Vulcano 2016
Frontier Objects in Astrophysics and Particle Physics

Editors:
R. Fusco Femiano, G. Mannocchi, A. Morselli, G. Trinchero
The sixteenth edition of the Vulcano workshop: Frontier objects in Astrophysics and Particle Physics was organized jointly by the National Institute for Nuclear Physics (INFN) and the National Institute for Astrophysics (INAF). The workshop was held in the Conference Room of the Therasia Resort (Vulcano Island, Sicily, Italy) from May 22th to May 28th and was attended by more than 80 scientists worldwide. This workshop is certainly one of the first that since 1986 has the aim to gather people from High Energy Astrophysics and Particle Physics to discuss the most recent highlights in these fields. It is well known that at the beginning of the 80’s the Universe was considered the greatest particle accelerator of the world to test the Grand Unified Theories ideas. Of course a machine hard to use because all the experiments happened only once, a long time ago. Today, gigantic underground accelerators and space crafts probe everyday this connection. As never before, these two fields of knowledge complement and integrate each other. The discovery of new particles may unveil some cosmic mysteries, conversely, astrophysical observations may give new information on the infinitely small. In particular Cosmology provides three fundamental questions for the Universe, namely: what is Dark Energy? What is Dark Matter? And why in the Universe we have matter and not antimatter, and therefore one of the goals of this workshop was to discuss if these questions may be resolved, at least partially, by particle physics. In September 2015 we had the spectacular detection of gravitational waves. Since the first edition in 1986 the Vulcano workshop has always reserved a session to the GW. In this edition of the Vulcano workshop we have discussed not only the results of these observations but also the possibilities to investigate the universe given by the new window that may be opened by the rising gravitational astronomy. Wednesday morning a session has been dedicated to the memory of Professor Guido Altarelli that two years ago was one of the participants to the Vulcano workshop. We have remembered his rich scientific activity. The final scientific program was selected by the Scientific Organizing Committee chaired by Roberto Fusco-Femiano (INAF) and Giampaolo Mannocchi (INFN) and composed by: Antonella Antonelli (INFN), Simone Dell’Agnello (INFN), Pino Di Sciascio (INFN), Aurelio Grillo (INFN), Aldo Morselli (INFN), Luigi Piro (INAF), Marco Ricci (INFN) and Gian Carlo Trinchero (INAF). The Local Organizing Committee was composed by Maria Cristina D’Amato (INFN), Roberto Fusco-Femiano, Giampaolo Mannocchi, and Lia Sabatini (INFN) with the precious help of Alessio Gorgi (INAF). A special thank to Maria Cristina D’Amato and Lia Sabatini for their fundamental work not only in the preparatory phase but also during and after the conclusion of the workshop, allowing us to receive numerous compliments on the level of the Conference.

Roberto Fusco-Femiano and Giampaolo Mannocchi
To our great friend and colleague Aurelio Grillo who has suddenly left us on February 16, 2017. He was member of the Scientific Organizing Committee since the first Vulcano Workshop in 1986 always bringing his enthusiastic and valuable scientific contribution.
Conference Program

Gravitational Waves

G.A. Prodi  Overview of gravitational wave observations by LIGO and Virgo
M.Bassan  Lisa pathfinder first step toward a gravitational wave space observatory
S.Capozziello  Beyond einstein: gravitational waves from extended gravity

Future Prospects

F. Linde  Astroparticle Research in Europe
Y. Wang  Astroparticle Research in China
P. Ubertini  INAF Programs for High Energy Astrophysics
M. Pallavicini  INFN Astroparticle programs
R. Forty  Future of the LHC
P. Ubertini  Space astronomy in future
B. Negri  ASI Science Programs: Current and Future Space Missions
M. Branchesi  EM Counterparts/Follow up of GW Sources

Gravity

M. Martini  Experimental gravity tests in the solar system
R. March  Nonminimally coupled gravity and planetary motion
R. Casadio  Black holes: a window into strong (quantum?) gravity
Y. Jack Ng  Holographic theory of gravity and cosmology

Cosmic Rays

P. S. Marrocchesi  The CALET Mission on the International Space Station
M. Teshima  Ground based Gamma Ray Astronomy Review
G. Brunetti  Cosmic Rays in galaxy clusters and constraints from gamma ray
C. Pittori  High Energy Astrophysics
L. Latronico  The Fermi Observatory - status and prospects
M. Mostafa  The High Altitude Water Cherenkov Observatory
C. Pittori  Search of MeV-GeV counterparts of TeV sources
M. Ahlers  The Power Spectrum of Cosmic Ray Arrival Directions
G. Giacinti  Particle Acceleration at Supernova Remnants and Supernovae
C. Evoli  Inhomogeneous CR transport in the Galaxy
P. Picozza  Cosmic-ray direct detection
E. S. Seo  The Cosmic Ray Energetics And Mass Project
A. Chiavassa  Review on galactic cosmic ray detection
Z. Cao Prospects of spectrum measurements for individual species of cosmic ray particles above 100 TeV using LHAASO array
J. Gonzalez IceTop Results on Cosmic Ray Detection
J. Hörandel High-energy cosmic rays: galactic or extragalactic?
P. Biermann The nature and origin of ultra-high energy cosmic ray particles
P. L. Ghia Review on UHECR (Experimental)
A. Castellina Selected results and perspectives from the Pierre Auger Observatory
J. Matthews Recent Results from the Telescope Array Ultra High Energy Cosmic Ray Observatory
M. Sitta Cosmic-ray studies using the alice detector at LHC
F. Schröder Review on Cosmic-Ray Radio Detection

Neutrino

A. Smirnov Overview on Neutrino Physics
J. Cao Reactor neutrinos and mass hierarchy
A. Giuliani Neutrinoless double beta decay: where we are and where we are going
P. Piattelli ANTARES + Km3Net
I. Taboada Understanding ICExUBE’s astrophysical neutrino observations
C. Virtue The Halo 2 Supernova Detector

In memory of Guido Altarelli

M. Greco Remembering Guido Altarelli*
R. Barbieri Beyond the Standard Model
F. Feruglio Neutrino masses and mixing angles: a tribute to Guido Altarelli
A. Masiero Guido’s contribution to the life of INFN

Dark Matter and Dark Energy

N. Menci Limits on warm dark matter particle masses from the abundance of ultra-faint galaxies at high redshifts in the Hubble frontier fields
P. Ullio Indirect and direct Dark Matter search
A. Morselli Gamma-rays signatures of dark matter: status and future prospect
M. Taoso A realistic model for cosmic-rays production and the galactic center GeV excess.
R. Bernabei Signals from the dark universe
E. Aprile The XENON1T Dark Matter Search at LNGS
R. Scaramella Euclid: the next ESA mission for Cosmology
M. Crosta The relativistic all-sky analysis with GAIA
**Astrophysics and Cosmology**

E. F. Branchini  the large scale structure path to the high energy universe
U. Maio  The first stars in the universe
M. Boer  Multimessenger astronomy, the dawn of a new era for the knowledge of the Universe
F. Nicastro  The Milky Way Hot Baryons and their Peculiar Density Distribution: a Relic of Nuclear Activity
P. Mazzotta  Observations of Galaxy Clusters in the Microwave and the X-rays
F. Haardt  The probes and sources of cosmic reionisation
E. Troja  GRBs multimessenger
C. Gustavino  Underground study of big bang nucleosynthesis in the precision era of cosmology
M. Burgay  Can we use Fast Radio Bursts as cosmological probes?
L. Feretti  Science with the Square Kilometer Array

**Particle Physics**

A. Petrukhin  TeV energy physics at LHC and in cosmic rays
M. Antonelli  Accelerator dark matter searches: comparison with direct and indirect searches
T. Spadaro  Dark Matter search at fixed target
V. Kozhuharov  New projects on dark photon search
G. Battistoni  Nucleusnucleus interactions and their application in medicine
R. Di Nardo  Recent Higgs Boson Results from the LHC
P. Campana  Flavor perspectives with the LHCb upgrade
M. Pepe Altarelli  Recent selected highlight from LHCb

Slides are available at: https://agenda.infn.it/conferenceOtherViews.py?view=nicecompact&confId=10576

*"The Early Days of QCD in Rome in the Seventies” video by G. Altarelli*
Fr ontier Objects in Astro p hysics and P ar ticle Phys ics
C onference Proceedi ngs

M. Bassan Lisa pathfinder first step toward a gravitational wave space observatory 1
S. Capozziello Beyond einstein: gravitational waves from extended gravity 11
R. Forty Future of the LHC 19
M. Martini Experimental gravity tests in the solar system 27
R. March Nonminimally coupled gravity and planetary motion 33
R. Casadio Black holes: a window into strong (quantum?) gravity 41
Y. Jack Ng Holographic theory of gravity and cosmology 49
P. S. Marrocchesi The CALET Mission on the International Space Station 57
P. Picozza Cosmic-ray direct detection 65
A. Chiavassa Review on galactic cosmic ray detection 76
Z. Cao Prospects of spectrum measurements for individual species of cosmic ray particles above 100 TeV using LHAASO array 85
J. Hörandel High-energy cosmic rays: galactic or extragalactic? 95
P. Biermann The nature and origin of ultra-high energy cosmic ray particles 103
A. Castellina Selected results and perspectives from the Pierre Auger Observatory 122
M. Sitta Cosmic-ray studies using the alice detector at LHC 131
F. Schröder Review on Cosmic-Ray Radio Detection 139
A. Giuliani Neutrinoless double beta decay: where we are and where we are going 147
I. Taboada Understanding ICECUBE’s astrophysical neutrino observations 156
M. Greco Remembering Guido Altarelli 164
F. Feruglio Neutrino masses and mixing angles: a tribute to Guido Altarelli 174
N. Menci Limits on warm dark matter particle masses from the abundance of ultra-faint galaxies at high redshifts in the Hubble frontier fields 183
A. Morselli Gamma-rays signatures of dark matter: status and future prospect 191
M. Taoso A realistic model for cosmic-ray production and the galactic center GeV excess. 203
R. Bernabei Signals from the dark universe 211
<table>
<thead>
<tr>
<th>Author</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>R. Scaramella</td>
<td>Euclid: the next ESA mission for Cosmology</td>
<td>222</td>
</tr>
<tr>
<td>M. Crosta</td>
<td>The relativistic all-sky analysis with GAIA</td>
<td>233</td>
</tr>
<tr>
<td>E. F. Branchini</td>
<td>the large scale structure path to the high energy universe</td>
<td>240</td>
</tr>
<tr>
<td>U. Maio</td>
<td>The first stars in the universe</td>
<td>247</td>
</tr>
<tr>
<td>F. Nicastro</td>
<td>The Milky Way Hot Baryons and their Peculiar Density Distribution: a Relic of Nuclear Activity</td>
<td>255</td>
</tr>
<tr>
<td>F. Haardt</td>
<td>The probes and sources of cosmic reionisation</td>
<td>264</td>
</tr>
<tr>
<td>C. Gustavino</td>
<td>Underground study of big bang nucleosynthesis in the precision era of cosmology</td>
<td>272</td>
</tr>
<tr>
<td>L. Feretti</td>
<td>Science with the Square Kilometer Array</td>
<td>283</td>
</tr>
<tr>
<td>A. Petrukhin</td>
<td>TeV energy physics at LHC and in cosmic rays</td>
<td>291</td>
</tr>
<tr>
<td>M. Antonelli</td>
<td>Accelerator dark matter searches: comparison with direct and indirect searches</td>
<td>301</td>
</tr>
<tr>
<td>V. Kozhuharov</td>
<td>New projects on dark photon search</td>
<td>310</td>
</tr>
<tr>
<td>G. Battistoni</td>
<td>Nucleusnucleus interactions and their application in medicine</td>
<td>319</td>
</tr>
<tr>
<td>R. Di Nardo</td>
<td>Recent Higgs Boson Results from the LHC</td>
<td>327</td>
</tr>
<tr>
<td>P. Campana</td>
<td>Flavor perspectives with the LHCb upgrade</td>
<td>335</td>
</tr>
</tbody>
</table>
LISA PATHFINDER
FIRST STEP TOWARD A GRAVITATIONAL WAVE SPACE OBSERVATORY

Massimo Bassan
Dipartimento di Fisica Universit`a di Roma Tor Vergata & INFN Roma Tor Vergata

On behalf of the LISA Pathfinder collaboration (*)

Abstract
We briefly review the concept of a space-based gravitational wave interferometer, and the science it can explore in the milliHertz frequency region. Then we discuss the LISA Pathfinder technology demonstrator mission that is currently flying and will soon deliver the first results.

