researchers and Centers
Agustín Camón (INMA, CSIC – UNIZAR)
Margarita Hernanz (ICE, CSIC – IEEC)
Miguel Mas-Hesse (CAB, CSIC-INTA)
Isabel Márquez (IAA, CSIC)
Giulio Pellegrini (CNM-IMB, CSIC)
Teresa Puig (ICMAB, CSIC)
Olof Tengblad (IEM, CSIC)
Iván Vila (IFCA, CSIC – UC)

1. INTRODUCTION

Challenges in physics come inevitably together with the associated challenges in the development of the appropriate tools for measurement. These tools comprise a wide range of aspects that go beyond sensors and include, among others, readout electronics and data acquisition systems, intelligent data filtering in real time, mechanical structures, control of the environment, etc. Examples are the development of compact accelerators to explore the energy frontier, new optics for telescopes or cryogenic systems.

Developments in the area of instrumentation enable tool-driven revolutions that can open the door to future discoveries. This is only possible with the appropriate support for innovation. This support includes not only funding but also the access to shared infrastructures, the existence of organizational networks, the support and recognition of the workforce engaged in instrumentation R&D activities and structures through which the community can build relationships with industry.

In order to cope with the current challenges in the field of instrumentation development, we need a framework where to harvest the technological seeds for the future with the goal of developing a "knowledge database" with technologies that might be the solutions of these and future challenges.

Projects in this field may last for decades and one of the most ambitious challenges is the management of the focus in the instrumentation activities. There are basically three categories: the now, the next and the horizon. Inevitably, the largest share should be for the short-term commitments, the now, and finding the solutions for the next projects. However, there should always be some effort dedicated to "blue sky research" which is at the basis of the future developments.

2. IMPACT ON BASIC SCIENCE AND POTENTIAL APPLICATIONS

Understanding the basic components of the Universe, its structure and evolution requires the understanding of all the profound connections underlying everything we see, including the smallest and the largest structures of the Universe. Investments have been rewarded recently with discoveries of the tiny masses of neutrinos, the Higgs boson, the first detection of gravitational waves and, very recently, the first image of a black hole. Advances in instrumentation are necessary to enable the pursuit of the main science drivers described in this White Book, so that the observation and measurement of the Universe from the smallest to the largest provide the right answers to the known questions and pave the way to the unknown.

Research projects in this area correspond to the so-called Big Science: science made in large collaborations with very ambitious goals and very long lifetimes. The time interval that covers conception, design and construction of such projects is very often longer than 10 years. This is the reason why the community prepares strategic plans that cover several decades. In Europe we have the European Strategy for Particle Physics [Ushijima et al., 2020], the European Astroparticle Physics Strategy [The European Astroparticle Physics Strategy, 2017], the Long Range Plan in Nuclear Physics [The Long Range Plan 2017 Perspectives in Nuclear Physics, 2017], ESA's *Cosmic Vision 2015-2035* [ESA Cosmic Vision], ESO's *Strategic Plan* [ESO Strategic Plan], the ASTRONET *Roadmap for Ground and Space Astronomy [Roadmap for Ground and Space Astronomy]* and the *Strategy Report on Research Infrastructures* (ESFRI) [Strategy Report on Research Infrastructures (ESFRI), 2018] . All of them have defined the science cases and required infrastructures for the next decades that align the efforts of the different communities and are revised periodically by the community at large. The next two sections will review the science goals and the planned instruments to meet them. The final section is devoted to briefly discuss the applications of the developed technologies to other areas of science and to cope with current societal challenges.

2.1. Particle, Astroparticle and Nuclear Physics

The main drivers in particle and astroparticle physics in the coming years are the use of the Higgs boson as a new tool for discovery, the physics of the neutrino mass and dark matter, and exploring the unknown (see chapter 1).

The High-Luminosity Large Hadron Collider (HL-LHC) at CERN, whose operation is expected to start around 2027, will take the first step towards the complete and precise determination of the Higgs boson properties. The most probable scenario foresees the construction of a Higgs factory to fully complete the study of the Higgs boson right after the HL-LHC. The Higgs factory can be implemented as a linear collider (ILC or CLIC) or as a circular collider (FCC-ee or CEPC). These colliders may be followed by a proton- proton collider reaching an energy of up to 100 TeV that would bring us to the unexplored energy frontier.

Neutrino physics has progressed dramatically in the last decades with a very diverse program that exploits particle astrophysics, accelerator, reactor and underground experiments. The landscape of the future neutrino experiments includes the Deep Underground Neutrino Experiment (DUNE) in the US, the KM3NeT neutrino telescope in the deepest seas of the Mediterranean and the neutrinoless double–beta decay experiment NEXT in the Laboratorio Subterráneo de Canfranc (LSC). While DUNE and KM3NeT will try to study the mass hierarchy of neutrinos, the charge parity violation or the sources of cosmic neutrinos, NEXT tries to determine the nature of this mass. Complementary to those main lines, nuclear physics is improving on the prediction of the anti-neutrino spectrum from reactors, of relevance for neutrino oscillation experiments of new generation like JUNO and TAO.

Dark matter is presumed to consist of one or more kinds of new particles that we are trying to find with direct and indirect detection techniques, observation of large-scale astrophysical effects and production in particle colliders. Direct detection experiments are sensitive to interactions of dark matter with ordinary particles in the laboratory. Example of this is the DAMIC-M experiment using fully depleted CCDs (Charge Coupled Device) to detect very low mass, eV or below, dark matter candidates. Indirect detection experiments, such as the Cherenkov Telescope Array (CTA) gamma-ray observatory, can spot the particle debris from interactions of relic dark matter particles in space.

The forefront of research activity in nuclear physics is to achieve a thorough understanding of the complex structure of nuclei, and the properties of strong-interacting matter under extreme conditions in astrophysical scenarios recreated in the laboratory. In order to address these questions, the scientific community is developing new and more sophisticated tools at the level of accelerators and detectors. The Nuclear Physics European Long Range Plan, issued in 2017, has thus established as main priorities the urgent completion of the ESFRI flagship FAIR (Facility for Antiproton and Ion Research) in Europe, the support for construction, augmentation and exploitation of world leading ISOL (Isotope Separation On-Line) facilities (HIE-ISOLDE, GANIL/SPIRAL2, SPES, JYFL) and the completion of AGATA (Advanced Gamma Tracking Array) in full geometry.

2.2. Astrophysics and Cosmology

Key scientific questions in astronomy, astrophysics and cosmology over the next 20 years will include, among others: fundamental laws (limits of General Relativity, symmetry violations, fundamental constants, short-range forces, etc.), the nature of dark matter and dark energy, the formation of the very first stars and galaxies and their evolution until today, the assembly of the large-scale structures we see in the Universe today, the role of black holes in shaping galaxies, extreme conditions of matter and energy, the formation of stars and planetary systems, and the characterization of extra-solar planets including the search for extraterrestrial life.

All of these questions should be addressed with a coordinated observational and experimental approach, covering all wavelengths of the electromagnetic spectrum with infrastructures on the ground and in space, but also exploring new observing windows to the Universe, like underground neutrino detectors, exotic nuclei facilities or neutron beams providing insight to astrophysical processes or space interferometers for gravitational wave detection. For instance, being able to detect gravitational waves gives us a powerful source of complementary information about the cosmos and the celestial objects it contains, so that this ability has been compared to giving the astronomers a new sense.

Cosmic Vision 2015-2025, ESA's scientific strategic program, with first missions launched very recently, will extend until 2035. It includes three large missions, devoted to the study of giant planets of the solar system and their moons (JUICE, expected launch in 2022), the hot and energetic Universe (Athena, expected launch in 2032) and the observation of gravitational waves (LISA, expected launch in 2034). It also includes 5 medium size missions. Four of them are already approved: Solar Orbiter (dedicated to solar and heliospheric physics, launched in February 2020), EUCLID (to study the evolution of cosmic structures and better understand dark matter and dark energy; expected launch in 2022), PLATO (to search and characterize small transiting planets and perform precision stellar variability studies, expected launch in 2026) and ARIEL (survey of exoplanet atmospheres, expected launch in 2028). Three proposals have been preselected for the fifth and last medium size mission, with final selection expected for 2021 and launch about 2032. They are EnVision (to study the different evolution of Venus and the Earth), SPICA (infrared survey to explore the origin and evolution of galaxies, stars and planets) and Theseus (to monitor transient events in the high-energy Universe across the whole sky and over the whole cosmic history). Finally, there are the newly defined Fast Missions, with the first of them, Comet Interceptor, approved recently. Towards the further future, ESA is presently defining its next science program, *Voyage 2050*.

ESO's present focus is on the construction of the ELT, which will be the largest optical/NIR telescope in the world with six powerful instruments, as well as on defining upgrades for the VLT (Very Large Telescope) and ALMA, which will come next. The European Low Frequency Survey (ELFS), aiming to provide an all-sky CMB polarization map from the ground, is also in definition.

Beyond European context, other space agencies such as NASA, the Japanese JAXA and the Chinese CAS have their own programs, including basically the same science cases. Other large international observatories which are expected to provide unprecedented new science are SKA and CTA, both with strong Spanish involvement, and included in the current ESFRI [Strategy Report on Research Infrastructures (ESFRI), 2018] roadmap.

For Spain and CSIC, the new instruments and facilities in the observatories in the Canary Islands (QUIJOTE, EST, GTC), Calar Alto (CARMENES and its upgrade) and Pico Veleta (NIKA2) also deserve a special mention.

2.3. Applications to other areas of science and societal challenges

The science cases described in this White Book have most often been powerful drivers of breakthrough technological advances that have found application in other areas of science and everyday life. Most popular examples would be the mobile phone cameras, cardiac pumps or spacecraft winglets. Technologies developed for accelerators, telescopes or space contribute therefore to the dynamization of the industrial fabric, through technology transfer and will help to overcome current societal challenges.

Recent advances in instrumentation have led to significant improvements in medical imaging. Fast photodetectors (PhotoMultiplier tubes (PMT) and Silicon PhotoMultipliers (SiPM)) and improvements in readout electronics have made possible the achievement of Time-of-Flight Positron Emission Tomography (TOF-PET), now implemented in commercial scanners with significantly improved image quality. Solid state photodetectors have recently enabled the combination of PET and Magnetic Resonance Imaging techniques, with much higher spatial resolution, soft tissue contrast and dose reduction to the patient.

Novel designs of silicon detectors have also had an impact in the medical imaging field. In hadron therapy the recently developed tissue-equivalent silicon microdosimeters are more appropriate for the characterization of the dose deposited in the tissue than gaseous microdosimeters, since the dose deposition volume matches the cell size. The fabrication of 2D and 3D arrays, together with appropriate readout electronics, in progress by CSIC groups will allow to characterize the dose deposition within the Bragg peak.

Compton cameras are re-emerging in different fields (e.g. localization of radioactive hotspots with drones after the Fukushima catastrophe) and continue to show great promise in medical imaging if the instrumentation requirements can be met. The high resolution and low Doppler Broadening of silicon detectors makes them excellent candidates in imaging and hadron therapy monitoring applications. In addition, the future use of Low Gain Avalanche Detectors (LGAD), with high timing resolution, can improve significantly their performance. Also in the hadron therapy field, compact accelerators are under development employing high gradient radio frequency acceleration which, in addition, finds application in free electron laser technologies.

The research and instrumentation developments carried out in this field have, in most cases, a direct application in other areas, such as “Clean, Safe and Efficient Energy”. Indeed, most neutron capture cross-sections of astrophysical interest are also required for the design of new, safer and cleaner nuclear reactors such as the new generation of accelerator-driven system MYRRHA.

Some of the gamma-ray imaging techniques and instruments that have been developed for astrophysical experiments, have found also industrial applications for the monitoring and control of radioactive waste in the decommissioning of nuclear power plants or in medicine.

The performance of the recently developed cryogenic detectors makes them very interesting for a variety of other applications. Transition-Edge Sensor (TES) sensitivity and spectral resolution find a variety of applications in materials science, nanotechnology, bioanalysis research and industry. TES detectors are in use as high resolution spectrometers in several beam lines and accelerators for materials science, where their efficiency and spectral resolution fill the gap between conventional semiconductor detectors and gratings; TES can also be used in mass spectrometers, and several applications related to security and characterization of nuclear waste. Cryogenic detectors are also in development for quantum communications, because of their single photon counting capabilities and low dark counts.

3. KEY CHALLENGES

Approved instruments and observatories under construction cover most of the science challenges to understand the Universe in the coming decades. They will start operation from now until 2040, and beyond, and will become the new generation of instruments to produce data and provide advances in Big Science in the next two decades. Many efforts are thus devoted to developing and maturing technologies which will make this possible. Work will be focused on these developments and later on the construction, calibration and test of these instruments for the coming years.

In addition, future science goals will require new technologies to meet the challenges of higher performance, reduced costs and new concepts. This demands a strong focus on research and technology. Several technological challenges beyond the horizon can already be identified; the list will certainly increase as the new instruments start to produce data. Listed below are the technology challenges for the planned instruments and identified longer term challenges beyond those. Because of the limited extension, focus is put on challenges where CSIC might have an increased specific weight or can play a leadership role, leaving apart possible global challenges that fall far from available CSIC expertise. Work on these longer term challenges, especially those associated to breakthrough hardware technologies, is of paramount importance to enhance CSIC position in future large infrastructures and consortia and the corresponding transfer to the industrial fabric.

3.1. New generation of sensing elements

Technological challenges depend many times on the specificity of the project but come from the wish of increasing the resolution and contrast of our measurement instruments. This is done, in general by increasing the "magnification" and "range of application" (e.g. energy and luminosity in accelerators, sensitivity in detectors), increasing the granularity of our sensors (resulting in more complex systems, more electronics channels), connecting or merging the information from different instruments and, certainly, developing innovative solutions.