1 Gravitational astronomy from space
Gravitational waves have finally been detected 1), one century after being predicted by A. Einstein 2). The very first signal has already shown the amazing power of gravitational astronomy, offering detailed information about the merging of two black holes of unexpected mass. In the years to come, Earth-based Advanced Interferometers will provide a great wealth of information about compact objects like neutron stars and black holes in their final stages of life.
What is the role then of a space observatory like LISA, expected to operate no sooner than fifteen years from now?

Theoretical predictions suggest that a wide variety of astrophysical sources emit gravitational waves (g.w.) in that part of the frequency spectrum, from 0.1 mHz to about 100 mHz, that is only accessible by a space interferometer. Signals gathered in this band can provide crucial insights about binary stars formation, Extreme Mass Ratio Inspirals, SuperMassive Black Holes and galaxy mergers. An observatory in that frequency range has moreover guaranteed signals for calibration, the so called Verification Binaries, of which all parameters are known by e.m. observations and whose g.w. flux is precisely predicted.

2 The LISA mission concept

The frequency range below 1 Hz, where all these sources are to be observed, is only accessible by a space-borne detector, that is immune from Earth-generated, low frequency disturbances, like seismic and Newtonian noise. LISA is a space mission concept, developed over two decades through several different configurations, that addresses this request. LISA’s basic scientific and technological fundamentals are well established since many years. Its actual design has undergone several revisions, but has reached a stable configuration since 1998; it relies on a few pillars:

- **The orbital configuration:** a redundant set of three interferometers with arm length of 1-5 Gm (million kilometers), hosted by a constellation of three satellites arranged at the vertices of a gigantic triangle. Each spacecraft (S/C) passively move on a different "smart" heliocentric orbit; while keeping the relative distance roughly constant, the triangular pattern rotates around its center, that moves on the ecliptic, trailing the Earth by some 20°. This motion modulates the g.w. signals, giving them a unique signature.

- **The optical transponders:** the divergence of a laser beam is such that, even using a large telescope, only a few parts in $10^9$ of the incoming beam can be received at the far end of such a long interferometric arm, thus making it unfeasible to reflect the light back with a mirror. A transponder is a...
phase-preserving amplifier that will regenerate the beam to its full power (1-5 W) and send it back to the other S/C for phase comparison.

- **Time delay interferometry**: The distances between S/C’s vary in time, due to Keplerian dynamics, of 1-2 %, i.e. up to $10^5$ km: the unequal arm length gives rise to frequency noise in the interferometer. This noise can be depressed, by several order of magnitude, by recording the phase signal of each arm and then, off-line, synthesizing linear combinations of the signals emitted at different times $^7$, before comparing them, thus canceling most of the frequency noise. So, the usual optical path difference with equal arms $X(t) = y_1(t) - y_2(t)$ is substituted by:

$$X(t) = y_1(t) - y_2(t) - y_1(t - 2L_2/c) + y_2(t - 2L_1/c) \quad (1)$$

More complex relations (second order TDIs) would apply if one also wants to cancel the Doppler effect of different S/C velocities.

- **Test masses in free fall**: aboard each S/C, two test masses (TM), one for each interferometer arm ending at that node, are in geodesic motion, shielded by the spacecraft from any external disturbance, and responding only to space-time perturbation. A feedback loop, acting on the S/C microthrusters, recenters the S/C on the TM along the sensitive axis. All other Degrees of Freedom (DoFs) are stabilized by electrostatic forces. Each interferometric arm acts then as a huge differential accelerometer.

In 2011, financial constraints forced NASA to withdraw its support for the LISA mission. European scientists undertook an intense effort to "rescope" the mission so that it could fit in the "ESA only" budget: the ambitious goal was to achieve the same physics with half the budget. This study produced a new configuration, called "eLISA": still with three spacecrafts, but with only one interferometer (two arms, therefore giving up redundancy and polarization detection capabilities), a reduced (1 or 2 Gm) arm length (thus making do with less laser power and without a pointing mechanism for the telescopes) and a shorter mission duration. This project, presented to ESA in 2013 $^8$, gained for the theme The Gravitational Universe the selection to third large ESA mission of the Cosmic Vision Plan (L3), for launch in 2034. Nevertheless, recent re-analyses of the science, technology and finances for the mission make
us hopeful that the full, three arm configuration can be regained before the
detailed mission formulation.

Figure 1: The *smart orbits* of the space g.w. observatory: each spacecraft moves
on an independent, passive, Keplerian orbit, and the constellation maintains a
triangular formation, while rotating around its center. Picture from ref. 8)

3 LISA Pathfinder, the technology demonstrator

The technical difficulty of the LISA project made it sensible to test as much
as possible of the enabling technology in a dedicated mission. Particularly
challenging is the assessment of the quality of free fall that can be achieved
in space, despite extensive laboratory investigations, carried out with torsion
pendulums 9), of the residual forces that can locally disturb the inertial motion
of the TM. Other items that need to be space-tested are the interferometric
read-out of the TM positions, the Gravitational Reference Sensor (GRS) that
senses motion in the other degrees of freedom and applies electrostatic feedback,
its front-end electronics (FEE) the microthrusters, TM decaging and release
(with minimum velocity) and more.

Therefore ESA has approved and realized a dedicated space mission to
verify all the flight hardware that can be tested in a single-spacecraft mission,
and to evaluate the level of geodesic motion that current technology can achieve

4
and measure. Such mission, called LISA Pathfinder \textsuperscript{10}, is active and operational now.

In LISA Pathfinder, a single spacecraft holds two independent TMs: once spaceborne, each TM is released within its GRS and their relative distance is measured by a differential interferometer: in many respects, this is a LISA arm shrunk from few Gm down to 38 cm. Feedback loops act on the spacecraft and on the second TM to make sure that they both follow the first TM in its free fall. Most of the hardware (with few exceptions like the telescope, the pointing mechanism and the TDI) are therefore tested in an environment that is representative of the final LISA mission. Moreover, a large number of noise sources, i.e. all local sources, like thermal, magnetic, laser shot and radiation pressure, long term changes in local gravity etc. can be accurately assessed and measured, so that a reliable noise budget can be formulated for the LISA mission.

LPF was launched with the sixth VEGA launcher on Dec 3rd 2015, one century after publication of Einstein’s first paper on the \textit{General Theory of Relativity} and few weeks after the detection of the first g.w. signal by LIGO, thus marking an unforgettable year for Gravitational research. After a few

Figure 2: Schematics of the LISA Pathfinder payload: two Test Masses, each surrounded by its GRS (electrostatic sensors and actuators) and, in between them, the monolithic Optical Bench, with all the optics needed to operate 4 different interferometers
orbits around the Earth, raising each time its apogee, it has started its journey toward the first Lagrangian point (L1) of the Sun-Earth system, where the gravitational gradient is a minimum, about 1.5 Gm from the Earth. It has now reached its target orbit, a large, slow orbit around L1 where it has undergone testing and commissioning. At the time of this talk, operations are under way and the first results are extremely promising. However, no data has yet been published, and expectations are high for the first release due in June.

4 Post-conference update: first results from LISA Pathfinder

![Figure 3: Spectrum of the differential acceleration noise of the two Test Masses of LISA Pathfinder. We also show the requirements, well exceeded the present mission and closely approached for the future LISA observatory. Some known sources of noise, measured and subtracted, are also indicated. The rise on the right is readout noise (white in displacement), that grows as $\omega^2$ when converted to acceleration. Figure from ref. 11).](image)

On June 7th 2016, about two weeks after the workshop, the first LPF data were released 11). The results are excellent, beyond the most optimistic expectations: the interferometer readout noise is $35 \text{fm}/\sqrt{\text{Hz}}$, that is 100 times better than the requirements and than noise measured on ground, where mirrors could be carefully hand-aligned. The purity of free-fall is gauged by the differ-
ential acceleration of the two test masses, and is measured to be $5.2 \times 10^{-15} \text{ms}^{-2}$ i.e. below the femto-g range. This is also lower, by a factor 5, than the mission requirements and closely approaches the tougher spec required for LISA. Even more striking, the specs are exceeded not only on the required frequency region, i.e. down to 1 mHz, but on a band extending almost down to 0.1 mHz, i.e. on the full LISA sensitivity band. Figure 3 shows the residual acceleration noise, together with the LISA Pathfinder requirements and the LISA requirements, that are approached to within a factor 1.5. Moreover, the noise in the “flat” region 2-8 mHz, that is mainly of Brownian origin, appears to be decreasing with time: a probable reason for this is the continuing decrease in pressure due to venting to open space.

The tests and investigations on LISA Pathfinder will continue for the next several months, with the aim of characterizing all components of the residual noise, but we can already confidently state that LISA Pathfinder is doing its job and the road to LISA looks brighter than ever.

References

3. See e.g. K. Danzmann, A. Rüdiger, Class. Quantum Grav. 20 (2003) S159
4. LISA Pre-Phase A Report, MPQ 233, July 1998
   M. Bassan et al. Phys Rev Lett 116, 051104 (2016) and references therein
(* ) The LISA Pathfinder collaboration:

M Armano\(^a\), H Audley\(^b\), G Auger\(^c\), J T Baird\(^d\), M Bassan\(^e\), P Binetruy\(^c\), M Born\(^b\), D Bortoluzzi\(^f\),
N Brandt\(^g\), M Caleno\(^b\), L Carbone\(^i\), A Cavalleri\(^bb\), A Cesarini\(^i\), G Ciani\(^i\),\(^cc\),
G Congedo\(^i\),\(^cc\), A M Cruise\(^j\), K Danzmann\(^b\), M de Deus Silva\(^a\), R De Rosa\(^g\), L Di Fiore\(^y\),
M Diaz-Aguiló\(^n\), I Diepholz\(^b\), G Dixon\(^j\), R Dolesi\(^i\), N Dunbar\(^k\), L Ferraioli\(^l\), V Ferroni\(^i\), D Flatscher\(^g\),
N Brandt\(^g\), M Caleno\(^h\), L Carbone\(^i\), A Cavalleri\(^bb\), A Cesarini\(^i\), G Ciani\(^i\),\(^dd\),
G Congedo\(^i\),\(^cc\), A M Cruise\(^j\), K Danzmann\(^b\), M de Deus Silva\(^a\), R De Rosa\(^g\), L Di Fiore\(^y\),
M Diaz-Aguiló\(^n\), I Diepholz\(^b\), G Dixon\(^j\), R Dolesi\(^i\), N Dunbar\(^k\), L Ferraioli\(^l\), V Ferroni\(^i\), D Flatscher\(^g\),
N Brandt\(^g\), M Caleno\(^h\), L Carbone\(^i\), A Cavalleri\(^bb\), A Cesarini\(^i\), G Ciani\(^i\),\(^dd\),
G Congedo\(^i\),\(^cc\), A M Cruise\(^j\), K Danzmann\(^b\), M de Deus Silva\(^a\), R De Rosa\(^g\), L Di Fiore\(^y\),
M Diaz-Aguiló\(^n\), I Diepholz\(^b\), G Dixon\(^j\), R Dolesi\(^i\), N Dunbar\(^k\), L Ferraioli\(^l\), V Ferroni\(^i\), D Flatscher\(^g\),
N Brandt\(^g\), M Caleno\(^h\), L Carbone\(^i\), A Cavalleri\(^bb\), A Cesarini\(^i\), G Ciani\(^i\),\(^dd\),
G Congedo\(^i\),\(^cc\), A M Cruise\(^j\), K Danzmann\(^b\), M de Deus Silva\(^a\), R De Rosa\(^g\), L Di Fiore\(^y\),
M Diaz-Aguiló\(^n\), I Diepholz\(^b\), G Dixon\(^j\), R Dolesi\(^i\), N Dunbar\(^k\), L Ferraioli\(^l\), V Ferroni\(^i\), D Flatscher\(^g\),
N Brandt\(^g\), M Caleno\(^h\), L Carbone\(^i\), A Cavalleri\(^bb\), A Cesarini\(^i\), G Ciani\(^i\),\(^dd\),
G Congedo\(^i\),\(^cc\), A M Cruise\(^j\), K Danzmann\(^b\), M de Deus Silva\(^a\), R De Rosa\(^g\), L Di Fiore\(^y\),
M Diaz-Aguiló\(^n\), I Diepholz\(^b\), G Dixon\(^j\), R Dolesi\(^i\), N Dunbar\(^k\), L Ferraioli\(^l\), V Ferroni\(^i\), D Flatscher\(^g\),
BEYOND EINSTEIN: GRAVITATIONAL WAVES FROM EXTENDED GRAVITY

Salvatore Capozziello
Università di Napoli "Federico II" and INFN Sez. di Napoli, Italy

Mariafelicia De Laurentis
Institute for Theoretical Physics, Goethe University, Frankfurt, Germany

Abstract

We show that linearizing Extended Theories of Gravity, further gravitational modes emerge. Besides massless spin-2, also spin-0 and spin-2 massive and ghost fields have to be considered as soon as one is considering the full curvature budget of generic metric theories of gravity. Such additional modes give rise to further polarizations that could be of interest for direct detection by the forthcoming Advanced LIGO-VIRGO and other collaborations.

1 Introduction

The recent discovery of gravitational waves 1) pointed out several new perspectives for some key questions of fundamental physics, astrophysics and cosmology. They range from the validity of Equivalence Principle, to black hole physics to the nature of dark energy and dark matter. Combined gravitational-wave, neutrino and electromagnetic observations can be used to understand the
main characteristics of several astrophysical systems and to map in detail the observed Universe. Furthermore, observations indicate consistent upper bounds on the graviton mass \(^1\) allowing the possibility that further metric theories of gravity can be investigated besides General Relativity (GR).

Given these facts and the lack of a final self-consistent theory of Quantum Gravity, alternative theories of gravity can be pursued as part of a semi-classical approach where GR and its positive results should be retained. The approach of Extended Theories of Gravity (ETG), based on corrections and enlargements of the Einstein theory, has become a sort of paradigm in the study of the gravitational interaction. These theories have received a lot of interest in cosmology since they “naturally” exhibit inflationary and dark energy behaviors \(^2\). At a fundamental level, detecting new gravitational modes could be a sort of *experimentum crucis* in order to discriminate among competing models since this possible detection could be the “signature” that GR should be enlarged, modified or retained as it is \(^3\).

In this report, we discuss the problem of gravitational waves in ETG, showing that new polarizations are derived besides the two standard ones of GR. The theoretical set up of the approach is reported together with some consideration on the actual detectability of such new modes.

## 2 \textit{Gravitational waves in Extended Gravity}

Let us generalize the action of GR by adding curvature invariants other than the standard Ricci scalar. Specifically, we are considering the action \(^1\)

\[
S = \int d^4x \sqrt{-g} f(R, P, Q)
\]

where

\[
P \equiv R_{ab} R^{ab}, \quad Q \equiv R_{abcd} R^{abcd}
\]

In other words, we are taking into account the full curvature budget of generic metric theories of gravity. Varying with respect to the metric, one gets the field equations:

---

\(^1\)\textit{Conventions: } g_{ab} = (-1,1,1,1), \quad R^a_{bcd} = \Gamma^a_{bd,c} - \Gamma^a_{bc,d} + \ldots, \quad R_{ab} = R^c_{acb}, \quad G_{ab} = 8\pi G_N T_{ab} \text{ and all indices run from 0 to 3.}
\[ FG_{\mu\nu} = \frac{1}{2} g_{\mu\nu} (f - R F) - (g_{\mu\nu} \Box - \nabla_\mu \nabla_\nu) F - 2 (f_P R^{a}_{\mu} R_{a\nu} + f_Q R_{abc\mu} R^{abc}_{\nu}) \]
\[ \quad - g_{\mu\nu} \nabla_a \nabla_b (f_P R^{ab}) - \Box (f_P R_{\mu\nu}) + 2 \nabla_a \nabla_b \left( f_P R^a_{(\mu} R^{b)_{\nu}} + 2 f_Q R^a_{(\mu\nu)} \right) \]

(3)

where we have defined
\[ F \equiv \frac{\partial f}{\partial R}, \quad f_P \equiv \frac{\partial f}{\partial P}, \quad f_Q \equiv \frac{\partial f}{\partial Q} \]

and \( \Box = g^{ab} \nabla_a \nabla_b \) is the d’Alembert operator. The notation \( T_{(ij)} = \frac{1}{2} (T_{ij} + T_{ji}) \) denotes symmetrization with respect to the indices \((i, j)\). Considering the trace of eq. (3), we find:

\[ \Box \left( F + \frac{2}{3} (f_P + f_Q) R \right) = \]
\[ \quad = \frac{1}{3} \left[ 2f - RF - 2R^{ab} \nabla_a \nabla_b (f_P + 2f_Q) - R \Box (f_P + 2f_Q) - 2 (f_P P + f_Q Q) \right] \]

If we define
\[ \Phi \equiv F + \frac{2}{3} (f_P + f_Q) R \quad \text{and} \quad \frac{dV}{d\Phi} \equiv \text{RHS of (6)} \]

we get a Klein-Gordon equation for the scalar field \( \Phi \):

\[ \Box \Phi = \frac{dV}{d\Phi} \]

(6)

In order to find the gravitational modes as perturbations, we need to linearize the field around the Minkowski background:

\[ g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \quad \text{and} \quad \Phi = \Phi_0 + \delta \Phi \]

(7)

From eq. (6), we get

\[ \delta \Phi = \delta F + \frac{2}{3} (\delta f_P + \delta f_Q) R_0 + \frac{2}{3} (f_{P0} + f_{Q0}) \delta R \]

(8)

where \( R_0 \equiv R(\eta_{\mu\nu}) = 0 \) and similarly \( f_{P0} = \frac{\partial f}{\partial P}|_{\eta_{\mu\nu}} \) (note that the 0 indicates the value around the Minkowski metric) which is either constant or zero. The first term of eq. (8) is

\[ \delta F = \left. \frac{\partial F}{\partial R} \right|_0 \delta R + \left. \frac{\partial F}{\partial P} \right|_0 \delta P + \left. \frac{\partial F}{\partial Q} \right|_0 \delta Q \]

(9)
However, since $\delta P$ and $\delta Q$ are second order, we get $\delta F \simeq F_{,R0} \delta R$ and

$$
\delta \Phi = \left( F_{,R0} + \frac{2}{3}(f_{P0} + f_{Q0}) \right) \delta R
$$

(10)

Finally, from eq. (6), we get the Klein-Gordon equation for the scalar perturbation $\delta \Phi$

$$
\Box \delta \Phi = \frac{1}{3} \frac{F_0}{F_{,R0} + \frac{2}{3}(f_{P0} + f_{Q0})} \delta \Phi - \frac{2}{3} \delta R a \partial_a \partial_b (f_{P0} + 2f_{Q0}) - \frac{1}{3} \delta R \Box (f_{P0} + 2f_{Q0})
$$

$$
= m^2_2 \delta \Phi
$$

(11)

The last two terms in the first line are actually zero since the terms $f_{P0}, f_{Q0}$ are constants and we have defined the scalar mass as $m^2_2 \equiv \frac{1}{3} \frac{F_0}{F_{,R0} + \frac{2}{3}(f_{P0} + f_{Q0})}$.

Perturbing the field equations (3) and working in Fourier space, we can rewrite the metric perturbation as

$$
h_{\mu\nu} = \bar{h}_{\mu\nu} - \frac{\eta_{\mu\nu}}{2} \eta_{\mu\nu} h_f
$$

(12)

and use the gauge freedom to demand that the standard conditions $\partial_\mu \bar{h}^{\mu\nu} = 0$ and $\bar{h} = 0$ hold. The first of these conditions implies that $k_\mu \bar{h}^{\mu\nu} = 0$ while the second that

$$
h_{\mu\nu} = \bar{h}_{\mu\nu} + \eta_{\mu\nu} h_f \quad \text{and} \quad h = 4h_f
$$

(13)

With these considerations in mind, after some algebra, we get:

$$
\frac{1}{2} \left( k^2 - k^4 \frac{f_{P0} + 4f_{Q0}}{F_0} \right) \bar{h}_{\mu\nu} = (\eta_{\mu\nu} k^2 - k_\mu k_\nu) \frac{\delta \Phi}{F_0} + (\eta_{\mu\nu} k^2 - k_\mu k_\nu) h_f
$$

(14)

Defining $h_f \equiv -\frac{\delta \Phi}{F_0}$, the equation for the perturbations is

$$
\left( k^2 + \frac{k^4}{m^2_{\text{spin}}/2} \right) \bar{h}_{\mu\nu} = 0
$$

(15)

\footnote{It is convenient to work in Fourier space so that, for example, $\partial_\gamma h_{\mu\nu} \rightarrow ik_\gamma \bar{h}_{\mu\nu}$ and $\Box h_{\mu\nu} \rightarrow -k^2 \bar{h}_{\mu\nu}$.}
where we have defined \( m_{\text{spin}2}^2 \equiv -F_0 f_{P0} + 4f_{Q0} \), while from eq. (11), one obtains:

\[
\Box h_f = m_{\text{spin}2}^2 h_f \tag{16}
\]

From equation (15) it is easy to see that a modified dispersion relation is achieved. It corresponds to a massless spin-2 field \((k^2 = 0)\) and a massive spin-2 ghost mode \( k^2 = -F_0 f_{P0} + 4f_{Q0} \equiv -m_{\text{spin}2}^2 \) with mass \( m_{\text{spin}2}^2 \). To see this, note that the propagator for \( \bar{h}_{\mu\nu} \) can be rewritten as

\[
G(k) \propto \frac{1}{k^2} - \frac{1}{k^2 + m_{\text{spin}2}^2} \tag{17}
\]

Clearly the second term has the opposite sign, which indicates the presence of a ghost mode. Also, as a sanity check, we can see that for the Gauss-Bonnet term \( \mathcal{L}_{GB} = Q - 4P + R^2 \) we have \( f_{P0} = -4 \) and \( f_{Q0} = 1 \). Then, eq. (15) simplifies to \( k^2 \bar{h}_{\mu\nu} = 0 \) and, in this case, we have no ghosts as expected. The solution to eqs. (15) and (16) can be written in terms of plane waves

\[
\bar{h}_{\mu\nu} = A_{\mu\nu}(\vec{p}) \cdot \exp(i k^\alpha x_\alpha) + \text{cc} \tag{18}
\]

\[
h_f = a(\vec{q}) \cdot \exp(i q^\alpha x_\alpha) + \text{cc} \tag{19}
\]

where

\[
k^\alpha \equiv (\omega_{m_{\text{spin}2}}, \vec{p}) , \quad \omega_{m_{\text{spin}2}} = \sqrt{m_{\text{spin}2}^2 + p^2} \tag{20}
\]

\[
q^\alpha \equiv (\omega_m, \vec{q}) , \quad \omega_m = \sqrt{m^2 + p^2} .
\]

and where \( m_{\text{spin}2} \) is zero (non-zero) in the case of massless (massive) spin-2 mode. The polarization tensors \( A_{\mu\nu}(\vec{p}) \) can be found in Ref. 4). Eqs. (15) and (18) mean that the standard waves of GR 5) can be obtained, while eqs. (16) and (19) represent further massive gravitational modes 6, 7).

### 3 Polarization states of gravitational waves

Considering the above equations, we can note that there are two conditions for eq. (11) that depend on the value of \( k^2 \). In fact, we have a \( k^2 = 0 \) mode that corresponds to a massless spin-2 field with two independent polarizations plus a scalar mode, while if we have \( k^2 \neq 0 \) we have a massive spin-2 ghost mode.
and there are five independent polarization tensors plus a scalar mode. Taking $\vec{p}$ in the $z$ direction, a gauge where only $A_{11}$, $A_{22}$, and $A_{12} = A_{21}$ are different to zero can be chosen. The condition $\bar{h} = 0$ gives $A_{11} = -A_{22}$. In this frame we can take the bases of polarizations defined as

\[
e^{(+)\mu\nu} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \
e^{(x)\mu\nu} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}
\]

\[
e^{(B)\mu\nu} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \
e^{(C)\mu\nu} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}
\]

\[
e^{(D)\mu\nu} = \frac{\sqrt{2}}{3} \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}, \
e^{(s)\mu\nu} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}
\]

and the amplitude can be written in terms of the 6 polarization states as

\[
h_{\mu\nu}(t,z) = A^{+}(t - v_{G,z}z)e^{(+)\mu\nu} + A^{\times}(t - v_{G,z}z)e^{(x)\mu\nu} \\
+ B^{(B)}(t - v_{G,z}z)e^{(B)\mu\nu} + C^{(C)}(t - v_{G,z}z)e^{(C)\mu\nu} \\
+ D^{(D)}(t - v_{G,z}z)e^{(D)\mu\nu} + h_{s}(t - v_{G}z)e^{(s)\mu\nu}.
\]

(21)

where $v_{G,z}$ is the group velocity of the massive spin-2 field. The terms $A^{+}(t - z)e^{(+)\mu\nu} + A^{\times}(t - z)e^{(x)\mu\nu}$ describe the two standard polarizations of gravitational waves which arise from GR, while the other terms arise from the generic extended models, involving any curvature invariants, that we considered here.