Although "sensors" is a very broad concept and very difficult to categorize, we can distinguish between "environment" sensors measuring forces, electromagnetic fields, temperatures and radiation sensors that include charged particle sensors, aiming to reconstruct the trajectory of charged particles, and photon detectors, that have to cover a very broad spectrum, from microwave to the highest energy gamma rays coming from space or particle accelerators. We list below a number of directions that will require progress in the coming years.

Radiation-hard detectors with improved time and spatial resolution, for future accelerators

Particle physics in future accelerators will need unprecedented resolution on the momentum of the particles, which is affected by the spatial resolution and the amount of material the particle traverses, and will also need enough granularity to resolve the particle jets produced in the decay of the unstable particles, which can contain hundreds of them. On top of that, each collision might yield on the order of a few tens of thousands of particles, with a very high collision rate, that need to be resolved, both in space and time, to identify the topology of the events. This sets very hard requirements to the granularity of the sensors and the time resolution they can achieve.

As a consequence of the huge amount of particles produced, the detectors will have to be radiation tolerant up to fluences of 10^{17} 1 MeV neq/cm^2 and doses of 100 MGy, about 2 orders of magnitude larger that the running experiments at the LHC. They should provide spatial resolutions of 10 μm and provide very precise timing information (on the order of picoseconds).

The detectors should cover very large areas, for instance the trackers of ATLAS and CMS at the HL-LHC will deploy of the order of 200 m^2 of silicon detectors each. Research is needed in the development of services (power, cooling) and mechanical support structures that should not deteriorate the performance nor the physics yield of the detector. Examples are the development of new, very low density yet rigid, materials for the support structures, embedded low mass cooling for sensors and electronics, etc.

Several technologies are being explored each covering some of the aspects required: 3D pixels for the high radiation tolerance and spatial resolution, LGAD and iLGAD to provide accurate timing resolution or Depleted Monolithic Active Pixel Sensors (DMAPS) for an affordable technology covering large areas. The sensor technologies will need validation and solutions for their integration in the detector systems.

Extremely sensitive sensors for gravitational wave detection in space

Detection of gravitational waves in space requires measuring forces of the order of femto Newton in free fall masses inside the satellites. This involves high precision measurements of small environment perturbations (temperature, magnetic field and radiation) that might influence the measurement of free fall masses. For this reason, high precision and high stability sensors are needed. Key challenges for LISA are:

- High precision magnetic field sensors with precisions of 10^{-9} T. Two technologies are in development, based respectively on Anisotropic Magneto-Resistors and on Microelectromechanical Systems.
- Thermal sensors with sensitivities of the order of 1 μK/sqrt(Hz) at 1 mHz and extreme stability, along the long time periods demanded by LISA.

Detectors with sub-eV resolution for cosmology and ultralight dark matter search

In dark matter search and cosmology, low energy particles and rare processes challenge available detectors. The same happens in some astrophysics areas, especially those expecting low count rates and requiring high resolution. For these cases, new detector concepts or readouts play a significant role. For instance, fully depleted skipper CCDs with non- destructive repetitive readout, are in development for achieving a sub-eV noise figure and suppressing dark current. Alternatively, cryogenic detectors, operating below 1 K, are natural candidates to detect ultralight dark matter, since operation at such low temperatures gives access to excitations of the order of the meV (phonons and quasiparticles in superconductors), thus resulting in lower detection thresholds and better energy resolution. Superconducting bolometers could either be used to directly measure the extremely low powers arising from photons generated in axion haloscopes or could be applied to detect dark matter-generated quasiparticles that diffuse into the bolometers. They could also be used for neutrino mass studies such as the double–beta decay.

Low dark count detectors with high efficiency and spectral resolution, for NIR-VIS- UV astronomy

Future NIR-UV observatories in space will require high spectral resolution, large band- width, high optical efficiency and nearly zero count rates to address key science cases such as comparative planetology, cosmology, galaxy evolution and search for biosignatures in exoplanets. Cryogenic detectors mentioned above can provide single photon detection combined with spectroscopic capabilities, low dark count rates and excellent sensitivity, over a wide energy range. TESs, with their high efficiency, and achieved outstanding spectral resolution for soft X-rays, can combine the requirements for these challenges at lower energies, eliminating, for instance, the need of dispersive optics for biosignature search in exoplanets. However, they need further development to improve the spectral resolution or the time response in this energy band, while approaching efficiency values to 100%.

Large array detectors with improved resolution for cosmic microwave background observatories

Future ground based and space observatories to study the Cosmic Microwave Background will be focused on the study of the B-mode signal to detect the primordial gravitational waves. This will require cameras with a large number of detectors ($>10^4$), very good sensitivity and also high contrast in polarization selectivity. Measurement sensitivity is limited by radiation fluctuations coming from system unstabilities and atmosphere; large arrays reduce them. Superconducting Kinetic Inductance Detectors (KIDs) are ideal candidates as they exhibit very good sensitivity and are intrinsically multiplexable: KID arrays have already been implemented for instance in NIKA2 at the IRAM 20 m telescope. However, still key developments are needed to fulfill all the requirements: improvement of the sensitivity and extension of the operation frequency band to lower frequencies (50 to 100 GHz); regarding the detector performance, the study of the polarization response is at a very early stage of development.

3.2. Accelerator technology

Particle accelerators of the future need a number of developments to reach the required energies and luminosities or collision rates. There are many open fronts in this field. Here we only list those where CSIC is already contributing considerably.

Conductors with improved surface resistance for future accelerators

Exploring the unknown in the realm of particles physics will necessitate particle accelerators capable of reaching energies of the order of 100 TeV. Synchrotron radiation will be a serious problem at these energies (35.4 W/m/beam synchrotron radiation emitted by the proton beams at the FCC-hh). This is usually absorbed by a screen layer to isolate the rest of the system. This layer should survive the harsh conditions imposed by the intense synchrotron radiation and the electromagnetic fields present without distorting them and should guarantee the stability of the beam. Current studies estimate that mirror currents with peak values as high as 25 A will be induced with a frequency spectrum in the range of 0 to 1 GHz.

New materials with surface resistance below that of presently used Cu are required; coatings with high temperature superconductors are a most promising technology. They would be able to carry 50 to 100 A in a 1 cm wide tape at 50 K and within a 16 T magnetic field, offering a factor 50 improvement in surface resistance as compared to Cu in conditions very similar to those of the FCC-hh.

High-gradient radio-frequency cavities

High-energy linear colliders with a centre-of-mass energy of the order of a TeV and beyond require high-gradient accelerating cavities. High-frequency acceleration cavities operated at room temperature are being investigated, as this allows for higher acceleration gradients than superconducting cavities. Work needs to be done, however, on the understanding and mitigation of high-voltage breakdown phenomena, caused by the high electric fields on the cavity surface (> 200 MV/m), which limit the gradient of accelerating structures.

3.3. Technologies for high resolution imaging and spectroscopy for exoplanets

Towards the farther future, exoplanet science does require significant research and development of new technology given that the goal is to obtain direct images and spectroscopy of small, terrestrial planets at very small angular separations from the central star. Two main technology areas that would require development are:

- *Technologies for high contrast imaging at optical and near infrared wavelengths*. The main complexity comes from the fact that diffractive 'ghosts' resemble planets, thus greatly reducing the achievable contrast ratio. To address this, adaptive optics systems need to be installed in large space telescopes. No adaptive optics system has been tested in space. The leading concepts under study are LUVOIR and HabEx. There is also interest in developing small compact high-resolution spectrometers with the goal of feeding them via an adaptive optics system in a 40 m class telescope on the ground, which would allow combining high spatial and high spectral resolutions.
- *High precision interferometry from free floating satellites in space*. The goal here is to operate mid-IR interferometers with baselines of tens to hundreds of meters. The free-floating flight formation technologies are being developed for other programs (e.g., LISA), but it remains to be demonstrated to the required precision. Other challenges are to compensate the beams as perfectly as possible to ensure that the interference is of high quality. Operating such interferometer from the surface of the Moon may also be considered, but that would require the construction and development of a lunar base (which is also under consideration, see Chapter 1 of Volume 12).

Exoplanet science is related to the search of extraterrestrial life (Chapter 5 of Volume 12), and thus these technology developments have clear synergies with Chapter 6 of Volume 12.

3.4. Real time data reduction and distributed data acquisition systems

When it comes to data produced by the instruments of the future, a number of challenges will need to be faced: the amount of data produced given the expected event rates and granularity of the sensors, the fact that data producers might be distributed in space requiring data synchronization at unprecedented levels and the fact that the detectors will be operating in very harsh conditions, with very limited access, that will require complex redundant, fault-tolerant and highly redundant systems.

Distributed acquisition of huge amounts of data from many data producers spread over large areas is going to be a challenge in the future. Examples are

KM3NeT and SKA, where the distributed acquisition should be orchestrated and synchronized with very high accuracy. One of the challenges of the future acquisition systems is to increase the bandwidth of transmission from the nodes and, very importantly, the precision of node synchronization (to sub-nanosecond levels), even for distances of hundreds of kilometers. Another example is the acquisition of huge amounts of data with high rates from the future particle detectors: HL-LHC and FCC detectors will necessitate systems capable of sustaining bandwidth of transmission of the order of hundreds of Tb/s with a low latency of about 1 microsecond. On the other hand, SKA will be delivering around 600 PB of public data each year to the community.

The data sets produced by the new, complex systems will be of considerable size and algorithms need to be developed to reduce their size when the data are transferred for permanent storage, to apply calibration corrections, for image stabilization in imaging systems, etc. Examples are the onboard processing algorithms for PHI/Solar Orbiter, and X-IFU/Athena, those in the HL-LHC experiments, or the incorporation of artificial intelligence in data processing and a variety of other tools such as telescope control (focusing, scheduling, guiding) or in quality assessment in collider physics.

This "intelligence" (or algorithms) needs to be executed synchronously to the data acquisition, in real time, and therefore new processing units are required that have to operate in very difficult conditions due to high radiation levels and very restricting constrains in terms of power consumption, compactness and performance. Examples of this are the new generation of Digital Processing Units (DPUs) based on new, powerful and reconfigurable FPGA devices that will allow the execution of the required algorithms.

Furthermore, many experiments will be located in sites that are difficult or impossible to access (deep sea, space, remote regions...) or are operated under harsh conditions (under heavy radiation levels) making any kind a maintenance impossible. It is crucial to design robust, fault-tolerant and highly reliable systems.

3.5. Power supplies

Power supply design for space is an important technological challenge. This is so because of the high level of redundancy required and the need to cope with very strict requirements such as the limited mass, dimensions and power consumption with high efficiency. Other requirements, such as the environmental, radiation and electromagnetic compatibility (EMC) need to be taken into account as well. It is very important to design systems that can work correctly in presence of the perturbations emitted by other systems in the spacecraft and at the same time to ensure that the designed power supply does not affect other systems performance.

3.6. Cryogenic systems

There is an increasing demand of cryo-cooling systems in order to improve sensitivity or capability of instruments. These include large cryogenic systems to cool accelerator magnets or mirrors in gravitational wave detection systems, and also for operation of a variety of sensors and detectors, such as the Liquid Argon TPC (Time Projection Chamber) for DUNE, the Phase Array Feeds of SKA, the high resolution spectrometers of CARMENES or ELT, and the subKelvin detectors in CMB telescopes and space missions such as Athena (X-rays) and SPICA-like (FIR). All these systems, in spite of their specificities, share demands to provide stable environments, reliability, longer life, ever lower power consumption and lower weight (specially for space operation). Two challenges are specifically identified:

- More efficient cooling systems for the Phase Array Feeds of SKA, in order to improve sensitivity and reduce dramatically power consumption.
- Miniaturized on-chip cooling for subKelvin refrigerators: the use of cryogenic detectors in astronomic instruments has made crucial the development of new cooling techniques at temperatures below 1 K. One of the most promising strategies is on-chip cooling, which would allow the development of miniaturized adiabatic demagnetization (ADR) refrigerators. These new ADRs would be much lighter, smaller and with a much lower power consumption than current refrigerating systems, making them most suitable for cryogenic instruments in space.

CHAPTER 9

ABSTRACT

It is a fact that deep new mathematics will be needed to understand some of the new challenges that will shape the Physics of the XXI century. In this chapter we will outline how we aim to develop new mathematics to understand the interplay of analytic, geometric and algebraic properties that arise in the study of the equations of physics, and how emerging aspects of mathematics are shaping our understanding of various aspects of quantum field theory, supergravity and string theory.

KEYWORDS

Differential geometry | Algebraic geometry | Analysis | Partial differential equations | Supergravity | String theory | Supergeometry | Dualities | Quantum computation

CHAPTER 9

NEW DEVELOPMENTS IN GEOMETRY AND ANALYSIS INSPIRED BY THE EQUATIONS OF PHYSICS

Coordinators

Alberto Enciso
(ICMAT, CSIC – UAM – UCM – UC3M, Coordinator)

Fernando Marchesano
(IFT, CSIC – UAM, Adjunct Coordinator)

Researchers and Centers

Emilio Elizalde (ICE, CSIC – IEEC)

Óscar García-Prada
(ICMAT, CSIC – UAM – UCM – UC3M)

Antonio González-Arroyo
(IFT, CSIC – UAM)

María A. Lledó (IFIC, CSIC – UV)

Tomás Ortín (IFT, CSIC – UAM)

Javier Parcet
(ICMAT, CSIC – UAM – UCM – UC3M)

Daniel Peralta-Salas
(ICMAT, CSIC – UAM – UCM – UC3M)

Agustín Sabio-Vera
(IFT, CSIC – UAM)

Germán Sierra (IFT, CSIC – UAM)

1. INTRODUCTION

In this chapter we elaborate on the relation between mathematics and physics, which goes back to the very first days of science. Specifically, our objective is twofold. On the one hand, we aim to develop new mathematics pertaining to the fields of geometry and mathematical analysis, which is essential to face some of the new challenges in physics. On the other hand, we aim to explore a range of applications of cutting-edge mathematics to physics, particularly in formal aspects of quantum field theory, supergravity and string theory, where mathematical concepts and results are oftentimes pivotal to reformulate and solve key problems. In turn, such formal areas frequently stimulate existing research in mathematics, and even motivate the creation of new research lines. This follows the well- established principle that physical problems offer an invaluable guide to develop powerful new mathematics, and that forefront mathematics is essential to make progress in many formal and deep areas of theoretical physics.