The first two polarizations are the same as in the massless case, inducing tidal deformations on the x-y plane. In Fig.1, we illustrate how each GW polarization affects test masses arranged on a circle.

---

3 The polarizations are defined in our 3-space, not in a spacetime with extra dimensions. Each polarization mode is orthogonal to another one and it is normalized as $e_{\mu\nu}e^{\rho\upsilon} = 2\delta$. Note that other modes are non-traceless, in contrast to the ordinary plus and cross polarization modes of GR.
Figure 1: The six polarization modes of gravitational waves. The picture shows the displacement that each mode induces on a sphere of test particles at the moments of different phases by $\pi$. The wave propagates out of the plane in (a), (b), (c), and it propagates in the plane in (d), (e) and (f). Where in (a) and (b) we have respectively the plus mode and cross mode, in (c) the scalar mode, in (d), (e) and (f) the D, B and C mode.
4 Conclusions

We considered a generic gravitational Lagrangian with any possible combination of curvature invariants. The only assumption is that the gravitational Lagrangian is analytic. We have linearized the field equations around the Minkowski background and found that, besides the massless spin-2 field, there are also spin-0 and spin-2 massive modes with the latter being, in general, ghosts. Then, we have classified the additional polarization modes. However, a point has to be stressed. If the interferometer is directionally sensitive and we also know the orientation of the source (and of course if the source is coherent) the situation is straightforward. In this case, the massive modes would induce longitudinal displacements along the direction of propagation which should be detectable and the amplitude due to the scalar mode would be a possible "new" detectable signal 6). The other modes should be disentangled according to particular features of the sources 7). As a final remark, it is worth noticing that detecting further gravitational modes, besides the two standard of GR, could be a formidable challenge for gravitational physics in view to select the final theory of gravity. In this perspective, Advanced Virgo-LIGO, and the other running GW experiments should be correlated in a sort of global interferometer to investigate polarizations other than the two standard of GR.

References

FUTURE OF THE LHC

Roger Forty
CERN, Geneva, Switzerland

Abstract

The future of the LHC is discussed, considering aspects related to the accelerator, experiments, and physics. This includes the recently-approved high luminosity upgrade, HL-LHC, and beyond.

1 Accelerator

The Large Hadron Collider (LHC) at CERN is the highest energy collider in the world, with 27 km circumference—it has been called the world’s largest scientific instrument. It is primarily a proton-proton collider (although heavier ions can also be accelerated), with design centre-of-mass energy of 14 TeV. The dipole magnets shown in Fig. 1 feature a two-in-one design to allow the acceleration of same-charged particles in both directions around the ring. The superconducting magnets use niobium titanium (NbTi) alloy cable, which gives 8.3 T
field in the main dipoles. They are cooled with liquid helium to a temperature of 1.9 K. The design luminosity is $10^{34} \text{cm}^{-2}\text{s}^{-1}$, achieved with 2808 circulating bunches, each with $\sim 10^{11}$ protons, and focussed at the high-luminosity interaction points with $\beta^* = 40 \text{cm}$, giving a transverse beam size of order $10 \mu\text{m}$.

The bunches are spaced by 25 ns, corresponding to a collision rate of 40 MHz at each of the four interaction points.

The LHC was conceived in the early 1980s, and re-uses the tunnel of the previous machine, LEP ($e^+e^-$, 1989–2000). An incident during commissioning in 2008, due to failure of a magnet interconnect, delayed the start-up. The accelerator restarted at $\sqrt{s} = 7 \text{TeV}$ in 2010 then 8 TeV in 2012. The inte-
grated luminosity is illustrated in Fig. 2, and amounted to $5 + 24 \, \text{fb}^{-1}$ at the high-luminosity experiments (Run 1). The discovery of the Higgs Boson was announced in 2012. This was followed by the first long shutdown (LS1) in 2013–14, during which consolidation was completed for all of the magnet interconnects. The machine restarted in 2015 at 13 TeV. This required a significant number of training quenches for the magnets: up to 50 in the most difficult sector. $4 \, \text{fb}^{-1}$ had been integrated so far in Run 2.

2016 is intended to be a “luminosity production year”, aiming to reach the nominal luminosity of $10^{34} \, \text{cm}^{-2} \text{s}^{-1}$ (that has since been achieved). The accelerator complex had recently recovered from power failure during start-up, caused by a weasel. At the time of the conference $\sim 900$ bunches were circulating, and the machine was ramping up fast to higher luminosity. Pushing up the energy to 14 TeV would require many further training quenches, so it has been decided to stay at 13 TeV for this year (at least). The planned schedule is shown in Fig. 3.

1.1 HL-LHC

According to the European Strategy for Particle Physics, updated in 2013, “Europe’s top priority should be the exploitation of the full potential of the LHC,
including the high-luminosity upgrade of the machine and detectors with a view to collecting ten times more data than in the initial design, by around 2030."

High-luminosity LHC (HL-LHC) is a project to increase the peak luminosity by a factor 5 and integrate 3000 fb\(^{-1}\) of data.

The project is mostly focused on entirely renovating the insertion regions around the high-luminosity experiments (i.e. about 1.2 km of accelerator, as shown in Fig. 4). To achieve stronger focusing the low-beta triplet quadrupoles will be replaced with higher field and larger aperture versions. Low $\beta^*$ requires a larger crossing angle, which would reduce the luminosity by a geometrical factor, so crab cavities will be introduced to rotate bunches to collide head on. Overall this is a 1 BCHF-scale project. R&D on high-field magnets in progress using niobium-tin (Nb\(_3\)Sn) alloy as superconductor: in December 2015 a two-in-one dipole using this material (1.8 m long) reached 11.3 T without quench. This will allow space for extra collimation in dispersion suppressor region.

2 Experiments

There are four major experiments at the LHC. Two are general-purpose high-$p_T$ experiments, ATLAS and CMS, that perform precision studies of the Standard Model (including the new field of Higgs properties) and search for physics beyond the Standard Model. There is an enormous rate of $b$ and $c$ hadrons produced at LHC, dominantly in forward region, and the dedicated flavour experiment, LHCb, has exquisite proper-time resolution (40 fs). Luminosity is levelled for LHCb by adjusting the separation of the beams, as shown in Fig. 5; this technique will be important for HL-LHC. The LHC also accelerates heavy ions (Pb\(^{82+}\)). Typically it is run with Pb-Pb or Pb-p collisions for one month at the end of each year. The dedicated experiment for this physics is
ALICE, although by now all experiments participate. Such collisions allow the properties of matter to be studied at high temperature/density, with a total collision energy of over 1 PeV.

Other smaller experiments include MoEDAL, a monopole search experiment (at the LHCb IP) which surrounds the interaction region with plastic sheets that reveal tracks of highly-ionizing particles after etching. There are also experiments studying forward physics: LHCf (at the ATLAS IP) uses a zero-degree calorimeter to study neutral production, relevant for cosmic rays; TOTEM (at the CMS IP) uses silicon tracking detectors in Roman Pots to study elastic and diffractive scattering of protons.

2.1 Detector upgrades

The recent increase in energy brings less for LHCb and ALICE than high-\(p_T\) experiments, so they have major upgrades planned already for 2019 (during LS2), known as Phase 1. Major upgrades for ATLAS and CMS are foreseen to prepare for HL-LHC in LS3 (Phase 2), with an agreed funding scale of \(\sim 250\) MCHF for each experiment. LHCb and ALICE are also expected to continue during HL-LHC. The major challenges for the high-luminosity phase are the radiation dose, and occupancy/pile-up.

The increase in occupancy will be fought using higher granularity. All experiments will replace their silicon trackers, and CMS is preparing a High-
Figure 6: Higgs couplings vs. mass for the various Standard Model particles: (left) as measured today, (right) precision that will be achieved with HL-LHC.

Granularity Calorimeter in the forward region. Another theme of the detector upgrades concerns increased speed: the ALICE TPC wire chamber readout currently limits their data-taking rate, and will be replaced with GEM end-plates allowing 50kHz readout (i.e. 20× higher). The LHCb signal yield is currently limited for hadronic modes by their first-level trigger; for the upgrade they will remove the hardware trigger and read out the full detector at 40 MHz. This will give an enormous data rate ∼ 5 TB/s, via 12,000 optical links to the CPU farm on the surface. Fast timing detectors are studied by all experiments to fight pile-up: since the beam-spot spreads over ∼ 300 ps, if could be divided into O(25 ps) slices this would reduce the occupancy back to its current level.

3 Physics

The Higgs Boson was discovered in the γγ and ZZ decay modes. ATLAS and CMS results have now been combined, e.g. for the mass measurement: 125.09 ± 0.21 (stat) ± 0.11 (syst) GeV. Alternative spin-parities are disfavoured at over 99.9%—it behaves like Standard Model Higgs, so far (see Fig. 6).

The major focus at the LHC is on the search for physics beyond the Standard Model. Some hints of anomalies were seen in the Run 1 data: in flavour physics, e.g. lepton-flavour violation in $B^0 \rightarrow D(\ast)\tau\nu$, and the angular analysis
Assuming dark matter is made of particles that couple to quarks via a mediator, it may be produced at LHC. It would leave no trace in the detector, so to tag its production a particle is needed from initial state radiation, leading to a monojet search (with missing $E_T$). One can also expect that mediator would couple to quarks in the final state, leading to a dijet resonance search. This will continue to be a very active field at ATLAS and CMS.

LHCb integrated 3 fb$^{-1}$ of data in Run 1 with levelled luminosity of $4 \times 10^{32}$ cm$^{-2}$s$^{-1}$. Precision measurements were made of rare decays and CP violation for many $b$ and $c$ hadrons. In the LHCb upgrade the luminosity will be increased to a few $\times 10^{33}$ cm$^{-2}$s$^{-1}$, aiming to integrate 50 fb$^{-1}$. Examples of the precision expected include $\text{BR}(B_s \rightarrow \mu^+\mu^-)$ at the $10^{-10}$ level; $\phi_s$ (the phase of $B_s$ oscillation) $\pm 0.008$; and the unitarity angle gamma to $\pm 1^\circ$. Another active field is the study of exotic spectroscopy, where there is a zoo of possible exotic hadron states: LHCb has established a pentaquark state

Figure 7: (left) Sensitivity to Supersymmetry in the plane of neutralino vs. chargino mass, for different integrated luminosities. (right) Predicted signal for $H \rightarrow \mu^+\mu^-$ at HL-LHC.
\( p_c(4450)^+ \rightarrow J/\psi p \) using a full angular analysis.

At HL-LHC, if new physics discovered in Run 2 or 3, its first detailed exploration will be possible with well-understood accelerator and experiments. Otherwise, the direct discovery potential will be extended by 20–30\% in mass reach, as illustrated in Fig. 7. In either case over 100 million Higgs Bosons will be produced, allowing the Higgs couplings to be measured to a few percent, and including the 2\(^{\text{nd}}\) generation via the observation of \( H \rightarrow \mu^+\mu^- \).

3.1 Far future

Results from Run 2 will hopefully clarify the best choice for the next energy-frontier machine in time for the next update of the European Strategy in 2019–20. One option is the Future Circular Collider (FCC): a 100 TeV-scale pp collider (with an e\(^+\)e\(^-\) machine as a possible first step), for which the LHC is likely to be reused as injector. Key R\&D for the FCC is to develop 16 T magnets to reach 100 TeV in a 80–100 km tunnel. Using such magnets in the existing tunnel would give \( \sqrt{s} \sim 30 \text{ TeV} \). Investigation of this possible High-Energy LHC (HE-LHC) is now included as part of the FCC study.

4 Conclusions

The LHC at CERN is the flag-ship facility of world-wide particle physics. It has been operating successfully over the last 5 years, providing the Higgs Boson discovery, as well as a vast array of other results: over 1500 scientific publications (and counting). This is a very exciting time for particle physics, as the recent increase in energy is the last such major step for some time, and there are strong hopes for discoveries over the coming years. An upgrade program is in preparation for both machine and experiments, to integrate over 100 times the current data-set, and exploit the LHC to its full potential over the next 20 years. Results from the LHC will play a key role in defining the future direction; the long lead time means that the choice of its successor will need to be made soon.

Acknowledgments

I would like to thank the organizers for the invitation to speak at such an interesting workshop, in such beautiful surroundings.
EXPERIMENTAL GRAVITY TESTS IN THE SOLAR SYSTEM

Manuele Martini, Emanuele Ciocci and Simone Dell’Agnello

Laboratori Nazionali di Frascati dell’INFN

Abstract

Since 1969, Lunar Laser Ranging (LLR) with the Apollo Corner Cube Retroreflectors (CCRs) has supplied almost all significant tests of General Relativity (GR). We performed an analysis of measured LLR data both from stations and dummy observations, using the Planetary Ephemeris Program (PEP), developed by the Center for Astrophysics (CfA). In addition, we are starting to study possible improvements to GR using not only the Moon but also other solar system bodies like Mars. In order to do this, we are simulating ranging measurements to INRRI payload (INstrument for landing-Roving laser Retroreflector Investigation), which is scheduled to reach the Red Planet during October 2016 with ExoMars mission. This kind of simulations will be the stepping stone for Mars test of GR at 1.5 AU with INRRIs, similarly to the Apollo/Lunokhod lunar laser retroreflectors. The work done with the Viking experiments for Mars and with lunar reflectors can be extended. We are studying improvements, progressively achievable with any additional INRRI on Mars surface, on: tests of the spacetime curvature (Post Newtonian Parameter γ), variations of the
gravitational constant and violations of the inverse-square force-law. We have already planned possible improvements in the PEP software (computation of PPN parameter $\eta$, Center of Mass of the Earth and Mars). The simulations show an improvement about one order of magnitude for the general relativity tests for the Moon and Mars.

1 Introduction

Satellite/Lunar laser ranging (SLR/LLR) technique consists in a time-of-flight measurement of short laser pulses fired from ground stations of the International Laser Ranging Service (ILRS) towards payloads of CCRs on the Moon and Satellites, and then retroreflected back to the stations. Nowadays laser ranging technique provides precise, metrologically absolute positioning measurements in the gravitational laboratory of the Sun-Earth-Moon system with a space segment based on cost-effective, passive, maintenance-free payloads. Using LLR is possible to provide one of the best tests for every parameter in table 1.

Table 1: Science constraints using LLR updated to 2004.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>GR value</th>
<th>Up to 2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPN $\beta - 1$</td>
<td>0</td>
<td>$1.1 \times 10^{-4}$</td>
</tr>
<tr>
<td>Weak Equivalence Principle</td>
<td>0</td>
<td>$1.4 \times 10^{-13}$</td>
</tr>
<tr>
<td>Strong Equivalence Principle, $\eta$</td>
<td>0</td>
<td>$4.4 \times 10^{-4}$</td>
</tr>
<tr>
<td>$\dot{G}$</td>
<td>0</td>
<td>$9 \times 10^{-13}$</td>
</tr>
<tr>
<td>Geodetic Precession</td>
<td>0</td>
<td>$6.5 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

We are interested in these kind of parameters because any possible departure from their value in GR would imply $2)$:

- A change in the lunar orbit;
- other gravity theories;
- implications in dark energy and cosmological constant from geodetic precession;
- possible new metrics theories of gravity (from equivalence principle through Nordtvedt parameter $\eta$, $\beta$ and $\gamma$) $1)$. 

28
2 Data analysis

In order to analyze laser ranging data we use the Planetary Ephemeris Program (PEP), developed by the Center for Astrophysics (CfA), by I. Shapiro et al. starting from 1970s. PEP was designed not only to generate ephemerides of the Planets and the Moon, but also to compare model with observations. One of the early uses of this software was the first measurement of the geodetic precession of the Moon 3).

2.1 Physics simulations

2.1.1 Moon

We performed two different numerical simulations using every LLR data available until 2015 (fig.1) plus dummy observations on new retroreflectors on the lunar surface.

![Figure 1: LLR data available until 2015.](image)

All the dummy observations were computed by PEP after defining new CCRs...
positions on the lunar surface and the accuracy of the ranging data. As ground station we use: APOLLO (Apache Point Observatory Lunar Laser-ranging Operation), MLRS (McDonald Laser Ranging Station), CERGA (Centre de recherches en géodynamique et astrométrie) and MLRO (Matera Laser Ranging Observatory). As lunar retroreflector, in addition to the dummy CCRs, we use the real Apollo arrays (11, 14 and 15) and the Lunokhod arrays (1 and 2). For the round-trip accuracies of the dummy data we have used two different sets of parameters: Standard (STD) and Half Standard (H-STD). The round-trip timing uncertainties for the STD set are defined as: 16 ps for APOLLO and 33 ps for other sites on existing reflectors, and 3 ps for APOLLO and 7 ps for other sites on the proposed reflectors. H-STD is simply the half of the previous set (STD).

The expected improvements obtained are shown in table 2. The results shown represent the very pessimistic case in which we are not considering any PEP software update, any station upgrade and only few new CCRs. On the contrary, the progressive optimizations and upgrades of PEP and LLR stations, following progressive deployment of next-generation CCRs on the Moon and analysis of new associated LLR data, will allow for a more significant improvement in the accuracy of the gravity tests, up to a factor 100. We also note that the laser ranging station of the Italian space Agency (ASI) is already undergoing an upgrade for high altitude targets, which will be very useful for LLR science.

<table>
<thead>
<tr>
<th>Test of GR</th>
<th>Improvement (STD)</th>
<th>Improvement (H-STD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPN $\beta - 1$</td>
<td>$\times 8.2$</td>
<td>$\times 12.3$</td>
</tr>
<tr>
<td>$\hat{G}$</td>
<td>$\times 8.8$</td>
<td>$\times 16.9$</td>
</tr>
<tr>
<td>Geodetic Precession</td>
<td>$\times 8.6$</td>
<td>$\times 16.6$</td>
</tr>
</tbody>
</table>

2.2 Mars

In order to study the possibility to use Mars as a test body with laser ranging, we performed a preliminary simulation using five new reflector arrays. The coordinates of these new arrays are the same as Phoenix, Viking 1 & 2, Curiosity and Opportunity missions. As station we used a simulated orbiter around
Mars and for the round-trip accuracies of the simulated data we have used two different, very conservative choices of parameters: in the first case the round-trip timing uncertainties correspond to 100 meters; in the second case they correspond to 10 meters. The preliminary results are shown in fig. 2 and fig. 3.

Figure 2: Improvements on $\beta$.

3 Conclusions

Using new CCRs on the Moon surface we can improve the accuracy of gravity measurements up to two orders of magnitude. For Mars, we are only presenting preliminary and very conservative results, while we are consolidating the overall analysis framework. For Mars, additional work not presented here indicates that improvements in the gravity tests by a factor up to 100 are also possible with a positioning accuracy on the surface of Mars below 10 meters.

References


2. S. Merkowitz et al, A white paper to the planetary science decadal survey.
Figure 3: Improvements on $\frac{\dot{G}}{G}$.


NONMINIMALLY COUPLED GRAVITY
AND PLANETARY MOTION

Riccardo March
Istituto per le Applicazioni del Calcolo, CNR,
Via dei Taurini 19, 00185 Roma, Italy
INFN - Laboratori Nazionali di Frascati (LNF),
Via E. Fermi 40, Frascati 00044 Roma, Italy

Jorge Páramos and Orfeu Bertolami
Departamento de Física e Astronomia and Centro de Física do Porto,
Faculdade de Ciências da Universidade do Porto,
Rua do Campo Alegre 687, 4169-007, Porto, Portugal
Simone Dell’Agnello
INFN - Laboratori Nazionali di Frascati (LNF),
Via E. Fermi 40, Frascati 00044 Roma, Italy

Abstract

The effects of a nonminimally coupled curvature-matter model of gravity on planetary orbits are computed. The parameters of the model are then constrained by the observation of Mercury orbit.

1 Introduction

We consider effects in the Solar System of a model of gravity with a nonminimal coupling (NMC) between geometry and matter 1).

NMC gravity has been applied to several astrophysical and cosmological problems such as, for instance, dark matter 2), dark energy 3), cosmological perturbations 4), post-inflationary reheating 5). For other NMC gravity theories and their potential applications, see 6, 7, 8).
The model admits Minkowski spacetime as a background, and we derive the $1/c$ expansion of the metric assuming a general distribution of matter with mass density, pressure and velocity. The nonrelativistic limit of the model is not Newtonian, but contains a Yukawa correction. We compute a parameterized post-Newton plus Yukawa (PPNY) approximation of the NMC model of gravity. We compute the metric around a static, spherically symmetric body and look for trajectories of a test body around the spherical body. We use the NMC gravity model to compute the perihelion precession of planets, then we constrain the parameters of the model from observation of Mercury orbit.

Further details about the subject of the present communication can be found in the manuscript 9).

2 The NMC gravity model

The action functional of NMC gravity is given by 1)

$$S = \int \left[ \frac{1}{2} f^1(R) + [1 + f^2(R)]L \right] \sqrt{-g} d^4x,$$

where $f^1(R), f^2(R)$ are functions of the spacetime curvature $R$, $g$ is the metric determinant, $L = -\rho c^2$ is the Lagrangian density of matter, and $\rho$ is mass density.

The first variation of the action functional with respect to the metric yields the field equations

$$(f^2_R + 2f^1_R L) R_{\mu\nu} - \frac{1}{2} f^1 g_{\mu\nu} = \nabla^\alpha (f^1_R + 2f^1_R L) + (1 + f^2) T_{\mu\nu},$$

where $f^i_R = df^i/dR$ and $\nabla_{\mu\nu} = \nabla_\mu \nabla_\nu - g_{\mu\nu} g^{\sigma\eta} \nabla_\sigma \nabla_\eta$.

Such equations will be solved by a perturbative method.  

$G$ being Newton’s gravitational constant.

The first variation of the action functional with respect to the metric yields the field equations

$$S = \int \left[ \frac{1}{2} f^1(R) + [1 + f^2(R)]L \right] \sqrt{-g} d^4x,$$

where $f^1(R), f^2(R)$ are functions of the spacetime curvature $R$, $g$ is the metric determinant, $L = -\rho c^2$ is the Lagrangian density of matter, and $\rho$ is mass density.

The function $f^2(R)$ yields a NMC between geometry and matter, and the class of $f(R)$ gravity theories is recovered in the case $f^2(R) = 0$. General Relativity (GR) is recovered by taking:

$$f^1(R) = 2\kappa R, \quad f^2(R) = 0, \quad \kappa = c^4/16\pi G,$$

$G$ being Newton’s gravitational constant.

The first variation of the action functional with respect to the metric yields the field equations

$$\left( f^2_R + 2f^1_R L \right) R_{\mu\nu} - \frac{1}{2} f^1 g_{\mu\nu} = \nabla^\alpha (f^1_R + 2f^1_R L) + (1 + f^2) T_{\mu\nu},$$

where $f^i_R = df^i/dR$ and $\nabla_{\mu\nu} = \nabla_\mu \nabla_\nu - g_{\mu\nu} g^{\sigma\eta} \nabla_\sigma \nabla_\eta$.