Achieving this goal involves a natural and substantial interaction among the different groups whose expertise encompasses many aspects of pure and applied mathematics and of theoretical physics. Indeed, theoretical physicists and mathematicians have learned in the last century that the key to progress in one of these realms is often to be found in the other one. As a result, progress in mathematics and theoretical physics has become intertwined in various instances, with programs in both domains addressing related questions using their own techniques and perspectives, but with a common, final goal.

The main research lines along which the challenge will be addressed are the following:

1.1. Differential and algebraic geometry

Moduli spaces are central objects that appear in the most natural classification problems in geometry. Their importance has been accentuated over the years due to the occurrence of these spaces in such diverse areas of mathematics as algebraic geometry, differential geometry, topology, algebra, and, perhaps more surprisingly, theoretical physics. Some of our main challenges revolve around the study of moduli spaces related to various types of algebraic, geometric and analytic structures.

Many of the moduli spaces considered emerge in the context of gauge theory, in particular Yang-Mills theory and other gauge theories involving connections as well as Higgs fields of various types. The moduli spaces appear first as spaces of equivalence classes of solutions of certain gauge-theoretic partial differential equations (PDEs), and in many cases, in particular when the base manifold is Kähler, these spaces have an algebraic geometric interpretation that links with fundamental problems in algebraic geometry, symplectic geometry and geometric invariant theory. These kind of correspondences is sometimes referred to as Hitchin-Kobayashi correspondences. In many other situations the moduli spaces have a purely topological interpretation. This interplay between differential and algebraic geometry, analysis and topology involved here is extremely rich and has produced many fundamental results in recent years. This is the case of nonabelian

Hodge theory (NAHT), a theory developed by Hitchin, Donaldson, Simpson, Corlette and others, that generalizes to nonabelian Lie groups the theory of Hodge of harmonic forms. On Riemann surfaces this theory links the moduli space of solutions of Hitchin self-duality equations (obtained as dimensional reduction of the instanton equations), to algebraic geometric objects known as Higgs bundles, and to representations of the fundamental group of the surface in a given Lie group G — the G-character variety. This theory plays a fundamental role in the study of integrable systems, hyperkähler geometry, higher Teichmueller theory, the geometric Langlands correspondence and mirror symmetry.

The study of the geometry and topology of the moduli spaces of G-Higgs bundles is an extremely active area of research of great international interest with many challenging problems for the next years to come. Among these one can mention the following. (1) Computation of the cohomology of the moduli spaces of G-Higgs bundles, in particular the computation of Betti and Hodge numbers. This is only for $G = GL(n, \mathbb{C})$ after more than three decades of collective intensive work. But very little is known for other groups. (2) Study of higher Teichmueller spaces and their relation to geometric structures. Higher Teichmueller spaces appear as a generalization for higher rank Lie groups of the usual Teichmüller space of a compact surface, parametrizing complex structures on the surface, which can be identified with the character variety for the group $PSL(2, \mathbb{R})$. Teichmüller theory is central in the study of hyperbolic geometry, Thurston theory of 3-manifolds, and Riemann's moduli space of the surface, which is a central object in algebraic geometry and plays an important role in string theory, in physics. Again, after more than 30 years of investigation there is now a full classification of the higher rank groups for which higher Teichmueller components exist. The challenge is now to study these spaces and to identify the geometric structures they may parametrize. (3) Prove the mirror symmetry correspondence described by Kapustin-Witten between the moduli space of Higgs bundles for a complex reductive Lie group and its Langlands dual. This is related to magnetic- electric duality in physics.

A similar Hitchin-Kobayashi type correspondence for Kähler-Einstein metrics and metrics with constant scalar curvature is described in the Yau-Tian-Donaldson conjecture, relating the existence of such metrics to a certain stability condition in algebraic geometry. This has been recently proved by Chen-Donaldson-Sun for the case of Fano manifolds. The general case remains a fundamental challenge in differential geometry. There are other equations for a metric that emerge, like the constant scalar curvature condition, as the vanishing condition for a certain moment map in an infinite dimensional context. Some of these equations, like the Kähler-Yang-Mills equations, combine the Yang-Mills theory for a connection on a bundle with the Donaldson theory for a metric. It is believed that these various equations relate also to a stability condition still to be discovered. This applies also to other set of equations like the Strominger system. Proving a similar theorem to Yau-Tian-Donaldson correspondence for these other equations is indeed a major challenge.

Other important topics in geometry that are relevant to this challenge and which will be considered too are: symplectic and contact topology, geometric mechanics and control theory, and the theory of dynamical systems.

1.2. Mathematical analysis and partial differential equations

The use of analysis and differential equations to understand fundamental problems of physics harkens back to Newton, who developed calculus to formulate his extraordinary theory of gravity. In this block, a central research topic is harmonic analysis understood in a very wide sense. This includes hard classical problems (Kakeya conjecture, Bochner-Riesz multipliers, restriction of the Fourier transform. . .); noncommutative harmonic analysis (Fourier *Lp* convergence in group von Neumann algebras, non-commutative martingales, connections with geometric group theory ...); harmonic analysis and PDEs (elliptic PDEs in rough domains, connections with geometric measure theory, in- verse problems. . .); or new generalizations of the Calderón-Zygmund theory in nonstandard scenarios. Other topics beyond harmonic analysis include the geometry of Banach lattices, smooth interpolation of data, complex analysis and operator theory.

Harmonic analysis is intimately related to the other central subject of this block, partial differential equations. We are deeply involved in both the development of the theory and its application in fluid mechanics and mathematical physics, and our expertise covers a wide spectrum allowing a multidisciplinary approach to the problems and the promotion of applications to physics. Analytic, asymptotic and numerical techniques, mathematical modeling, and a priori estimates are applied to solve problems in the areas of fluid mechanics, relativity, wave propagation and quantum mechanics.

Our objective is to sharpen these tools in order to face some fundamental problems motivated by physics: what is the origin of turbulence and its connections with chaos theory? How can an incompressible fluid develop singularities? How to estimate the dispersive effects of waves, and their effects for the long-term evolution of the universe through the Einstein equations? What do the nodal sets in spectral geometry look like? How can one extend probabilistic and harmonic analysis methods to the quantum realm?

1.3. Mathematical structures in QFT, supergravity and superstring theory. Supergeometry

Quantum Field Theory (QFT) is the framework to describe elementary particles and their interactions (except possibly gravity). Physicists have set up a methodology that allows calculations in powers of small couplings that in some cases have yielded spectacular success in matching the experimental measurements. However, the basic formulation of QFT becomes loosely defined from the mathematical stand-point. For example, the afore-mentioned expansion is at best asymptotic and in some cases non-Borel summable. The path-integral formulation in its Minkowski version is thought to be merely formal. A very general statement of what a Quantum Field Theory is from that stand-point is a measure in distribution space satisfying certain axions (the Glimm-Jaffe axions) which implies the existence of euclidean Green functions which can be analytically continued to Minkowski space. This produces on the way a connection between the original axions with those of Ostwerwalder-Schrader and eventually those of Wightman. This procedure is too general, and attempts to produce a more explicit approach such as Constructive Field Theory are far from being able to deal with the most relevant examples of these field theories. Radically new ideas are needed. These might come in the case of certain supersymmetric field theories which have milder divergences in the ultraviolet and look like a more feasible goal, like N=4 Super Yang-Mills. Then, non-supersymmetric theories could be approached by deformations of such supersymmetric theories. Other developments affect the possibility of extending the path integrals to complex paths in a kind of infinite dimensional generalizations of steepest descent and the so-called Lefschetz thimbles. This might allow a procedure for resumming the perturbative power series. These ideas are gathered around the notion of Resurgence, on which both mathematicians and physicists are working at present.

Also in relation with the QFT vacua, and stemming from seminal activity by Kontsevich, important work has been carried out at ICE and IFT on the regularization of the contribution of the fluctuations of the quantum vacuum to different kinds of physical processes, by means of analytic methods. In particular, by using the zeta functions corresponding to the relevant physical, pseudo differential operators with different sorts of boundary conditions. Again, this is an extremely beautiful and fruitful interplay between higher mathematics and relevant QFTs. As a separate but related work line, several attempts at the solution of the Riemann conjecture starting from physical models have been pursued. The aim is to try to approach and clarify this very difficult issue from different perspectives.

Supergravity theories can be thought of as extensions of General Relativity in which gravity is coupled in a very subtle way to bosonic and fermionic matter fields. In most instances they arise as effective field theories of superstring theories. They have an intricate mathematical structure that has not yet been completely understood and formalized. This structure encompasses many different kinds of fields (pinors, spinors, p-forms, gauge fields, metric field, ...) and symmetries (gauge, geometrical ...) and dualities which are entangled in a highly non-trivial way that connect traditionally separated realms of mathematical research, especially in differential geometry. The huge amount of symmetry of these theories allows for classifications of their classical solutions much more refined than those which are possible in General Relativity. Many of these solutions are of great interest in physics (black holes, branes, instantons, cosmologies, U-folds...) and also in mathematics since their very existence often depends on delicate relations between the symmetries, dualities and the field content of the theory. Finally, they provide powerful tools for obtaining mathematical results in more conventional theories. Witten's simple proof of the Positive Energy Theorem in General Relativity using supergravity techniques (many details of which were completed later by Taubes) provides a perfect example of what can be achieved in this direction.

The mathematical foundations of supersymmetric, supergravity and superstring theories have a large intersection with the field known as super geometry, which is an area of mathematical research on its own. The main idea of super geometry is to extend classical geometry by allowing for odd or anticommuting coordinates, or more precisely Z2-graded modules and sheaves over Z2-graded commutative algebras. The global object which one then obtains from gluing such extended coordinate systems are supermanifolds. A particularly interesting case of supermanifolds are super Riemann surfaces, which are directly related to superstring theory. Indeed, superstring perturbation theory expresses scattering amplitudes as integrals over the moduli of punctured super Riemann surfaces, which gives a powerful framework for understanding important properties of string theory such as spacetime supersymmetry. Development in this area requires the development of appropriate algebro-geometric foundations and is currently an active area of research for both physicists and mathematicians.

1.4. String theory vacua, D-branes, dualities and effective field theories

Superstring theory mixes very high energy particle physics with vibrant research areas of pure mathematics. This does not only happen through its perturbative formulation, but also via the equations that are derived from this theory. Indeed, solving the equations of motion of string theory, and more precisely finding string theory vacua, has led to one of the richest interactions between theoretical physics and mathematics in past few decades. The most relevant class of vacua for high energy physics consists of a ten-dimensional space-time with four large dimensions (the observable ones) and six small compact dimensions, whose size makes them inaccessible to current experiments. An observer at a present-day experiment would only perceive these six compact dimensions through the four-dimensional low-energy effective field theory (EFT) that results from them, in the sense that such an EFT depends on the topology and geometry of the six compact dimensions. This results in a very deep connection between high energy physics and sophisticated areas in modern mathematics, most notably with algebraic geometry. Indeed, the simplest class of solutions to the vacuum equations specify that the six compact dimensions correspond to a manifold with a Calabi-Yau metric. Algebraic geometric techniques can then be used to build a dictionary between geometric and physical quantities, as well as to construct and classify the different families of vacua. It is via this classification that one of the most striking features of string compactifications arises, an equivalence known as mirror symmetry. Mirror symmetry identifies two Calabi-Yau manifolds which are very different, but when used as extra dimensions in string theory they correspond to the same EFT. From the mathematical viewpoint, this equivalence can be used to solve long-standing problems in enumerative geometry, and as a powerful tool to compute geometric invariants. A standard approach to mirror symmetry is based on the homological mirror symmetry program of Maxim Kontsevich, which from the physics viewpoint can be understood including the objects known as D-branes in the construction of vacua. Such objects are of prime importance in string theory, because of their property to localise gauge theories in submanifolds of the six compact dimensions. From the mathematical viewpoint, their proper definition and general description is yet another active field of research, involving geometric constructions such as coherent sheaves and Higgs bundles.

Mirror symmetry is one of many equivalences that occur between string theory vacua, equivalences known as string dualities. For physicists, the most interesting class are those known as weak-strong coupling dualities, which map physics captured by the perturbative expansion of QFTs to non-perturbative regimes. Exploiting such a duality is at the very heart of the framework known as F-theory, a non-perturbative realization of string theory that is written in the language of algebraic geometry. Standard F-theory solutions are described in terms of Calabi-Yau four-folds, and provides an overall description of the set of string theory vacua within a single, unifying framework. The brane sector contains 7-branes (a generalization of D7-branes) which can be described by Vafa-Witten systems, a particular class of Higgs bundles whose properties are currently under investigation. More general compactifications require leaving the Calabi-Yau realm, and to turn to manifolds with *G*-structure described by generalized complex geometry, another area with multitude of applications and open problems in physics and mathematics. Again, string dualities permit to define even more general compactifications like U-folds. The mathematical description of such compactifications is still being developed, in terms of the field of extended geometry.

An important feature of each family of vacua is the moduli space of solutions that it corresponds to, like for instance the moduli space of Calabi-Yau metrics on a manifold, which is itself a manifold with a certain metric. From the four-dimensional EFT view- point, this moduli space translates into assigning different vacuum expectation values to massless scalar fields, on which the properties of the EFT depend. It has been recently realized that EFTs that are compatible with a quantum theory of gravity like string theory cannot have arbitrary moduli spaces. More precisely, an EFT whose moduli space does not satisfy certain properties it is said to belong to the *Swampland*, which is the set of consistent field theories which, nevertheless, cannot be consistently coupled to quantum gravity, and can never be obtained from a consistent string compactification. In general terms, the Swampland program aims to draw general constraints on EFTs that can be coupled to quantum gravity, and in particular on their moduli spaces. Particular effort has been invested to characterize the physics along moduli space geodesics of infinite distance, by means of the Nilpotent Orbit Theorem of Schmid, other mixed Hodge structures results and further algebro-geometric techniques. This quickly developing field of research has been of high interest for the string theory community in the past few years, and it is a promising direction to unveil new and deep links between fundamental physics and mathematics.

1.4. Quantum Computation

Quantum Simulation and Quantum Computation are based on the developments carried out in Quantum Information theory and Atomic, Molecular and Optics Physics (AMO) in the last two decades. They have led us to the first quantum computers that are still small and noisy, but whose power is expected to increase exponentially in the coming years.

This is an interdisciplinary field where many areas of Physics, Mathematics, Engineering and Computer Sciences converge. On the one hand, mathematics is an essential part of this quest, one of whose goals is to lay the foundations of quantum information. On the other hand, theoretical physics has been contributing by deepening our understanding of fundamental concepts such as entanglement that is the basic resource in quantum computation. Tensor networks is another topic that came from Condensed Matter Physics and is now being actively investigated in Quantum Field Theory and Quantum Gravity. A feature of the exciting period we are living in is the cross fertilization of ideas and techniques between many areas in science. This has been called the second quantum revolution.

2. IMPACT ON BASIC SCIENCE AND POSSIBLE APPLICATIONS

The impact of this topic on fundamental science is huge. Indeed, many of the hardest and most influential challenges in mathematics are directly or indirectly related to the equations of physics. It is no exaggeration to say that these questions have been key to shape mathematics and theoretical/mathematical physics, and no doubt they will continue playing a central role in the evolution of these fields.

This is obvious even from a historical perspective, as the mutual impact between mathematics and physics is impossible to overestimate. Well known examples are Riemannian geometry and general relativity, quantum mechanics and operator algebra, or algebraic, differential and super geometry and string theory.

It should be emphasized that the challenge under consideration is the backbone of modern mathematics. To put things in perspective, note that up to 13 Fields medalists have done some of their best work in these topics, and that thousands of mathematicians and theoretical physicists work on these questions. And at the core of this program are two of the seven celebrated “Millenium problems” proposed by the Clay Mathematics Institute: the mass gap problem in Yang–Mills theory, and the question on finite-time blowup for the 3D Navier–Stokes equation of fluid mechanics.

Although the questions that drive us are purely theoretical, the techniques developed to address these questions may lead to important industrial applications. To mention just a few examples from the last decades: medical imaging algorithms and tomography (elliptic inverse problems, harmonic analysis); methods for financial risk analysis (free-boundary problems); software for the design of cars, ships and airplanes (analysis of the equations of fluid mechanics), cryptography (algebraic geometry).

Last, but not least, the impact and applications of quantum computation on future technology and industry are countless. Building a quantum computer that can be used practically is in itself an outstanding challenge that has become the 'new race to the moon'. Next to researchers and vendors of future computing technologies, national authorities are showing strong interest in bringing the techniques to maturity due to its known potential to break many of today's encryption techniques, which would have significant impact on our society. Finally, it is expected that quantum computing has substantial beneficial impact on many computational disciplines. Overall, quantum computation is expected to generate a revolution on its own.

3. KEY CHALLENGES

A large fraction of the research lines described in the introduction are rooted in two of the famous Millenium Problems of the Clay Mathematics Institute (the regularity problem for the 3D Navier–Stokes equation and the mass gap problem). These can be considered fundamental challenges driving research on mathematics and physics at international level. It is definitely expected that there should be a solution to these challenges in a reasonable period of time.

Several of synergic research lines lie at the intersection of mathematics with quantum field theory and supergravity. Indeed, topological recursions are ubiquitous entities related to enumerative geometry of surfaces and to the asymptotic expansion of integrable systems. Quite remarkably, they also appear when resumming sets of Feynman diagrams in different quantum field theories and supergravity. In the coming years it will be a challenge to investigate the topological properties of these diagrams using a branch of homology theory called rooted maps theory, in particular of the role of genus in these expansions. Another important area of research where physicists and mathematicians can work together is to further understand the deeper connection between cluster algebras and the underlying geometry of scattering amplitudes. Each cluster algebra has an associated polyhedron (associahedron) which characterizes the algebra including its cluster adjacency information. The likely outcome of this interaction will lead to new methods in combinatorics and a better understanding of the calculation of physical observables. In the coming years it will be interesting to apply the cluster-algebraic techniques to better grasp the analytic structure of scattering amplitudes beyond the planar sector. These ideas can be readily applied to gravity via the colour-kinematics duality linking gauge and gravitational scattering amplitudes in supersymmetric theories.

Another example of synergy would be to lay the mathematical foundations of the Swampland Program, by the joint effort of string theorist and geometers. The Swampland Program is currently one of the most active lines of research in theoretical physics and it is of prime interest for the high energy physics community. This is because it may lead to obtaining testable predictions from a theory of quantum gravity like string theory. Progress in this program has so far been made by formulating a series of conjectures that are then tested with explicit theoretical constructions. In many cases like for the Swampland Distance Conjecture, it has been shown that there is oftentimes a beautiful mathematical structure underlying some of these conjectures. In some instances, these structures are related to well-known mathematical results, like Nilpotent Orbit Theorem of Schmid on limiting mixed Hodge structures. In other cases, physicists have developed their own results to lay mathematical ground towards the proof of such conjectures. In any event, there is a widespread consensus that substantial progress in a complicated field like string theory needs to build up on a solid mathematical fundament: something that the Swampland program is lacking at this moment.

CHAPTER 10

ABSTRACT

Neutron stars and black holes are unique objects to study matter at the highest magnetic fields and densities. This chapter focuses on these compact objects from the smallest scales of accretion disks close to a black hole event horizon, or the radiation coming from the neutron star surface and magnetosphere, to the largest ones, such as the interaction with the surrounding medium of black hole jets, tidal disruption events, gamma-ray bursts, and other related transient events. Furthermore, it is in the vicinity of those compact objects that cosmic rays and other particles are accelerated to ultra-high relativistic energies.
The understanding of astrophysical systems of all kinds, from planets to stars, to galaxies and clusters, as well as the Universe as a whole, is linked to the study of the interaction between matter and radiation. In this context compact objects represent unique laboratories to study how matter and radiation behave under such extreme conditions.

KEYWORDS

Supermassive Black Holes | Relativistic jets | Active Galactic Nuclei | Neutron Stars | Pulsars | Magnetars | X-ray Binaries | Tidal Disruption | Events | Supernovae | Gamma-ray Bursts | Cosmic Rays

CHAPTER 10

UNDERSTANDING MATTER AND RADIATION UNDER EXTREME CONDITIONS

Coordinators
Nanda Rea
(ICE, CSIC – IEEC, Coordinator)
Miguel Pérez-Torres
(IAA, CSIC, Adjunct Coordinator)

Researchers and Centers
Antxon Alberdi Odriozola
(IAA, CSIC)
Iván Agudo (IAA, CSIC)
Alberto Castro-Tirado (IAA, CSIC)
Juan José Hernández-Rey
(IFIC, CSIC – UV)
Giovanni Miniutti
(CAB, CSIC – INTA)
Diego F. Torres
(ICE, CSIC – IEEC)
Daniele Viganò
(ICE, CSIC – IEEC)

1. INTRODUCTION

For centuries physicists have studied the interaction between matter and radiation in different regimes, from atomic scales to the largest cosmological scales. To test our theoretical predictions for new physical processes and the state of matter in its most extreme conditions, we have only one possibility: we need to turn to astronomical observations. Understanding matter and radiation under extreme conditions requires a multi-disciplinary effort, combining theory, numerical modeling, multi-messenger observations, and the most advanced facilities. In this Chapter we want to follow-up key questions on neutron stars and black holes:

- How do matter and radiation behave around a black hole?
- What physical processes are at work close to neutron stars?
- What are the characteristics of relativistic jets, winds and explosions?
- What is the relation between compact objects and the transient astronomical sources?
- Where cosmic rays are produced and accelerated?

Compact objects, in particular of neutron stars and black holes, have a varied phenomenology, that is directly related to, e.g., dense matter, stellar evolution, supernovae, galaxy feedback, small-scale and large-scale jets, and to the acceleration of particles up to ultra-high energies. A better understanding of the physical processes operating in the most compact sources will lead to a better knowledge of the most energetic and catastrophic events in the Universe, all of them of a transient nature, such as supernovae and hypernovae explosions, gamma-ray bursts, fast radio bursts, quasar and microquasar jets, tidal disruption events, cosmic-ray production and propagation, neutron stars composition, flares, magnetospheres, and winds.

A typical neutron star might have a radius of about 12 km and a mass of about 1.4 times the solar mass, a density exceeding that of nuclear matter, a magnetic field of the order of $10^8 - 10^{15}$ Gauss, and a rotation period between 1.5 ms and 1000 s, depending on the different type and evolutionary path of its progenitor star. Nearly three thousand neutron stars are known today in different systems and environments: isolated or in binary systems with a large variety of companions, in twins of double neutron star systems, in the Galactic disk, in globular clusters, and in other nearby galaxies. Neutron stars' extreme gravitational, rotational and magnetic energy are the fuel for their large variety of emissions, which encompass all available multi-messenger tracers: electromagnetic waves, cosmic rays, neutrinos, and gravitational waves. Neutron stars are the only laboratories where the most extreme phases of matter can be studied: not only probing extremes of gravity and electromagnetism, but also the strong and weak interactions in regimes that have no hope to be explored on Earth. However, the main difficulty in studying these objects became more and more clear in the past few years, and stands mainly in the realization that they transcend the traditional astrophysical approach and require a multidisciplinary effort that spans from particle and nuclear physics to astrophysics, from experiment to theory, from gravitational waves to the electromagnetic spectrum. Despite decades of theoretical and observational efforts, however, we still have a great deal to learn about neutron stars. We do not know, for example, the equation of state that relates the pressure and density in their interior, how these strong magnetic fields actually form, how they are ordered in a dipolar field component, from which kind of progenitor stars the different neutron star classes are formed, or the exact details of how they interact with the environment forming powerful pulsar wind nebulae, how exactly they accelerate particles until TeV energies, and even how their radio emission is produced. Despite being presumably governed by a single equation of state, the neutron star zoo manifests itself as a puzzling multi-colored class, whose bewildering variety of observational properties is still largely unexplained. Furthermore, it was clear all along that these different classes had to be somehow evolutionary related to cope with the limits of the core collapse supernova rates in our Galaxy, but until a few years ago a physical based evolutionary model had not been developed. We propose here a series of challenges in the field of pulsars to understand their diversity, and their tangled magnetospheres, as well as their relations with the most energetic Universe events such as: Gamma Ray Bursts, Kilonovae, Super Luminous Supernovae and Fast Radio Bursts.

Most, if not all, galaxies with a bulge component in the Universe host a supermassive black hole (SMBH) in their very center, whose mass is correlated with the properties of the galactic bulge, suggesting a link between the central BH and its host galaxy. The details on exactly how such relationship is built over cosmic time are one of the main research topics in galaxy formation and evolution. Most research in this field revolves around the idea of feedback in which energy injection via radiation, outflows, and jets by SMBHs during their phases as Active Galactic Nuclei (AGN) is a key ingredient. While the interaction between the AGN radiative/mechanical output and the surrounding interstellar medium can be studied at relatively large galactic scales (Challenge 5 of this Thematic), a proper understanding of AGN energy injection has to start in the immediate BH vicinity, where most of the action actually takes place. Simple energetic arguments, as well as observed properties, imply that most of the power is released from the innermost few tens of gravitational radii of the accretion flow. Despite significant progress in the past decades, our understanding of the BH accretion phenomenon in this relativistic region of space-time is still far from complete and self-consistent. On a smaller Galactic scale, tens of BH binaries are known to show large outbursts due to accretion of matter from a companion star. These systems are key to study on smaller but closer scales what happens in AGN accretion disks and jets. We propose here a series of representative challenges in the field of Galactic BHs and AGNs with the goal of building a deep and self-consistent picture of the physical processes occurring in the general relativistic space-time close to accreting SMBHs.

Furthermore, extreme environments near compact objects inevitably lead to the acceleration of particles to relativistic energies. In particular, cosmic rays (i.e. protons and nuclei) do not point back to their sources due to the intervening galactic and extragalactic magnetic fields, and thus cannot be localized. It is therefore very difficult to correlate cosmic-ray events with observations obtained at other bands of the electromagnetic spectrum. Moreover, at ultra-high-energies cosmic rays interact with the cosmic microwave background, preventing their detection from very distant astrophysical objects. Despite these drawbacks, cosmic-ray observations have brought about a great deal of information about their energy spectrum and composition and the realization of the existence of ultra- high energy cosmic accelerators. γ-rays, unlike cosmic rays, point back to the sources where the emission was originated; thus, the information can be correlated with other electromagnetic wavelengths and messengers. On the other hand, neutrinos point back to their source. However, their detection is difficult and therefore neutrino observations are at present limited by statistics. In the future, more sensitive experiments will clear the path to neutrino astrophysics, since their little interaction with matter allows to study them in a huge variety of compact objects. Furthermore, currently, the main sources of gravitational waves are black holes and neutron stars, and they have been used to probe the hitherto uncharted territory of the merging of two stellar black holes and two neutron stars.