Such equations will be solved by a perturbative method.
3 Working assumptions

The metric tensor is assumed as a perturbation of the Minkowski metric $\eta_{\mu\nu}$:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad \text{with } |h_{\mu\nu}| \ll 1,$$

and we look for an $1/c$ expansion of $g_{\mu\nu}$ as in the PPN formalism 10) which, in the diagonal case, is represented by

$$
\begin{align*}
g_{00} &= -1 + h_{00}^{(2)} + h_{00}^{(4)} + O \left( \frac{1}{c^6} \right), \\
g_{ij} &= \delta_{ij} + h_{ij}^{(2)} + O \left( \frac{1}{c^4} \right),
\end{align*}
$$

where $h_{\mu\nu}^{(n)} = O(1/c^n)$, for $n = 2, 4$. Matter will be considered as a perfect fluid with density $\rho$, velocity $v_i$, pressure $p$, and specific energy density $\Pi$:

$$
\begin{align*}
T_{00} &= \rho c^2 \left( 1 + \frac{v^2}{c^2} + \frac{\Pi}{c^2} - h_{00}^{(2)} \right) + O \left( \frac{1}{c^6} \right), \\
T_{ij} &= \rho v_i v_j + p \delta_{ij} + O \left( \frac{1}{c^4} \right).
\end{align*}
$$

3.1 Assumptions on functions of curvature

We assume the functions $f^1(R)$ and $f^2(R)$ to be analytic at $R = 0$, so that they admit the Taylor expansions:

$$
\begin{align*}
f^1(R) &= 2\kappa \sum_{i=1}^{\infty} a_i R^i, \quad a_1 = 1, \\
f^2(R) &= \sum_{j=1}^{\infty} q_j R^j.
\end{align*}
$$

If $a_i = 0$ for any $i > 1$ and $q_j = 0$ for any $j$, then the action of GR is recovered.

The coefficients $a_2, a_3, q_1, q_2$ will be used to compute the metric at the required order and they will be considered as parameters of the NMC gravity model.
4 Perturbative solution of the field equations

At order $O(1/c^2)$ we obtain the following system of equations of Yukawa type for curvature $R^{(2)}$ at second order and the $0-0$ component of the metric:

\[
\nabla^2 R^{(2)} - \frac{R^{(2)}}{6a_2} = \frac{4\pi G}{3c^2a_2} (\rho - 6q_1 \nabla^2 \rho), \quad (10)
\]

\[
\nabla^2 \left( h^{(2)}_{00} - 2a_2 R^{(2)} + \frac{16\pi G}{c^2} q_1 \rho \right) = - \frac{8\pi G}{c^2} \rho. \quad (11)
\]

The solution for $h^{(2)}_{00}$ yields the nonrelativistic limit of the model and consists of the Newtonian potential $U$ plus a Yukawa perturbation $\mathcal{Y}$:

\[
h^{(2)}_{00} = 2 \frac{U}{c^2} + (1 - \theta) \frac{2}{3c^2} \mathcal{Y}, \quad \mathcal{Y} = G \int \rho(t, y) \frac{e^{-m|x-y|}}{|x-y|} d^3y, \quad (12)
\]

where the ratio $\theta$ will play a crucial role:

\[
\theta = q_1 a_2. \quad (13)
\]

The range $\lambda$ and the strength $\alpha_{\text{NR}}$ of the Yukawa potential are given by

\[
\lambda = \frac{1}{m} = \sqrt{6a_2}, \quad \alpha_{\text{NR}} = \frac{1}{3} (1 - \theta). \quad (14)
\]

The range $\lambda$ is the same found for $f(R)$ gravity models 11). If $q_1 = a_2$, then $\theta = 1$ and the Yukawa strength $\alpha_{\text{NR}}$ vanishes. Hence, if $q_1 \simeq a_2$, then long range (astronomical) effects are possible 12).

The solution of the field equations for further components of the metric and higher order then yields a parameterized post-Newton plus Yukawa (PPNY) approximation of NMC gravity.

The $i-j$ (spatial) components of the field equations at order $O(1/c^2)$ are

\[
\nabla^2 \left( \frac{1}{2} h^{(2)}_{ij} - 2a_2 \delta_{ij} R^{(2)} + \frac{16\pi G}{c^2} q_1 \rho \delta_{ij} \right) + \frac{1}{2} \delta_{ij} R^{(2)} + 2a_2 R^{(2)}_{ij} = \frac{c^2}{\kappa} q_1 \rho_{ij}, \quad (15)
\]

and the solution, in diagonal form after a gauge transformation, is given by

\[
h^{(2)}_{ij} = \left[ 2 \frac{U}{c^2} - (1 - \theta) \frac{2}{3c^2} \mathcal{Y} \right] \delta_{ij}. \quad (16)
\]
If $q_1 \simeq a_2$ we have PPN parameter $\gamma \simeq 1.12$, while in the case $f^2(R) = 0$ (pure $f(R)$ gravity) it is known that $\gamma = 1/2$ if the Yukawa potential is long range.

The solution for $h^{(4)}_{00}$ is long and we give the leading term:

$$h^{(4)}_{00} = h^{(4)-GR}_{00} + \frac{8\pi G^2}{c^4} \theta \left( -2q_1 + \frac{a_2 q_1}{a_2} - \frac{4}{3} \frac{q_2}{a_2} \right) X(\rho^2) + \ldots,$$

where $h^{(4)-GR}_{00}$ is the contribution from GR,

$$X(\rho^2) = \int \rho^2(t, y) \frac{e^{-m|x-y|}}{|x-y|} d^3y,$$

and dots... denote the contribution from further potentials which can be found in the manuscript.

4.1 Metric around a static, spherical body

Using the PPNY metric, given in the previous section, the metric in vacuum around a spherical body (Sun) can be computed. Under the assumption of a static, uniform mass density $\rho$, the metric is the following ($g_{0i} = 0$):

$$g_{00} = -1 + 2\frac{GM_S}{rc^2} \left( 1 + \alpha e^{-r/\lambda} \right) + \frac{2}{c^4} F(r),$$

$$g_{ij} = \left[ 1 + 2\frac{GM_S}{rc^2} \left( 1 - \alpha e^{-r/\lambda} \right) \right] \delta_{ij},$$

where $M_S$ is the mass of the spherical body, $r$ is the radial coordinate, and $F(r)$ is a radial potential which can be found in the manuscript. The strength $\alpha$ of the Yukawa potential is computed for $\lambda \gg R_S$, where $R_S$ is the radius of the spherical body. The leading term of $\alpha$ is the following:

$$\alpha = \frac{1}{3} (1 - \theta) + \frac{GM_S}{c^2 R_S^2} \left[ \theta \left( \frac{\mu}{2} - 1 \right) - \frac{2}{3} \nu \right] \left( \frac{\lambda}{R_S} \right)^2 + \ldots,$$

where $\mu, \nu$ are the following dimensionless ratios: $\mu = a_3/a_2^2, \nu = q_2/a_2^2$, and dots... denote further contributions which can be found in 9).
5 Planetary precession

By using the metric around a spherical body (Sun) the effect of NMC gravity on the orbit of a planet is computed. In NMC gravity the energy-momentum tensor is not covariantly conserved
\[ \nabla_\mu T^{\mu\nu} = \frac{f_2}{f_1 + f_2} (g^{\mu\nu} L - T^{\mu\nu}) \nabla_\mu R \neq 0 \quad \text{if} \ f^2(R) \neq 0, \tag{21} \]
consequently, the trajectories deviate from geodesics:
\[ \frac{d^2x^\alpha}{ds^2} + \Gamma^\alpha_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} = \frac{f_2^2(R)}{1 + f^2(R)} g^{\alpha\beta} R_{,\beta}. \tag{22} \]
Moreover, geodesics are different from GR. The orbit of a planet around the Sun is computed under the assumption of a long range Yukawa perturbation, \( \lambda \gg L \), where \( L \) is the semilatus rectum of the unperturbed orbit of the planet.

Perihelion precession is computed by means of the variation of the Runge-Lenz vector \( \Delta \) along the line perpendicular to both \( A \) and the angular momen-
tum (per mass) \( h = r \times v \):
\[ \frac{d\phi_P}{dt} = (h \times \Delta)^2, \quad \frac{d\Delta}{dt} = \Delta \times h + v \times (r \times \Delta), \tag{23} \]
where \( \phi_P \) denotes the angular coordinate of perihelion in the orbital plane, and \( \Delta \) is the perturbation of the Newtonian force. An integration then yields
\[ \delta \phi_P = \int_0^{2\pi} \frac{d\phi_P}{dt} \frac{L^2}{|h| [1 + e \cos(\phi_P - \phi_p)]} d\phi, \tag{24} \]
where \( e \) is the eccentricity of the orbit.

The leading term in formula for perihelion precession of a planet is the following:
\[ \delta \phi_P = \frac{6\pi GM_S L c^2}{L^2 c^2} + (1 - \theta)^2 \frac{2\pi}{3} \left( \frac{L}{R_S} \right)^2 e^{-L/\lambda} \tag{25} \]
\[ + \ (1 - \theta) \frac{\pi GM_S}{3 L c^2} \theta \left[ 3\theta \left( \frac{\mu}{2} - 1 \right) - 2\nu \right] \left( 1 - \frac{L}{R_S} \right)^3 \left( \frac{L}{R_S} \right)^3, \]
where the terms in the first row are the GR precession and the nonrelativistic Yukawa precession, respectively, and the term in the second row is the leading contribution from the NMC relativistic correction. Dots \( \ldots \) denote further contributions which can be found in the manuscript.

Note that the above formula reduces to the GR expression if \( \theta = 1 \).
6 Constraints on parameters of the NMC gravity model

The prediction for perihelion precession assuming a PPN metric 10) is given by

\[
\delta \phi_P = \left[ \frac{2(1 + \gamma) - \frac{\beta}{3} + 3 \times 10^3 J_2}{6 \pi GM_S \frac{Lc^2}{\lambda^2}} \right] \epsilon_1, \quad (26)
\]

where \( \gamma, \beta \) are PPN parameters, and \( J_2 \) is the quadrupole moment of the Sun. The bounds on \( \gamma \) from Cassini experiment and bounds on \( \beta \) from fits to planetary data, including data from Messenger spacecraft 14) orbiting around Mercury, yield

\[
\gamma - 1 = (2.1 \pm 2.3) \times 10^{-5}, \quad \beta - 1 = (-4.1 \pm 7.8) \times 10^{-5}. \quad (27)
\]

Using such bounds we find that the additional perihelion precession due to NMC deviations from GR, in the case of Mercury orbit, is bounded by

\[
-5.87537 \times 10^{-4} < \delta \phi_P - \delta \phi_P \text{ GR} < 2.96635 \times 10^{-3}. \quad (28)
\]

These inequalities define an admissible region in the four-dimensional parameter space with dimensionless coordinates

\[
\theta = \frac{q_1}{a_2}, \quad \mu = \frac{a_3}{a_2}, \quad \nu = \frac{q_2}{a_2}, \quad \frac{R_S}{\lambda}. \quad (29)
\]

Exclusion plots obtained by slicing the admissible region with two-dimensional planes can be found in the manuscript 9).

We remark that the admissible region in three-dimensional parameter subspace with coordinates \((\theta, \mu, \nu)\), for \(0 < |1 - \theta| \ll 1\) and a given \(\lambda \gg L\), can be approximated by the region enclosed within the degenerate quadric surfaces

\[
\nu = \frac{3}{4} \mu - \frac{3}{2} - 9 \left( \frac{R_S}{L} \right)^3 \frac{\varepsilon_i}{(1 - L/\lambda)(1 - \theta)}, \quad (30)
\]

with \(i = 1, 2\), and

\[
\varepsilon_1 \frac{6 \pi GM_S}{Lc^2} = -5.87537 \times 10^{-4}, \quad \varepsilon_2 \frac{6 \pi GM_S}{Lc^2} = 2.96635 \times 10^{-3}. \quad (31)
\]

Using such an approximation it follows that the intersection of the three-dimensional admissible subregion with a plane \(\theta = \text{constant}\), with \(0 < |1 - \theta| \ll 1\), is a strip enclosed between two lines in the \((\mu, \nu)\) plane. The intersections
with the planes $\mu = constant$ and $\nu = constant$ are regions enclosed by pairs of hyperbolae \(^9\).

We conclude by observing that the forthcoming BepiColombo mission to Mercury should allow a reduction on the above bounds by approximately one order of magnitude.

7 Acknowledgements

The work of R.M. is partially supported by INFN (Istituto Nazionale di Fisica Nucleare, Italy), as part of the MoonLIGHT-2 experiment in the framework of the research activities of the Commissione Scientifica Nazionale n. 2 (CSN2).