In this Chapter we present the challenges we want to study concerning galactic and extragalactic compact objects, using both an experimental/observational and theoretical approach.

2. IMPACT ON BASIC SCIENCE AND POSSIBLE APPLICATIONS

The importance of this field has been consistently identified by all recent exercises at the international level, e.g., Astronet, ESA Cosmic Vision, or the USA Decadal Review. The multi-disciplinary reach of the complex physics driving matter and radiation interaction, also reflects the large impact that this research line has from basic science going as far as to induce technological and medical applications.

The densities and magnetic fields close to black holes and neutron stars probe a radically different physical regime than what can be done in terrestrial laboratories, such as accelerators and heavy-ion colliders, which probe low density, high temperature aspects of both strong and weak interactions. Theoretical models cannot yet fully extend to this regime, and therefore our understanding of how matter and radiation interact under extreme conditions is still limited, with the strongest constraints to theory coming mainly from astrophysical observations of neutron stars and black holes, and recently from gravitational wave detections.

Significant investments are dedicated to the study of black holes and neutron stars, to their formation and evolution channels, to their connections with the most extreme explosions in the universe and transient events, to their capacity to produce cosmic rays.

3. KEY CHALLENGES

3.1. Neutron Stars: the strongest magnets in the Universe

Neutron stars are the compact remnants of supernova explosions, and they represent the most extreme case of a stable astrophysical body. We know more than 3000 of them, thanks to their radiation emitted in radio, X-rays and γ-rays. They show a remarkable variety of phenomena. Neutron stars are probably behind a plethora of so far poorly understood astrophysical phenomena, like fast radio bursts or hypernovae, and certainly crucial sources of gravitational waves.

Is there a real boundary between rotation-powered pulsars and magnetars?

Until a couple of decades ago, the population of isolated neutron stars was divided in a number of sub-classes, labelled after their observational properties. As detections went on, we saw these boundaries faded away: First of all, the distinction between a Soft Gamma Repeater and an Anomalous X-rays Pulsar disappeared. In the last decade we have witnessed how rotation-powered pulsars manifested magnetar-like outbursts, and how magnetars can be visible in the radio regime. The fundamental dividing line is in principle the energy budget: In one case, the emission is thought to be powered by rotational energy, in the other case by magnetic energy. How- ever, there are probably many cases in which both contribute. The discovery and follow-up of both multi-wavelength persistent emission and magnetar-like transient behavior will be essential in the next years. We are entering an era that is going to be marked by an astonishing amount of radio-pulsar data, provided by new radio facilities (MeerKAT, SKA), as well as by the upgrade of X-ray programs through the launch of a new generation of satellites.

What can we learn about the supernova rate from pulsar statistics?

The increasing amount of available data on pulsars impose us to tackle a population approach, besides studying exceptional phenomena from a selection of them. Timing and spectral properties are in this sense essential to understand the underlying existing zoo of neutron stars. However, the challenge is that such studies need to consider both the physical evolution and mechanisms, and the observational biases. Historical all-sky surveys are limited and relatively shallow, so that the available sample is somehow inhomogeneous, biased by observational and physical filters, often difficult to assess. One example is the poorly known fundamental properties of both radio and high-energy emission from neutron stars: we still lack an understanding of the geometry of the emitting region, and, therefore, we only have a vague idea of the intrinsic energy emitted by neutron stars. Population synthesis approaches can lead to an estimation of the total number of neutron stars actually present in the Galaxy. The actual number is likely to be about 100 times larger than the number of currently detected neutron stars, and is important because it can constrain the Galactic supernova rate, on which we have empirical evidence from historical records, and theoretical estimates from evolutionary scenarios.

What can we learn from X-ray radiation emitted by young neutron stars?
For several tens of young isolated neutron stars, we detect thermal radiation. Neutron stars are born extremely hot, and they cool down on relatively long timescales, remaining potentially visible in X rays for up to $10^4 - 10^6$ years. The inferred surface heat comes mainly slow release of the heat stored in the star's interior and from the dissipation of magnetospheric currents partially analogous to what happens in the Solar corona. Such currents can be intense especially in magnetically active neutron stars (magnetars). The spectral and timing study of magnetars' sporadic outbursts can shed light on these contributions and on the physical instabilities in the crust and in the magnetosphere that trigger such transient events. These mechanisms are related to the poorly known extreme conditions of the magnetospheric plasma and mechanical properties of neutron star's crust, which is subject to intense magnetic stresses and therefore could periodically fail. A continuous follow-up and characterization of the evolution of the transient radiation of upcoming events will be fundamental to enlarge the sample (currently limited to a couple of dozen of events). Outbursts are currently seen on average once-twice per year, but such frequency, and the quality of exploitable data, are destined to increase with the launch of Athena and eXTP.

How large-scale magnetic fields of up to 10^{15} *G can be produced?* Magnetars' timing period allow us to infer dipolar magnetic field up to 10^{15} G. After what was postulated by the seminal works of Thompson and Duncan in the nineties, little progress has been made to explain how to create an efficient dynamo, able not only to amplify the magnetic field intensity by several orders of magnitude, but also to "order" it to a large scale. In order to answer this, one needs magneto-hydrodynamic (MHD) simulations able to characterize the expected topology and intensity of the magnetic field after the collapse. Magnetic field will be easily amplified at small scales, where dynamo is inherently more efficient, and the differential rotation will be supposedly able to operate an inverse cascade of the magnetic energy, concentrating it in the large-scale dipolar and possibly toroidal field. The post-supernova magnetic configuration is a key element also if one wants to understand the long- term evolution of the magnetic field of isolated neutron stars.

How do pulsars interact with their environment? Cosmic accelerators.

The pulsar wind impact on the environment and is eventually stopped either by the interstellar medium and the reverse shock wave of the supernova remnant shock that earlier generated the pulsar, or by the more dramatic interaction with a close companion in a binary system. These processes lead to the formation of nebulae of different classes. Studying these nebulae is critical for understanding the pulsar complex (supernova remnant - pulsar wind nebula - the pulsar itself), the electrodynamics of the magnetized rotators, how their magnetospheres generate the wind, the acceleration of leptons and hadrons up to very high energies, their energy distribution, and how the latter feedback on the surrounding interstellar medium. Pulsars and their wind nebulae provide the possibility of studying plasma-field interaction processes in conditions far beyond what is achievable in the lab. Connecting nebulae with pulsar properties is also an essentially unexplored territory, and whether and how the magnetospheric emission of the pulsar is indicative of pulsar wind nebular emission is still an open question.

Accreting and transitional millisecond pulsars: the holy grail for accretion studies.

Binary millisecond pulsars are spun up to their extremely short spin period through disk accretion of the mass transferred by a low-mass companion star. When the rate of mass transfer declines at the end of such an Gyr-long, X-ray bright evolutionary phase, the pulsar magnetosphere expands and drives a mass outflow, until it sparks emission as a radio (and gamma-ray) pulsar powered by the rotation of its magnetic field. Millisecond pulsars in binary systems are fossils of such a complex evolutionary history. Understanding how a pulsar is spun up, and how fast it can go, gives immediate constraints on the capability of the neutron star structure to endure the centrifugal pull exerted by its rotation. The possibility of detecting continuous gravitational waves also crucially depends on the maximal spin that can be reached. Finally, since their discovery in 1998, accreting millisecond pulsars (AMXPs) have provided a vast body of observational data on extreme phenomena occurring in and near neutron stars, a wealth of information on the behavior of hot plasma accreting onto a magnetized object plus key insights on the evolution of stars exposed to the high-energy radiation emitted by the neutron star. AMXPs have also been recently linked with the enigmatic behavior of the so-called transitional millisecond pulsars (tMSPs), i.e., AMXPs that occasionally turn on as rotation-powered radio pulsars halting accretion for long periods of time. The latter also bears resemblance with gamma-ray binaries, a special class of binaries where the spectral energy distribution peaks in gamma-rays, and the emission is orbitally variable. A few of these objects are known in the Galaxy, and a common setting for their study together with tMSPs and AMXPs may unveil underlying connections. The study of these systems are key to understand the interaction of matter and radiation during different accretion regimes.

3.2. Black holes and Relativistic Jets

Black holes with masses ranging from a few million up to a few billion solar masses lurk in the centers of most, if not all, major galaxies. Those supermassive black holes accrete matter from their surroundings and convert it in energy in such an efficient way that the nuclei of these galaxies can be detected up to cosmological distances. These Active Galactic Nuclei (AGNs) are the most energetic objects known so far in the Universe, and emit across the entire electromagnetic spectrum, from radio wavelengths to γ-rays. They contribute to shape their surroundings not only via radiation, but also via the ejection of mechanical energy in highly collimated jets and/or wide-angle outflows and winds. The ultimate origin of such an energetic phenomenon resides in the innermost accretion flow close to the central black hole where gravity/space-time distortion is strong. Despite decades of intense study by the astrophysical community, there are still key fundamental questions that remain unanswered, some of which are:

Producing the first real time movies of Black Holes.

As described in detail in Chapter 7, the upgrade of the EHT (i.e., the next generation EHT, ngEHT), thanks to its superb angular resolution and sensitivity, will provide the first dynamic image of a Black Hole, being able to discern any deviation from the Kerr metric near the Event Horizon. The ngEHT will study in great detail the BH-jet connection, specifically the active accretion and relativistic jet launching.

What is the mass and spin of supermassive black holes?

We have a basic understanding of the nature of the innermost accretion flow in AGNs that can be roughly decomposed into the UV-emitting accretion disc and the X-ray emitting corona, plus reprocessing and foregrounds. One of the main challenges is to go from this basic "black-and-white" cartoon to a physically-motivated and self-consistent full-color movie of the innermost accretion flow around black holes. The best way forward is likely represented by the physically-driven modeling of combined X-ray variability and spectral data. The next generation of large collecting area X-ray observatories (such as *Athena* and *eXTP*) will enable us to determine the most crucial system parameters, such as black hole mass and spin. In addition, the spin encodes information on the black hole growth across cosmic time.

How are relativistic jets formed, collimated and accelerated? Are AGN jets launched magnetically?

The process of accretion onto supermassive black holes leads to the formation of powerful relativistic jets. Although the process is still largely unknown, our current understanding relies is that the magnetic field anchored in the disk extracts part of the material that is not "swallowed" by the black hole. Very- long-baseline interferometry (VLBI) observations show that the actual jet formation takes place in the innermost regions of the AGN (and most probably within the inner few Schwarzschild radii), while the collimation and acceleration zone seems to extend up to tens of thousands of Schwarzschild radii. However, it is still unclear whether the jets are powered by extraction of rotational energy from the black hole, from the accretion flow, or from both.

High-sensitivity polarimetric mm-VLBI observations with the EHT will provide unique information on the connection between the SMBH and the relativistic jet, and on the nature of the jet launching mechanism. Namely, these observations will potentially provide the final answer to the crucial questions of where the jets are formed, collimated and accelerated, and whether the AGN jets are launched magnetically, or not.

What is the composition of relativistic jets?

Electron-proton and electron-positron jets will produce differences in the magnetic field structure and in the polarization properties, including circular polarization, that may be probe by polarimetric VLBI observations. The future participation of the SKA will be a key element of the global cm-VLBI array thanks to its deep sensitivity and full-polarization precision. Similarly, high sensitivity VLBI campaigns at (sub-)millimeter wavelengths, including ALMA and IRAM, will probe the actual composition of relativistic jets in AGNs in their innermost regions.

What is/are the mechanism/s for high energy emission in relativistic jets?

There is a wide consensus about the low energy (radio to IR) emission of relativistic jets from supermassive black holes as produced by synchrotron emission from relativistic electrons in the presence of magnetic fields in the jets. The emission mechanism of the higher energy emission, up to very high (TeV) energies, is still hotly debated: inverse Compton up–scattering of low-energy photons vs. synchrotron emission of very high energy protons. The dominance of those emission models is directly related to the locus of the high energy emission, as well as the jet composition (i.e. electron–positron vs. electron–proton). The unprecedented sensitivity and wide field of view of the Cherenkov Telescope Array (CTA) at very high energy gamma- rays is expected to resolve this problem, complemented by multi-spectral-range observations across the electromagnetic spectrum.

Are jets at different scales the same?

It is believed that jets, whether at large scales (e.g. relativistic jets in AGNs) or at small scales (e.g. the relativistic jets in X-ray binaries or the non-relativistic jets in Young Stellar Objects, YSOs,), are ejected through a magneto-centrifugal mechanism. However, the magnetic field, which plays a fundamental role in this mechanism, is still a major unknown. We advocate for a systematic study of synchrotron-dominated jets (as in AGNs and X-ray binaries), and thermal-dominated jets in YSOs (in synergy with challenge 10 of this Thematic). Such study would help to better understand the ejection/collimation mechanism, since synchrotron emission allows to measure and map the magnetic field, while thermal emission allows to determine additional physical parameters (e.g. density and temperature).

3.3. The transient explosive Universe

Why do AGN change their look?

The classification of active galactic nuclei has been unified based on the inclination of the accreting supermassive black hole to the line of sight. In this "unified standard model", AGNs showing extremely broad optical lines are Type 1 AGNs while those showing very narrow lines are Type 2 AGNs. Recently, some AGNs have been shown to change from Type 1 to Type 2, and back again, in just a few years' time (changing-look AGN). This is at odds with the standard model, and challenges our understanding of the whole AGN phenomenon. The CSIC is leading observational efforts, mostly in the radio (JVLA, VLBA, e-MERLIN, EVN) and X-rays (XMM-Newton, Chandra), to test and probe the different plausible mechanisms that can lead to the odd behavior of those AGNs.

Tidal disruption events: A new window into supermassive black hole physics.

Many supermassive black holes remain difficult to study, mainly because they appear to be dormant. However, if a hapless star passes too close to a supermassive black hole, its gravitational field will rip apart the star, producing a tidal disruption event (TDE). In a TDE, roughly half of the star's mass is ejected, whereas the other half is accreted onto the black hole, generating a bright flare that is normally detected at x-ray, ultraviolet, and optical wavelengths. The discovery of tidal disruption events (TDEs) has therefore provided us with a unique opportunity to unveil unique properties of supermassive black holes: TDEs can reveal otherwise dormant super- massive black holes and probe their occupation fraction in galaxies of all types. As TDEs are more common in black holes of 10^6 - 10^8 $M_\odot$, they may be the best way to find intermediate mass black holes. TDEs directly constrain the mass and even the spin of SMBHs, and reveal how accretion can power jets. Understanding the TDE rate and its evolution over cosmic time may solve the mystery of how SMBHs can exist as early as $z > 7$.

Are Quasi-Periodic Eruptions due to binary black-hole interaction?

AGNs show sometimes quasi-periodic eruptions, whose nature and physical interpretation is still unclear, but an exciting possibility is that they could be associated with the presence of a second black hole, which would induce recurrent eruptions via inter- actions with the accretion flow on the orbital period. Although this is not the only possibility, all plausible models can be tested for the two known quasi-periodic eruption-sources with X-ray observations. The CSIC has led the discovery and studies of those sources, which puts it in the best position to exploit the exciting possibility of unveiling a binary black holes in the Universe.

Short-duration gamma-ray bursts and gravitational waves.

The coalescence of compact binaries with at least one neutron star were expected to become sources of gravitational waves. A nice confirmation was provided in 2017, following the detection of GW170817 in the electromagnetic spectrum. Short-duration (< 2s) gamma-ray bursts are thus now known to be the electromagnetic counterparts of some of the gravitational wave sources. One of the most relevant questions to be addressed in the future is whether the merger of a neutron star–neutron star system results in another neutron star, or rather a black hole, and what are the reasons for this?

Core-collapse of massive stars: supernovae and long-duration gamma-ray bursts.

The central engines that power the long-duration (> 2s) gamma-ray bursts were unknown until, back in 1998, there was a firm association of a nearby long-duration gamma-ray bursts with a temporally and spatially coincident core-collapse supernova, which was later confirmed in 2003. Thus, any long gamma-ray burst, if close enough, should be associated with a detectable gravitational wave emission and thus offers a very interesting potential synergy between gamma-ray and gravitational wave detectors. The first joined gravitational wave/long-duration gamma-rayburst/supernova observations in the future, possibly combined also with neutrino detections will prove crucial to unravel the nature of these sources and their explosion mechanism. In short, which massive stars explode as long-duration gamma-ray bursts, and how?

3.4. Cosmic Rays

At present, the topic of High Energy (HECR, $E > 10^{15}$–10^{18} eV) and Ultra High Energy Cosmic Rays (UHECR, $E > 10^{18}$ eV) is full of conundrums. There is no clear scenario explaining the origin of HECRs and UHECRs, their acceleration mechanisms, their flux suppression at the highest energies, their composition and the relationship with other cosmic messengers. Although the progress made during the last decade is impressive, only a partial knowledge of the topic is available. Among the questions yet to be solved, we can mention the following.

Where do HECRs and UHECRs come from and which are the physical sources of HECRs and UHECRS (both galactic and extragalactic)?

Do they come from: Pulsars? Supernova remnants? Gamma-ray bursts? Blazars? Starburst galaxies? Powerful black hole jets in aggregates of galaxies? What is the acceleration mechanism? (Diffusive shock acceleration? Discrete shear acceleration? Others?). If the particle mass of UHECRs increases with energy, as the data from the Pierre Auger Observatory and the Telescope Array seem to indicate, is the decrease in their flux above $5 \cdot 10^{19}$ eV due to the GZK limit or to the closeness of the accelerating process to its maximum energy? Galactic cosmic rays stop being observed at $3 \cdot 10^{18}$ eV, why is this the same energy at which extragalactic cosmic rays start getting heavier with energy? Is there some new undiscovered physical phenomenon that affects the cosmic ray air showers above a certain energy that explains the previous question?

What is powering the observed flux of high-energy cosmic neutrinos and cosmic rays?
The extragalactic gamma-ray flux observed by Fermi and the high-energy cosmic neutrino flux observed by IceCube imply similar energy densities in both messengers. Does this indicate that a large fraction, possibly most, of the energy in the non-thermal universe originates in hadronic processes (much more than previously thought)? Or are there hidden neutrino sources? Is it possible that there exists a new subclass of high-energy gamma-ray sources that produce the high-energy cosmic neutrinos and cosmic rays observed? What could be the nature of such a special class of efficient sources? Are they related to radio galaxies and the frequent occurrence of mergers in those galaxies?

Gamma-ray burts as sources of ultra-high energy cosmic rays and neutrinos.
The detection of very high-energy photons in the TeV range from two gamma-ray burts in 2019 (MAGIC and HESS collaborations) has been one of the major discoveries in the field. Furthermore, gamma-ray bursts are historically addressed among the best candidates of Ultra High Energy Cosmic Rays (UHCR), together with Active Galactic Nuclei. Future multi-messenger campaigns with deeper detector sensitivities will likely further constrain gamma-ray burst progenitor models, clarifying the presence of a jet and its composition, and the relative neutrino/EM energy budgets and the role of gamma-ray bursts as sources of UHCRs.

There is a broad consensus that these questions should be tackled with a multi-messenger approach in which neutrinos will play a paramount role in the near future.

VOLUMEN 9 **REFERENCES**

AAD, G. et al. (2020). Combined measurements of Higgs boson production and decay using up to 80 fb−1 of proton-proton collision data at √s = 13 TeV collected with the ATLAS experiment. Phys.Rev., D101:012002.

ABBOTT, B. P. et al. (2016). Observation of Gravitational Waves from a Binary Black Hole Merger. Phys. Rev. Lett., 116(6):061102.

ABE, K. et al. (2020). Constraint on the Matter-Antimatter Symmetry-Violating Phase in Neutrino Oscillations. Nature, 580:339–344.

AKUTSU, T. et al. (2019). KAGRA: 2.5 Generation Interferometric Gravitational Wave Detector. Nat. Astron., 3(1):35–40.

ALMA PARTNERSHIP; BROGAN, C. L.; PÉREZ, L. M. et al. (2015). The 2014 ALMA Long Baseline Campaign: First Results from High Angular Resolution Observations toward the HL Tau Region. ApJ Letters, 808(1):L3.

AMARO-SEOANE, P. et al. (2017). Laser Interferometer Space Antenna 2.

ANDREASSEN, A.; FROST, W. AND SCHWARTZ, M. D. (2018). Scale-invariant in- stantons and the complete lifetime of the standard model. Phys. Rev. D, 97:056006.

AOKI, S. ET al. (2020). FLAG Review 2019. Eur. Phys. J. C, 80(2):113.

ARCONES, A. et al. (2017). White paper on nuclear astrophysics and low energy nuclear physics Part 1: Nuclear astrophysics. Progress in Particle and Nuclear Physics, 94:1–67.

ATLAS COLLABORATION. (2019). Constraint of the Higgs boson self-coupling from Higgs boson differential production and decay measurements. ATL-PHYS-PUB-2019-009.

BAAK, M. et al. (2014). The global electroweak fit at NNLO and prospects for the LHC and ILC. Eur. Phys. J. C, 74:3046.

BAÑULS, M. C. et al. (2019). Simulating Lattice Gauge Theories within Quantum Technologies.

Basic research needs for dark-matter small projects new initiatives (2019)..

BERN, Z.; CARRASCO, J. J. M. AND JOHANSSON, H. (2008). New Relations for Gauge-Theory Am- plitudes. Phys. Rev. D, 78:085011.

BERN, Z.; CARRASCO, J. J.; CHEN, W. M. et al. (2018). Ultraviolet Properties of N = 8 Supergravity at Five Loops. Phys. Rev. D, 98(8):086021.

BRUNO, M. et al. (2017). QCD Coupling from a Nonperturbative Determination of the Three- Flavour Λ Parameter. Phys. Rev. Lett., 119(10):102001.

BULL, P. et al. (2020). Fundamental Physics with the Square Kilometer Array. Publ. Astron. Soc. Austral., 37:e002.

CARRASCO-GONZÁLEZ, C.; HENNING, T.; CHANDLER, C. J. et al. (2016). The VLA View of the HL Tau Disk: Disk Mass, Grain Evolution, and Early Planet For- mation. ApJ letters, 821(1):L16.

CATANI, S.; DE FLORIAN, D. AND RODRIGO, G. (2012). Space-like (versus time-like) collinear limits in QCD: Is factorization violated? JHEP, 07:026.

CERRI, A. et al. (2019). Report from Working Group 4: Opportunities in Flavour Physics at the HL-LHC and HE-LHC, volume 7: 867–1158.

CMS AND ATLAS COLLABORATIONS. (2012). First observations of a new particle in the search for the Standard Model Higgs boson at the LHC. Physics Letters B, 716.

DAINESE, A. et al. (2019). Report on the Physics at the HL-LHC, and Perspectives for the HE-LHC. CERN-2019-007 Yellow Report.

DAINESE, A. ET al., editors. (2019). Report on the Physics at the HL-LHC,and Perspectives for the HE-LHC, volume 7 of CERN Yellow Reports: Monographs. CERN, Geneva, Switzerland.

DE BLAS, J. et al. (2020). Higgs Boson Studies at Future Particle Colliders. JHEP, 01:139.

DIXON, L J. (2011). Scattering amplitudes: the most perfect microscopic structures in the universe. J. Phys. A, 44:454001.

ELLIS, R. K. et al. (2019). Physics Briefing Book: Input for the European Strategy for Particle Physics Update 2020.

ESA. Cosmic Vision (https://sci.esa.int/web/cosmic-vision/-/ 46510-cosmic-vision).

ESFRI. Strategy Report on Research Infrastructures (ESFRI) (http://roadmap2018.esfri.eu/media/1066/esfri-roadmap-2018.pdf).

ESO. Strategic Plan (https://www.eso.org).

EUROPEAN STRATEGY FOR PARTICLE PHYSICS PREPARATORY Group. (2019). Physics briefing book.

FADIN, V. S.; KOTSKY, M. I. And Fiore, R. (1995). Gluon Reggeization in QCD in the next-to- leading order. Phys. Lett. B, 359:181–188.

FENG, J. L. (2005). Supersymmetry and cosmology. Annals of Physics, 315(1):2–51.

FEYNMAN, R. P. (1963). Quantum theory of gravitation. Acta Phys. Polon., 24:697–722.

GOICOECHEA, J. R.; PETY, J.; CUADRADO, S. et al. (2016). Compression and ablation of the photo-irradiated molecular cloud the Orion Bar. Nature, 537(7619): 207–209.

GUO, F. et al. (2018). Hadronic molecules. Rev. Mod. Phys., 90(1):015004.

HARLOW, D. (2016). Jerusalem Lectures on Black Holes and Quantum Information. Rev. Mod. Phys., 88:015002.

HEINEMANN, B. et al. (2019). Physics briefing book: Input for the european strategy for particle physics update 2020. CERN-ESU-004; arXiv:1910.11775.

HEINEMEYER, S. et al. (2013). Handbook of LHC higgs cross sections: 3. Higgs properties. CERN-2013-004.

http://www.linearcollider.org/.

http://www.nu-fit.org/.

https://clic.cern/.

https://globalfit.astroparticles.es/.

https://indico.cern.ch/event/808335/ contributions/ 3365090/,.

https://lsc-canfranc.es/.

https://next.ific.uv.es/next/.

https://twiki.cern.ch/twiki/bin/view/ LHCPhysics/ LHCHXSWG; https://twiki.cern.ch/twiki/bin/view/ LHCPhysics/LHCTopWG.

JOHNSTONE, D. (2018). Time-Domain Sub-mm Astronomy. Measuring the Accretion Variability of Deeply Embedded Protostars. In Atacama Large-Aperture Submm/mm Telescope (AtLAST): 16.

KAWAMURA, S. et al. (2011). The Japanese space gravitational wave antenna: DECIGO. Class. Quant. Grav., 28:094011.

LEE, A. et al. (2019). LiteBIRD: an all-sky cosmic microwave background probe of inflation. In BAAS, volume 51: 286.

LITIM, D. F. AND MANUEL, C. (2002). Semiclassical transport theory for nonAbelian plasmas. Phys. Rept., 364:451–539.

MAGGIORE, M. et al. (2020). Science Case for the Einstein Telescope. JCAP, 03(03):050.

MANUEL, C. AND TORRES-RINCON, J. M. (2014). Chiral transport equation from the quan- tum Dirac Hamiltonian and the on-shell effective field theory. Phys. Rev. D, 90(7):076007.

MARTIN, J. (2012). Everything You Always Wanted To Know About The Cosmological Constant Problem (But Were Afraid To Ask). Comptes Rendus Physique, 13:566– 665.

NATIONAL RESEARCH COUNCIL (CSIC). (2003). Connecting Quarks with the Cosmos: Eleven Science Questions for the New Century. The National Academies Press, Washington, DC.

NUCLEAR PHYSICS EUROPEAN COLLABORATION COMMITTEE. (2017). Long Range Plan 2017: Perspectives in Nuclear Physics.