References

BLACK HOLES: A WINDOW INTO STRONG (QUANTUM?) GRAVITY

Roberto Casadio

Dipartimento di Fisica e Astronomia, Università di Bologna,
via Irnerio 46, 40126 Bologna, Italy
I.N.F.N., Sezione di Bologna, IS - FLAG,
via B. Pichat 6/2, 40127 Bologna, Italy

Abstract

Black holes are regions of space-time where gravity becomes so strong to confine everything. Their classical general relativistic description however shows critical aspects when faced with the established quantum nature of matter. Alternative approaches and descriptions, like the horizon quantum mechanics and corpuscular models, have therefore been proposed in order to investigate their quantum structure, and search for new phenomenological signatures.

1 Gravitational collapse of quantum matter

The classical, general relativistic description of the gravitational collapse of a compact, massive object predicts the end-point of its evolution will be a space-time singularity, where the energy density diverges (along with tidal forces), provided a trapping surface forms at some point during the collapse and the weak energy condition is preserved all along 1). In other words, if a black hole forms, general relativity predicts there is going to be a real singularity at its centre (see left panel in Fig. 1). However, matter is quantum, and such a singularity simply clashes with the Heisenberg uncertainty principle. One also gets a flavour of the sort of effects that the quantum nature of matter implies, for example, from the famous Hawking’s discovery of black hole evaporation 2): the space-time around the collapsing matter evolves in time and particles are produced in the vacuum state of any quantum field on such a background (see right panel in Fig. 1).

The Hawking effect has raised a number of concerning paradoxes about the possibility of building a consistent quantum description of gravity. Most notably, the prediction that information stored in the initial state of the collapsing star will go lost after the complete evaporation of the hole hinders the unitarity of the whole process. However, one should notice that the Hawking effect is derived by quantising small perturbations around the classical model of the collapse, which
leaves us hope that a fully quantum treatment of the whole matter-gravity system will solve such issues. Of course, the big missing piece is now “quantum gravity”.

2 Quantum gravity and black holes

It is common lore that quantum gravity should become relevant at the Planck length and mass
\[ \ell_p = \sqrt{\hbar G_N} \simeq 10^{-35} \text{ m} \quad \text{and} \quad m_p = \sqrt{\hbar G_N} \simeq 10^{19} \text{ GeV}. \] (1)

This argument is not mere numerology, but follows from the classical key concept of the gravitational radius of a static and spherically symmetric self-gravitating source, for which this quantity determines the existence of horizons.

A static spherically symmetric metric can always be written as
\[ ds^2 = g_{ij}(r) dx^i dx^j + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) , \] (2)

where \( x^i = (t, r) \) and the area of a sphere parameterised by \( \theta \) and \( \phi \) is \( A = 4 \pi r^2 \). The location of a horizon is then determined by the vanishing of the null geodesic expansion,
\[ g^{ij} \nabla_i r \nabla_j r = g^{rr} = 0 . \] (3)

Moreover, Einstein equations yield \( g^{rr} = 1 - r_H(r)/r \), where \( r_H(r) = 2 \ell_p m(r)/m_p \) is the gravitational radius determined by the Misner-Sharp mass function
\[ m(r) = 4 \pi \int_0^r \rho(\bar{r}) \bar{r}^2 d\bar{r} , \] (4)

with \( \rho = \rho(r) \) the static matter density. A horizon then exists where the gravitational radius satisfies \( r_H(r) = r \), for some \( r > 0 \). In the vacuum far outside the region where the source is located, the Misner-Sharp mass approaches the Arnowitt-Deser-Misner (ADM) mass of the source, \( m(r) \to M \), and the gravitational radius likewise becomes the Schwarzschild radius
\[ R_H = 2 \ell_p \frac{M}{m_p} . \] (5)

\(^1\)I will use units with the speed of light \( c = 1 \).
The Heisenberg principle of quantum mechanics introduces an uncertainty in the particle’s spatial localisation of the order of the Compton-de Broglie length, $\lambda_M \simeq \ell_p m_p/M$. It follows that one expects $R_H$ only makes sense if

$$R_H \gtrsim \lambda_M \quad \text{or} \quad M \gtrsim m_p,$$

which immediately explains the key physical role played by the Planck scale and its relation with the existence of black holes in the quantum theory.

Our present knowledge is summarised in Fig. 2, where the relevant parameter is the energy density $\rho$ in units of the Planck density $\rho_p$. We live in a region of extremely low $\rho/\rho_p$, in the bottom right corner, where the quantum field theoretical Standard Model of particles and classical general relativity describe very well our world. The region denoted by QG is where we meet both Planck length and mass, for which we presumably need a full quantum theory of gravity. From this region starts the line corresponding to black holes, moving up along which the energy density inside the horizon decreases and a black holes with larger mass $E$ appear more and more classical. The yellow disk represents the starting point of, say, a collapsing star, which should produce a black hole (the black dot) according to general relativity. There are clearly two possibilities: either the star becomes a black hole by evolving through classical gravitational configurations (green line) or going through the quantum gravity region (blue line). The important point is that, according to classical general relativity, once the black hole forms, the matter in the star is forced to further contract and enter the quantum gravity regime. The conclusion is therefore that black holes are unavoidable and they make quantum gravity necessary as well.

If we look at black holes as bound states of gravity, we can draw an analogy with the (better understood) non-linear QCD theory:

1. like QCD confines quarks and gluons below the scale $\Lambda_{QCD} \simeq 220$ MeV, Einstein gravity confines everything within a horizon $R_H$. Both effects occur in the non-perturbative regime.
of the respective theories and there is no reason to believe that understanding black hole formation is going to be any easier than the (still open) problem of solving QCD;

2. like QCD becomes asymptotically free at energies much above $\Lambda_{\text{QCD}}$, black holes should become asymptotically classical for $M \gg m_p$.

A perhaps overlooked difference is that we have plenty of experimental data supporting the previous two points in QCD, whereas we have practically no data from the strong regime of gravity (beside the recently detected gravitational waves), which makes the above considerations about black holes purely theoretical expectations. Nonetheless, like one can consider effective descriptions of QCD around the scale of confinement, we could also envisage attempting at a quantum description of specific quantities of physical relevance for black holes, rather than insisting in deriving their properties from a candidate general theory of quantum gravity. In the following we shall describe one of such attempts.

3 Horizon Quantum Mechanics

As matter sources are described by quantum physics, the quantities that define the ADM mass $M$ should also be considered as quantum variables, and the Horizon Quantum Mechanics (HQM) was precisely proposed in order to describe the Schwarzschild radius $M$ quantum mechanically $^3$. It is important to emphasise that the HQM differs from most previous attempts in which the gravitational degrees of freedom of the horizon, or of the black hole metric, are quantised independently of the state of the source. In the HQM, the gravitational radius is instead quantised together with the matter source, which is more akin to the non-linear general relativistic description of the gravitational interaction in the strong regime, and to DeWitt’s mini-superspace approach $^4$.

We restrict our analysis to spherically symmetric sources which are both localised in space and at rest in the chosen reference frame. Let $\alpha$ denote the set of quantum numbers parametrising the spectral decomposition of the source, and write a matter state as

$$|\psi_S\rangle = \sum_\alpha C_S(E_\alpha)|E_\alpha\rangle, \quad (7)$$

where the sum is over the eigenstates of a given Hamiltonian $H$, 

$$\hat{H} = \sum_\alpha E_\alpha |E_\alpha\rangle\langle E_\alpha|. \quad (8)$$

We can then replace the ADM mass with the expectation value of this Hamiltonian,

$$M \rightarrow \langle \psi_S |\hat{H}| \psi_S \rangle = \langle \psi_S | \sum_\alpha E_\alpha |E_\alpha\rangle\langle E_\alpha | \psi_S \rangle = \sum_\alpha |C_S(E_\alpha)|^2 E_\alpha. \quad (9)$$

We also introduce the gravitational radius eigenstates

$$\hat{R}_H|R_{H\beta}\rangle = R_{H\beta}|R_{H\beta}\rangle, \quad (10)$$

so that a physical state for our system can be described by linear combinations

$$|\Psi\rangle = \sum_{\alpha,\beta} C(E_\alpha, R_{H\beta}) |E_\alpha\rangle|R_{H\beta}\rangle \quad (11)$$

which satisfy the algebraic (Hamiltonian) constraint (5), that is

$$0 = \left(\hat{H} - \frac{m_p^2}{2\ell_p}\hat{R}_H\right)|\Psi\rangle = \sum_{\alpha,\beta} \left(E_\alpha - \frac{m_p^2}{2\ell_p}R_{H\beta}\right)C(E_\alpha, R_{H\beta}) |E_\alpha\rangle|R_{H\beta}\rangle. \quad (12)$$
The solution is clearly given by
\[ C(E_\alpha, R_{H\beta}) = C(E_\alpha, 2 \ell_p E_\alpha/m_p) \delta_{\alpha\beta}, \]
where \( \delta_{\alpha\beta} \) is the identity in the space of our quantum numbers.

By tracing out the gravitational radius, we must recover the matter state (7), which implies
\[ C(E_\alpha, 2 \ell_p E_\alpha/m_p) = C_S(E_\alpha). \]

Likewise, by integrating out the matter states, we obtain the horizon wave-function
\[ |\psi_H\rangle = \sum_\alpha C_S(m_p R_{H\alpha}/2 \ell_p) |R_{H\alpha}\rangle, \]
where \( m_p R_{H\alpha}/2 \ell_p = E(R_{H\alpha}). \) In the continuum, the normalised wave-function
\[ \psi_H(R_{H}) = \langle R_{H} | \psi_H \rangle = N_H C_S(m_p R_{H}/2 \ell_p) \]
yields the probability to detect a gravitational radius of size \( R_{H} \) associated with the particle in the quantum state \( |\psi_S\rangle \). We can further define the conditional probability density that the particle lies inside its own gravitational radius as
\[ P_\prec(R_{H}) = P_S(R_{H}) P_H(R_{H}) \]
where
\[ P_S(R_{H}) = 4 \pi \int_0^{R_{H}} |\psi_S(r)|^2 r^2 dr \]
is the usual probability that the particle is found inside a sphere of radius \( r = R_{H} \), and
\[ P_H(R_{H}) = 4 \pi R_{H}^2 |\psi_H(R_{H})|^2 \]
is the probability density that the value of the gravitational radius is \( R_{H} \). One can view \( P_\prec(R_{H}) \) as the probability density that the sphere \( r = R_{H} \) is a trapping surface, and the probability that the particle is a black hole (regardless of the horizon size) will be obtained by integrating (17),
\[ P_{BH} = \int_0^{\infty} P_\prec(R_{H}) dR_{H}, \]
which will depend on the observables and parameters of the specific matter state \( |\psi_S\rangle \).

3.1 Single particle and GUP

Let us consider a massive particle at rest in the origin of the reference frame described by the spherically symmetric Gaussian wave-function
\[ \psi_S(r) = e^{-\frac{r^2}{\ell^2}} \frac{1}{\ell^{3/2} \pi^{3/4}}, \]
with \( \ell \simeq \lambda_m \simeq \ell_p m_p/m \). The corresponding momentum space wave-function
\[ \psi_S(p) = \frac{45}{\Delta^{3/2} \pi^{3/4}} e^{-\frac{p^2}{\Delta^2}}, \]
has a width $\Delta = m_p \ell_p / \ell \simeq m$. We can also assume the relativistic mass-shell relation in flat space, $E^2 = p^2 + m^2$, which yields the normalized horizon wave-function

$$\psi_H(R_H) = \frac{\ell_3}{2} e^{-m_p^2 p R_H^2} \frac{m_p^2}{\ell_2} \frac{\Delta p}{m_p^2}. \quad (23)$$

From the plot of the corresponding $P_{\text{BH}}$ in the left panel of Fig. 3, it appears pretty obvious that the particle is most likely a black hole if $m \gtrsim m_p$, in agreement with the qualitative result (6).

For the state (21), the uncertainty in radial size is given by

$$\Delta r^2 \simeq \ell_2^2 \simeq \frac{m_p^2}{\Delta p^2} \ell_p. \quad (24)$$

Analogously, the uncertainty in the horizon radius will be given by

$$\Delta R_H^2 \simeq \frac{\ell_4^2}{\ell_2^2} \simeq \frac{\Delta p^2}{m_p^2} \ell_p, \quad (25)$$

which, combined linearly with Eq. (24), yields the generalised uncertainty relation

$$\Delta r = \Delta r + \gamma \Delta R_H \simeq \ell_p \frac{m_p}{\Delta p} + \gamma \ell_p \frac{\Delta p}{m_p}. \quad (26)$$

From the plot in the right panel of Fig. 3 (for $\gamma = 1$), one can see there is a minimum measurable length $\Delta r \gtrsim 1.3 \sqrt{\ell_p}$ obtained for $\Delta p \simeq m_p$.