PLANCK COLLABORATION (2018). Planck 2018 results. I. Overview and the cosmological legacy of Planck. arXiv e-prints, arXiv:1807.06205.

Roadmap for Ground and Space Astronomy (http://www.astronet-eu. org).

ROBSON, I. et al. (2013). A science vision for european astronomy update. Scientific Reports.

SETO, N.; KAWAMURA, S. AND NAKAMURA, T. (2001). Possibility of direct measurement of the acceleration of the universe using 0.1-Hz band laser interferometer gravitational wave antenna in space. Phys. Rev. Lett., 87:221103.

TANABASHI, M et al. (2018). Review of Particle Physics. Phys.Rev., D98:030001.

THE EUROPEAN STRATEGY FOR PARTICLE PHYSICS. https:// europeanstrategy. cern/home. . The Physics Briefing Book (http://cds.cern.ch/record/ 2691414).

THE LONG RANGE PLAN (2017). Perspectives in Nuclear Physics (http://www. nupecc.org/lrp2016/Documents/lrp2017.pdf).

THE EUROPEAN ASTROPARTICLE PHYSICS STRATEGY. (https: //www.appec. org/wp-content/uploads/2017/08/ APPEC-Strategy-Book-Proof-23-Nov-2.pdf).

TOLOS, L. AND FABBIETTI, L. (2020). Strangeness in nuclei and neutron stars. Progress in Particle and Nuclear Physics, 112:103770.

USHIJIMA I.; OHMAE N.; TAKAMOTO, M. et al. (2020). Test of general relativity by a pair of transportable optical lattice clocks. Nat. Photonics.

WALD, R. M. (1995). Quantum Field Theory in Curved Space-Time and Black Hole Thermodynamics. Chicago Lectures in Physics. University of Chicago Press, Chicago, IL.

WATTS, A. L. Et al. (2016). Colloquium : Measuring the neutron star equation of state using x-ray timing. Rev. Mod. Phys., 88(2):021001.

WEINBERG, S. (1989). The Cosmological Constant Problem. Rev. Mod. Phys., 61:1–23.
WILL, C. M. (2014). The Confrontation between General Relativity and Experiment. Living Rev. Rel., 17:4

ANNEX

MISSIONS AND FACILITIES

SPACE MISSIONS

Ariel

ARIEL (Atmospheric Remote-Sensing Infrared Exoplanet Large-survey) is ESA's fourth medium-class mission, with launch estimated 2028. ARIEL is a dedicated survey mission capable of observing a large, diverse and well-defined sample of exoplanets around a range of stellar types. It is designed to perform high-accuracy transit, eclipse, and phase- curve observations employing simultaneous multiband photometry in visible wavelengths and spectroscopy in near infrared wavelengths. Its payload comprises a 1-metre class, three-mirror telescope, an infrared spectrometer, and a Fine Guidance System module providing three narrow-band photometry channels (two used as guidance sensors as well as for science) and a low-resolution near-infrared spectrometer.

Athena

Athena (Advanced Telescope for High-ENergy Astrophysics) is the second Large-class mission of ESA's Cosmic Vision program, expected to launch in 2031. By combining a large X-ray telescope with state-of-the-art scientific instruments, Athena has three key goals: to map and study large-scale gas structures in the Universe, to survey supermassive black holes, and to explore high-energy astrophysical events such as supernova explosions and energetic stellar flares.

The mirror, featuring Si-pore optics, a new low-mass technology developed by ESA, is an essential part of the telescope. Two instruments will be onboard Athena: the X-IFU (X-ray Integral Field Unit) will be an advanced, actively shielded X-ray spectrometer for high-spectral resolution imaging, constituted by superconducting microcalorimeters (TESs). The WFI (Wide Field Imager) is a silicon depleted p-channel field effect transistor (DEPFET) active pixel sensor camera. It has a large field of view, high count-rate capability and moderate resolution spectroscopic capability.

Bepi Colombo

BepiColombo is Europe's first mission to Mercury. Launched on 20 October 2018, it is on a seven year journey to the smallest and least explored terrestrial planet in our Solar System. When it arrives at Mercury in late 2025, it will endure temperatures in excess of 350 C and gather data during its one-year nominal mission, with a possible one-year extension. The mission comprises two spacecrafts: the Mercury Planetary Orbiter (MPO) and the Mercury Magnetospheric Orbiter (MMO), housing a total of 11 instruments.

CHEOPS

CHEOPS (CHaracterising ExOPlanet Satellite), launched in November 2019, is the first mission dedicated to searching for exoplanetary transits by performing ultra-high precision photometry on bright stars already known to host planets. The mission's main science goals are to measure the bulk density of super-Earths and Neptunes orbiting bright stars and to provide suitable targets for future in-depth characterization studies of exoplanets in these mass and size ranges. With an accurate determination of masses and radii for an unprecedented sample of planets, CHEOPS will set new constraints on the structure and hence on the formation and evolution of planets in the sub-Saturn mass range.

CHEOPS is a precision photometry instrument operating in white light at visual wavelengths; it has the capability to provide precise, time-differential photometric time series at high cadence (1 minute) of a wide range of variable light sources on timescales of ≤ 2 days.

Comet Interceptor

Comet Interceptor has been selected as ESA's new fast-class mission in its Cosmic Vision 2015-2025 Program; it is foreseen for launch as co-passenger with ARIEL spacecraft in 2028. It will be the first to visit a truly pristine comet or another interstellar object starting its journey into the inner Solar System, thus containing material that has not undergone much processing since the dawn of the Sun and planets; it will target the comet, making a flyby of the chosen target when it is on the approach to Earth's orbit. Its three spacecrafts include a total of 9 instruments which will perform simultaneous observations from multiple points around the comet, creating a 3D profile of a 'dynamically new' object that contains unprocessed material surviving from the dawn of the Solar System. CSIC participates in 4 instruments: CoCa (Comet Camera), a high resolution camera in the NIR-VIS, the mass spectrometer MANIaC (Mass Analyzer for Neutrals and Ions at Comets), the multispectral visible camera with polarimetric filters EnVisS (Entire Visible Sky coma mapper) and the multiwavelength camera OPIC (Optical Imager for Comets), to cartography the nucleus and dust jets.

EnVision

EnVision, one of the 3 selected candidates by ESA for M5 (final selection planned in 2021, launch expected for 2032), is a planned joint mission with NASA, to explore Venus. It aims at getting insights on why Earth and Venus evolved so differently. To this end it will determine the nature and current state of geological activity on Venus and its relationship with the atmosphere, to better understand the different evolutionary pathways of the two planets.

EUCLID

Euclid is the second M-class mission of the ESA Cosmic Vision Program, expected to be launched in 2022. Its goal is to map the geometry of the Universe and better understand the mysterious dark matter and dark energy, which make up most of the energy budget of the cosmos. The mission will investigate the distance-redshift relationship and the evolution of cosmic structures by measuring shapes and redshifts of galaxies and clusters of galaxies out to redshifts ~2.

The Euclid telescope is a 1.2 m on-axis 3-mirror Korsch cold telescope providing a field of view of 1.25×0.727 deg^2. The payload comprises two instruments, VIS and NISP, sharing a large common field of view. VIS provides high quality images to carry out the weak lensing galaxy shear measurements. NISP performs imaging photometry to provide near-infrared photometric measurements for photometric redshifts, and also carries out slitless spectroscopy to obtain spectroscopic redshifts.

eXTP

The enhanced X-ray Timing and Polarimetry mission (eXTP) is a China-led mission, to be launched in 2027, designed to study the state of matter under extreme conditions of density, gravity and magnetism. Primary goals are the determination of the equation of state of matter at supra-nuclear density, the measurement of QED effects in highly magnetized stars, and the study of accretion in the strong-field regime of gravity. Primary targets include isolated and binary neutron stars, strong magnetic field systems like magnetars, and stellar-mass and supermassive black holes. It will include 4 instruments.

JUICE

JUpiter ICy moons Explorer is the first large-class mission in ESA's Cosmic Vision 2015- 2025 program, planned for launch in 2022 and arrival at Jupiter in 2029. It will spend at least three years making detailed observations of the giant gaseous planet Jupiter and three of its largest moons, Ganymede, Callisto and Europa. JUICE will address two themes of ESA's Cosmic Vision program: What are the conditions for planet formation and emergence of life? and How does the Solar System work?

The JUICE spacecraft will carry the most powerful remote sensing, geophysical, and in situ payload complement ever flown to the outer Solar System. The payload consists of 10 state-of-the-art instruments plus one experiment that uses the spacecraft telecommunication system with ground-based instruments. Two of them are the remote sensing imaging instrument JANUS and the laser altimeter GALA.

LISA

The Laser Interferometer Space Antenna (LISA), selected to be ESA's third large-class mission and expected to launch in 2034, will be the first space-based gravitational wave observatory. It will be able to scour the entire Universe in its hunt for these elusive waves, from the very tiniest to the very largest of scales. To detect gravitational waves with lower frequencies down to 0.1 mHz, such as those from the merging of supermassive black holes at the center of massive galaxies, an observatory must span millions of kilometers – something that can only be achieved in space.

LISA will consist of three spacecraft separated by 2.5 million km in a triangular formation, following Earth in its orbit around the Sun. Launch is expected in 2034.

The science goals of LISA include observing the growth and merger history of massive black holes throughout the Universe, understanding the dynamics and characteristics of the unexplored regions surrounding black holes, precisely measuring tens of thousands of compact binary star systems within our Galaxy, and probing the 'dark' and early Universe by detecting primordial gravitational waves produced at the time of inflation.

LITEBIRD

LiteBIRD (Lite (Light) satellite for the studies of B-mode polarization and Inflation from cosmic background Radiation Detection) is a JAXA mission with participation of NASA and Europe, expected to launch in 2027. It aims to detect the footprint of the primordial gravitational wave on the CMB in a form of polarization pattern called B-mode. It is therefore the next space mission to study CMB, after Planck. The detector is constituted by an array with more than 3000 cryogenic bolometers (TESs).

PLATO

PLATO (PLAnetary Transits and Oscillations of stars) is the third medium-class mission in ESA's Cosmic Vision program, with expected launch in 2027. It is a wide field, high precision photometric instrument to search for and characterize small transiting planets, and to perform precision stellar oscillations studies. It will obtain very high precision photometry of more than 250,000 of stars (<50 ppm for the 15,000 stars in the core sample). PLATO's payload module provides a wide field-of-view (FoV) to maximize the number of the sparsely distributed bright stars in the sky within one pointing, and allows the satellite to cover a large part of the sky. In addition, it provides the required photometric accuracy to detect Earth-sized planets in the habitable zone of Sun-like stars, together with a high dynamic range. Its performance is achieved by a multi-telescope instrument concept, which is novel for a space telescope: 24 'normal' cameras with CCD-based focal planes, operating in white light, will be read out with a cadence of 25 s and will monitor stars with mV > 8; 2 additional 'fast' cameras (with a blue and a red filter) will monitor brighter stars with a cadence of 2.5 seconds.

Solar Orbiter

Solar Orbiter (the first medium-class mission of ESA's Cosmic Vision 2015-2025 Program, launched February 2020) is dedicated to solar and heliospheric physics. Solar Orbiter aims to make significant breakthroughs in our understanding both of how the inner heliosphere works, and of the effects of solar activity on it. It will operate both in and out of the ecliptic plane. Solar Orbiter will measure solar wind plasma, fields, waves and energetic particles close enough to the Sun to ensure that they are still relatively pristine. Solar Orbiter includes 4 in-situ instruments plus 6 remote sensing instruments. One of them is PHI, The Polarimetric and Helioseismic Imager: it will provide high-resolution and full-disc measurements of the photospheric vector magnetic field and line-of-sight (LOS) velocity as well as the continuum intensity in the visible wavelength range. The LOS velocity maps will have the accuracy and stability to allow detailed helioseismic investigations of the solar interior, in particular of the solar convection zone.

SPICA/SAFARI

SPICA (SPace IR telescope for Cosmology and Astrophysics) is a joint European-Japanese Project, one of the 3 selected candidates by ESA for M5 (final selection planned in 2021, launch expected for 2032). Its science goals are understanding the origin and evolution of galaxies, stars, planets and life itself. These topics can be explored with a sensitive infrared survey, peering through the clouds of dust that typically obscure the sites of star birth.

SPICA will have a mirror of around 2.0 meter in diameter, and a total size of 4.5 × 5.9 meter housing two instruments, being SAFARI the largest and most complex one.

SAFARI will provide low (R~300) to medium (R up to 11000) resolution spectroscopic capabilities at a uniquely high sensitivity of a few times 10-20 Wm^{-2} (5σ, 10hr), instantaneously covering the full 35 to 230 μm range. The SAFARI detectors which make possible this sensitivity are TES based superconducting bolometers.

STROBE-X

The Spectroscopic Time-Resolving Observatory for Broadband X-rays (STROBE-X) probes strong gravity for stellar mass to supermassive black holes and ultradense matter with unprecedented effective area, high time-resolution, and good spectral resolution, while providing a powerful time-domain X-ray observatory. STROBE-X was approved by NASA as a Probe-class Mission for the 2020 US decadal survey.

GROUND FACILITIES

The main ground observatories in development are the ELT and ESFRIs SKA, CTA and EST. These and several other smaller instruments are briefly described below.

CARMENES

CARMENES (Calar Alto high-Resolution search for M dwarfs with ExoEarths with Near- infrared and optical Échelle Spectrographs) is a next-generation instrument built for the 3.5m telescope at the Calar Alto Observatory by a consortium of German and Spanish institutions. It consists of two separated spectrographs covering the wavelength ranges from 0.52 to 0.96 μm and from 0.96 to 1.71 μm with spectral resolutions R = 80,000–100,000, each of which performs high-accuracy radial-velocity measurements (~1 m s−1) with long-term stability. The fundamental science objective of CARMENES is to carry out a survey of ~300 late-type main-sequence stars with the goal of detecting low-mass planets in their habitable zones. We aim at being able to detect 2 MEarth planets in the habitable zone of M5V stars. The CARMENES first light with the two NIR and VIS channels working simultaneously occurred in Nov 2015.

An upgrade of the instrument, CARMENES+, is in process of definition.

CTA (Cherenkov Telescope Array)

The Cherenkov Telescope Array (CTA) is the next generation ground-based observatory for gamma-ray astronomy at very-high energies. With more than 100 telescopes located in the northern and southern hemispheres, CTA will be the world's largest and most sensitive high-energy gamma-ray observatory. Its construction is expected to begin in 2022 and last until 2025.

CTA's key targets are multi-purpose observations designed to efficiently address the broad-ranging science questions of CTA's study themes. The initial list includes: the galactic center, the Large Magellanic Cloud, the galactic plane, galaxy clusters, CosmicRay PeVatrons, Star Forming Systems, Active Galactic Nuclei, Transient Phenomena.

CTA will be ten times more sensitive and have unprecedented accuracy in its detection of high-energy gamma rays. Current gamma-ray telescope arrays host up to five individual telescopes, but CTA is designed to detect gamma rays over a larger area and a wider range of views with more than 100 telescopes located in the northern and southern hemispheres.

DESI (LLNL) at Mayall Telescope (USA)

The Dark Energy Spectroscopic Instrument (DESI) will measure the effect of dark energy on the expansion of the Universe. It will obtain optical spectra for tens of millions of galaxies and quasars, constructing a 3D map spanning the nearby Universe to 11 billion light years.

The DESI Survey is conducted on the Mayall 4-meter telescope at Kitt Peak National Observatory since 2019.

ELT (Extremely Large Telescope)

Extremely large telescopes are considered worldwide to be one of the highest priorities in ground-based astronomy. They will vastly advance astrophysical knowledge, allowing detailed studies of subjects including planets around other stars, the first objects in the Universe, supermassive black holes, and the nature and distribution of the dark matter and dark energy which dominate the Universe.

ESO's ELT, under construction, will see the first light in 2026. The ELT will have a 39-metre main mirror and will be the largest optical/near-infrared telescope in the world, with a performance that is orders of magnitude better than currently existing facilities. It will host 6 instruments, with CSIC participation in three of them: MOSAIC (Multi- Object Spectrograph) will allow near-infrared spectroscopy of large samples of astronomical objects; HIRES (High Resolution Spectrograph) combines a high resolution and wide spectral range; and HARMONI (High Angular Resolution Monolithic Optical and Near-infrared Integral field spectrograph), which will be used to explore galaxies in the early Universe, study the constituents of the local Universe and characterize exoplanets in great detail. In many ways, it will be complementary to ALMA and the James Webb Space Telescope.

European Solar Telescope (EST)

The European Solar Telescope (EST), to be located in the Canary Islands, will be a next generation large-aperture solar telescope. This 4-metre telescope will be optimized for studies of the magnetic coupling between the deep photosphere and upper chromosphere. This will require diagnostics of the thermal, dynamic and magnetic properties of the plasma over many scale heights, by using multiple wavelength imaging, spectroscopy and spectropolarimetry. To achieve these goals, the EST will specialize in high spatial and temporal resolution using various instruments simultaneously that can efficiently produce 2D spectral information.

IRAM-NIKA2

The new KID Array 2 (NIKA2) installed at IRAM 30m telescope in Pico Veleta is a dual- band camera operating with three frequency-multiplexed kilopixels arrays of Lumped Element Kinetic Inductance Detectors (LEKID). The NIKA2 is a new generation instrument for millimeter astronomy. NIKA2 allows to observe the sky at 150 and 260 GHz in a wide field of view with high angular resolution, in polarization at 260 GHz and with an excellent sensitivity. The scientific program is dedicated to studying the internal structure of galaxy clusters the formation of stars in our Galaxy, studying the role of the magnetic field at the scale of one tenth of a parsec, and in distant galaxies, detecting hundreds of infrared galaxies during their major star formation episodes.

QUIJOTE

QUIJOTE (Q U I JOint TEnerife) operates from Teide Observatory with the goal of characterizing the polarization of the CMB and other galactic and extragalactic emission in the frequency range 10-40 GHz, and at large angular scales. The main objective of the QUIJOTE project is to cover a sky area of 5, 000 square degrees, with a sensitivity around 1 μK (at the highest frequencies) and an angular resolution of 1° at 11, 13, 17, 19, 30 and 40 GHz. These will be the most sensitive measurements obtained for characterization of the synchrotron and anomalous microwave emission in our Galaxy at those frequencies.

QUIJOTE features 2 telescopes and 3 instruments, two of them (the Thirty-GHz (TGI) and Fourty-GHz (FGI) Instruments, in commissioning phase. KISS is a KID camera to be installed in the first telescope.

SKA (Square Kilometer Array)

The SKA (Square Kilometer Array) will be the largest radio telescope array ever constructed, by combining the unprecedented sensitivity of the thousands of individual radio receivers. Its total collecting area will be over one square kilometer or 1,000,000 square meters. To achieve this, the SKA will use hundreds of dishes and hundreds of thousands of low-frequency aperture array telescopes. that will be arranged in multiple spiral arm configurations, with the dishes extending to vast distances from the central cores, creating what is known as a long baseline interferometer array. In such an array, the physical distance that separates the telescopes is calculated precisely using the time difference be- tween the arrival of radio signals at each receiver. Computers can then calculate how to combine these signals to synthesize something the equivalent size of a single dish measuring the width of the distance between the two scopes. The system can act either as one gigantic telescope, or multiple smaller telescopes and any combination in between.

Some of the main SKA science drivers include: Galaxy evolution, cosmology and dark energy, Strong-field tests of gravity using pulsars and black holes, the origin and evolution of cosmic magnetism, how were the first black holes and stars formed, and possible life beyond Earth. It will feature a flexible design to enable also exploration of the unknown.

3. PARTICLE, ASTROPARTICLE AND NUCLEAR PHYSICS FACILITIES

The Large Hadron Collider (LHC)
The LHC is a particle accelerator, at CERN, that collides proton beams at a center of mass energy of 14 TeV in four locations around the 27-kilometer ring corresponding to four particle detectors: ATLAS, CMS, LHCb and ALICE. ATLAS and CMS are two general purpose detectors. At 46 m long, 25 m high and 25 m wide, the 7000-tonne ATLAS detector is the largest volume particle detector ever constructed. The 14,000-tonne CMS detector is 21 meters long, 15 meters wide and 15 meters high. LHCb specializes in flavour physics and ALICE is a heavy ion collider.

The High Luminosity Large Hadron Collider (HL-LHC)
The High-Luminosity Large Hadron Collider (HL-LHC) at CERN, whose operation is expected to start around 2027, is an upgrade of the LHC that will increase the luminosity of the 14 TeV LHC by a factor 10 resulting in unprecedented particle multiplicities and operational radiation levels for the detectors that will need to be replaced for this upgrade.

The Future Higgs Factory (ILC, CLIC, FCC-ee, CEPC)
Although still discussed within the particle physics community, the most likely scenario for an accelerator after the HL-LHC is an electron-positron collider with center of mass energies around twice the Higgs boson mass. There are different projects addressing this which are:

- The International Linear Collider (ILC). The ILC is an international project to build a 31-kilometer-long linear collider in Japan. The accelerator will reach energies of 250 GeV, upgradable up to 500 GeV. It uses Superconduction Radio Cavities to provide around 30 MV/m accelerating gradients.

- The Compact Linear Collider (CLIC) is a project to build a multi-TeV accelerator at CERN. The first phase of the project aims at a center of mass energy of 380 GeV by using radiofrequency structures and a two-beam concept to produce accelerating fields as high as 100 V/m.
- FCC-ee is the electron-positron circular collider proposed by CERN. It would consist of a 100 km tunnel housing the accelerator that would have a center of mass energy as high as about 380 GeV.
- CECP (Circular Electron Positron Collider) is a similar proposal for an electron- positron collider at China. It would also consist of a 100 km tunnel for an accelerator with an energy of the order of 240 GeV.

Colliders for the Energy Frontier (FCC-hh, SppC)

Both FCC-ee and CEPC have plans to upgrade the machine to accelerate protons to energies as high as 100 TeV. FCC-hh is the hadron-hadron Future Circular Collider proposed by CERN and SppC (Super proton-proton Collider) is equivalent Chinese project. Both projects will need to extend the current magnet technology to provide the required 16-18 T to bend the beams into the 100 km circumference.

Deep Underground Neutrino Experiment (DUNE)

DUNE is a long baseline neutrino experiment providing a high intensity neutrino beam produced at Fermilab that will be detected 1300 km away by a very large, modular Liquid Argon Time-Projection Chamber (LArTPC) with a 40 kt fiducial mass located 1.5 km underground. This LAr technology will make it possible to reconstruct neutrino interactions with image-like precision and unprecedented resolution.

KM3NeT

KM3NeT is an underwater neutrino telescope designed for the detection of high-energy neutrinos of cosmic origin that will fill a cubic kilometer of sea water with light sensors to detect the Cherenkov light produced by neutrinos.

NEXT

NEXT is a neutrino less double–beta decay experiment consisting in a Time Projection Chamber (TPC) filled with high-pressure gaseous xenon and with separated-function capabilities for calorimetry and tracking.

AGATA

The Advanced GAmma Tracking Array (AGATA) is a European gamma-ray spectrometer used for nuclear structure studies with the aim of developing and building a 4π gamma-ray spectrometer of the next generation. AGATA is based on the principle of gamma-ray tracking, which is made possible by the advent of segmented high-purity germanium crystals, advanced digital electronics, and pulse-shape analysis. The full AGATA spectrometer will consist of an array of 180 large (9.0 cm length, 8.0 cm circular diameter) encapsulated high-purity germanium crystals.

FAIR

The FAIR (Facility for Antiproton and Ion Research in Europe) facility in Darmstadt, Germany, will be able to generate ion beams of all the natural elements in the periodic table. The key component of FAIR is a ring accelerator with a circumference of 1,100 meters. Connected to this is a complex system of storage rings and experimental stations. The existing GSI accelerators will serve as the first acceleration stage. One of the four basic pillars of FAIR is NuSTAR (NUclear STructure, Astrophysics and Reactions) which encompasses all the experiments that will be exploiting the Radioactive Ion Beams (RIBs) among which we encounter R3B and HISPEC/DESPEC.

SPIRAL2

SPIRAL2 (Système de Production d'Ions Radioactifs en Ligne de 2e generation) in Ganil, France, is a new facility to extend significantly the actual possibilities of Radioactive Ion Beam (RIB) physics and related applications. SPIRAL2 comprises a linear accelerator (LINAC) and experimental areas with three halls for experiments with high flux of fast neutrons, with very high intensity beams of heavy-ions and with low-energy exotic nuclei.

SPES

SPES (Selective Production of Exotic Species) is the new project in the Laboratori Nazionale di Legnaro, in Italy. The SPES facility is mainly concentrating on the production of neutron-rich radioactive nuclei having mass in the range 80-160 based on the ISOL technique.

ISOLDE

ISOLDE (Isotope mass Separator On-Line facility) at CERN, in Switzerland, is a unique source of low-energy beams of radioactive nuclides, those with too many or too few neutrons to be stable. The high intensity proton beam from the Proton Synchrotron Booster (PSB) is directed into specially developed thick targets, yielding a large variety of atomic fragments. Commissioned in October 2015, the new linear accelerator HIE-ISOLDE (High Intensity and Energy ISOLDE) brings the energy of the radioactive beams up to 7.5 MeV/nucleon. The new facility will ultimately accelerate the nuclei up to 10 MeV/nucleon.

n_TOF

The n_TOF (neutron Time Of Flight) facility is a pulsed neutron source coupled to a 200- meter flight path. It is designed to study neutron-nucleus interactions for neutron energies ranging from a few MeV to several GeV. The wide energy range and high-intensity neutron beams produced at n_TOF are used to make precise measurements of neutron-related processes. The neutrons are produced from a pulsed beam of protons form the CERN's Proton Synchrotron (PS) hitting a lead target.

RIKEN

RIKEN Radioactive Ion Beam Factory is a research center for nuclear physics located in Wako city (Japan). This research center houses the world's largest superconducting ring cyclotron "SRC" and the superconducting radioactive beam separator "BigRIPS". The facility presently provides radioactive beams of unparalleled intensities in the world for short-lived nuclei using the in-flight separation technique.

EURISOL

The EURISOL project is aimed at the design - and eventual construction - of a 'next- generation' European ISOL radioactive ion beam (RIB) facility capable of extending current research in atomic and nuclear physics by providing users with a wide variety of exotic ion beams at intensities far greater than those presently available. An intensive design study where IFIC and IEM are involved is ongoing. The facility is planned for the far future +2035.

Understanding the basic components of the Universe, its structure, and evolution is one of the noblest and most ambitious undertakings of humanity. The fundamental laws of nature are the basis of all technology. Finding and understanding them requires exploring both the elementary components of matter at the smallest scales and the observable Universe at the largest scales. Only through understanding the laws that dictated the first instances of the Universe will we be able to find the ultimate reason for its actual appearance and future fate. Challenges in physics are intimately associated with technological challenges for the design and building of telescopes, space missions or accelerators, reactor and underground experiments, as well as developments in mathematics. Their resolution requires a worldwide, transdisciplinary effort and the orchestrated involvement of researchers, engineers, and technicians. The challenges described here are fully aligned with the priorities identified in international and European strategies.

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS
EDITORIAL CSIC