A crucial observation is that $\Delta R_H \sim m \sim R_H$, which seems to imply that the horizon of very massive sources fluctuate wildly, contrary to the expectation that astrophysical black holes should be classical objects. This leads us to consider alternative models of black holes, whose source is not localised within a very narrow wave-function (limiting to a point-like singularity).

### 3.2 BEC black holes

In the corpuscular model introduced by Dvali and Gomez $5^\text{)}$, black holes are bound states of gravitons of spatial size $R_H$, effectively forming a Bose-Einstein condensate (BEC) at a critical point. This picture emerges by considering the Newtonian potential generated by a star of mass $M$ as made of $N$ (virtual) gravitons of effective mass $m \simeq m_p \ell_p / \lambda_m$,

$$V_N(r) \simeq -\frac{G_N M}{r} = -\frac{\ell_p N m}{r m_p}. \quad (27)$$
After the star collapses to form a black hole \(6\), these gravitons are contained within a ball of radius \(r \simeq R_H \simeq \lambda m\) and must be (at least) “marginally bound”, that is \(5\)
\[
E_K + U_m \simeq 0 .
\]  
(28)
where \(E_K \simeq m\) and the average potential energy per graviton is
\[
U_m \simeq m V_N(\lambda m) := -N \alpha m ,
\]  
(29)
with the effective gravitational coupling \(\alpha = \ell^2_p / \lambda^2_m = m^2 / m^2_p\). When the condition \(28\) is reached, the gravitons are “maximally packed”, and their number satisfies \(N \alpha \simeq 1\). The effective graviton mass correspondingly scales as \(m \simeq m_p / \sqrt{N}\), while the total mass of the black hole scales like
\[
M = N m \simeq \sqrt{N} m_p .
\]  
(30)
Moreover, the horizon area is spontaneously quantised as expected \(7\), that is
\[
4 \pi R^2_H \simeq \lambda^2_m \simeq \ell^2_p N .
\]  
(31)
This BEC black hole will emit gravitational Hawking radiation, since reciprocal \(2 \to 2\) graviton scatterings inside the condensate give rise to a depletion rate
\[
\dot{N} \sim -\frac{1}{N^2} \frac{N^2}{\sqrt{N}} \frac{1}{\ell^2_p} ,
\]  
(32)
where the factor \(N^{-2}\) comes from \(\alpha^2\), the \(N^2\) factor is combinatoric, and the last factor comes is the characteristic energy of the process \(\Delta E \sim m\). This rate reproduces the standard decay law
\[
\dot{M} \simeq m_p \frac{\dot{N}}{\sqrt{N}} \sim -m_p \frac{N}{N^2} \frac{1}{\ell^2_p M^2} ,
\]  
(33)
and allows one to read off the “effective” Hawking temperature
\[
T_H \simeq \frac{m^2_p}{8 \pi M} \sim m \sim \frac{m_p}{\sqrt{N}} .
\]  
(34)
A more refined model was analysed in Refs. \(8\), in which we introduced candidate quantum states for both the BEC black hole and the emitted Hawking quanta. Such states were analysed using the HQM and their horizon uncertainty decreases for larger \(N\),
\[
\frac{\Delta R_H}{R_H} \simeq \frac{1}{N} ,
\]  
(35)
which shows that such extended models of black holes correctly reproduce the expected behaviour in the macroscopic limit \(N \simeq M/m_p \gg 1\).

4 Summary and outlook

Given the difficulty in conceiving a full quantum theory of gravity, one can focus on a quantum description of particularly relevant quantities for specific problems. The HQM is precisely such an attempt for the gravitational radius of a matter source, which Einstein theory teaches us is a crucial quantity in black hole formation. This approach was applied to many different situations \(3, 8, 9\) and will be further investigated and extended in the future.
Acknowledgements

It is a real pleasure to thank G. Dvali, A.Y. Kamenshchik, W. M"uck, G. Venturi for many discussions and all of my collaborators in this project, X. Calmet, R.T. Cavalcanti, R. Da Rocha, A. Giugno, A. Giusti, O. Micu, J. Mureika, A. Orlandi, F. Scardigli, D. Stojkovic.

References


HOLOGRAPHIC THEORY OF GRAVITY AND COSMOLOGY

Y. Jack Ng
Institute of Field Physics, Department of Physics & Astronomy,
University of North Carolina, Chapel Hill, NC 27599-3255, USA

Abstract

According to the holographic principle, the maximum amount of information stored in a region of space scales as the area of its two-dimensional surface, like a hologram. We show that the holographic principle can be understood heuristically as originated from quantum fluctuations of spacetime. Applied to cosmology, this consideration leads to a dynamical cosmological constant $\Lambda$ of the observed magnitude, in agreement with the result obtained by using uni-modular gravity and causal-set theory for the present and recent cosmic epochs. By generalizing the concept of entropic gravity, we find a critical acceleration parameter related to $\Lambda$ in galactic dynamics, and we construct a phenomenological model of dark matter which we call “modified dark matter” (MDM). We provide successful observational tests of MDM at both the galactic and cluster scales. We also discuss the possibility that the quanta of both dark energy and dark matter obey the quantum Boltzmann statistics or infinite statistics as described by a curious average of the bosonic and fermionic algebras.
1 Introduction and Summary

In Vulcano 2004, in a talk titled “Space-time fluctuations,” I discussed some aspects of "space-time foam" – a foamy structure of spacetime arising from quantum fluctuations. 1) To examine how large the fluctuations are, I considered a gedanken experiment in which a light signal is sent from a clock to a mirror (at a distance $l$ away) and back to the clock in a timing experiment to measure $l$. From the jiggling of the clock’s position alone, the Heisenberg uncertainty principle yields $\delta l^2 \gtrsim \frac{\hbar}{m c}$, where $m$ is the mass of the clock. On the other hand, the clock must be large enough not to collapse into a black hole; this requires $\delta l \gtrsim \frac{Gm}{c^2}$. We conclude that the fluctuation of a distance $l$ scales as $\delta l \gtrsim l^{1/3} l_P^{2/3}$ (where $l_P = \sqrt{\hbar G/c^3}$ is the Planck length). 2) I also showed that this scaling of $\delta l$ is what the holographic principle 3) demands.

The present talk is a continuation of the talk I gave twelve years ago. I will start (in Section 2) by rederiving this scaling of $\delta l$ by another method 4) which can be generalized to the case of an expanding universe for which a dynamical cosmological constant is shown to emerge, 5) a result that was earlier obtained 6) by a consideration (in Section 3) of unimodular gravity 7) and Sorkin’s causal-set theory. This led me to my more recent work with Ho and Minic, and later also with Edmonds, Farrah and Takeuchi. We found it natural (see Section 4) to generalize Verlinde’s formulation 8) of entropic gravity/gravitational thermodynamics to de-Sitter space with a positive cosmological constant. The result was a dark matter model which we call modified dark matter (MDM). 9) Recently we have successfully tested MDM (see Section 5) with 30 galactic rotation curves and a sample of 93 galactic clusters. 10)

The take-home message from this talk is this: It is possible that the dark sector (dark energy and dark matter) has its origin in quantum gravity. And if the scenario to be sketched in Section 6 is correct, then we can expect some rather novel particle phenomenologies, for the quanta of the dark sector obey not the familiar Bose or Fermi statistics, but an exotic statistics that goes by the name infinite statistics 11) or quantum Boltzmann statistics. 12)

I would like to take this opportunity to make a disclaimer: In a recent paper “New Constraints on Quantum Gravity from X-ray and Gamma-Ray Observations” by Perlman et al. (ApJ. 805, 10 (2015)), it was claimed that detections of quasars at TeV energies with ground-based Cherenkov telescopes...
seem to have ruled out the holographic spacetime foam model (with $\delta l$ scaling as $l^{1/3}l_p^{2/3}$). But now I (one of the authors) believe this conclusion is conceivably premature when proper averaging is carried out (though presently there is no formalism yet for carrying out such averages.)

2 Spacetime Foam and the Cosmological Constant $\Lambda$

We can rederive the scaling of $\delta l$ by another argument. Let us consider mapping out the geometry of spacetime for a spherical volume of radius $l$ over the amount of time $2l/c$ it takes light to cross the volume. $^4$ One way to do this is to fill the space with clocks, exchanging signals with the other clocks and measuring the signals’ times of arrival. The total number of operations, including the ticks of the clocks and the measurements of signals, is bounded by the Margolus-Levitin theorem which stipulates that the rate of operations cannot exceed the amount of energy $E$ that is available for the operation divided by $\pi\hbar/2$. This theorem, combined with the bound on the total mass of the clocks to prevent black hole formation, implies that the total number of operations that can occur in this spacetime volume is no bigger than $2(l/l_p)^2/\pi$. To maximize spatial resolution, each clock must tick only once during the entire time period. If we regard the operations as partitioning the spacetime volume into “cells”, then on the average each cell occupies a spatial volume no less than $\sim l^3/(l^2/l_p^2) = l l_p^2$, yielding an average separation between neighboring cells no less than $\sim l^{1/3}l_p^{2/3}$.

This spatial separation can be interpreted as the average minimum uncertainty in the measurement of a distance $l$, that is, $\delta l \sim l^{1/3}l_p^{2/3}$.

It is straightforward to generalize $^5, ^{12}$ the above discussion for a static spacetime region with low spatial curvature to the case of an expanding universe by the substitution of $l$ by $H^{-1}$ in the expressions for energy and entropy densities, where $H$ is the Hubble parameter. (Henceforth we adopt $c = 1 = \hbar$ for convenience unless stated otherwise for clarity.) Applied to cosmology, the above argument leads to the prediction that (1) the cosmic energy density has the critical value $\rho \sim (H/l_p)^2$, and (2) the universe of Hubble size $R_H$ contains $I \sim (R_H/l_p)^2$ bits of information. It follows that the average energy carried by each particle/bit is $\rho R_H^3/I \sim R_H^{-1}$. Such long-wavelength constituents of dark energy give rise to a more or less spatially uniform distribution of cosmic energy density and act as a dynamical cosmological constant with the observed small but nonzero value $\Lambda \sim 3H^2$. 
3 Quantum (Generalized Unimodular) Gravity and (Dynamical) Λ

The dynamical cosmological constant we have just obtained will be seen to play an important role in our subsequent discussions. So let us “rederive” it by using another method based on quantum gravity. The idea makes use of the theory of unimodular gravity, more specifically its generalized action given by

$$S_{\text{unimod}} = -(16\pi G)^{-1} \int \sqrt{g} (R + 2\Lambda) - 2\Lambda \partial_{\mu} T^{\mu} (d^3x)dt.$$ 

In this theory, $\Lambda/G$ plays the role of “momentum” conjugate to the “coordinate” $\int d^3x T^0$ which can be identified as the spacetime volume $V$. Hence the fluctuations of $\Lambda/G$ and $V$ obey a quantum uncertainty principle, $\delta V \delta \Lambda/G \sim 1$.

Next we borrow an argument due to Sorkin, drawn from the causal-set theory, which stipulates that continuous geometries in classical gravity should be replaced by “causal-sets”, the discrete substratum of spacetime. In the framework of the causal-set theory, the fluctuation in the number of elements $N$ making up the set is of the Poisson type, i.e., $\delta N \sim \sqrt{N}$. For a causal set, the spacetime volume $V$ becomes $l_p^4 N$. It follows that $\delta V \sim l_p^4 \delta N \sim l_p^4 \sqrt{N} \sim l_p^4 \sqrt{V} = G \sqrt{V}$, and hence $\delta \Lambda \sim V^{-1/2}$. By following an argument due to Baum and Hawking, we argued that, in the framework of unimodular gravity, $\Lambda$ vanishes to the lowest order of approximation and that its first order correction is positive (at least for the the cosmic epoch corresponding to redshift $z \lesssim 1$). See the second paper of Ref. 6). We conclude that $\Lambda$ is positive with a magnitude of $V^{-1/2} \sim R^{-2}_H$, contributing a cosmic energy density $\rho$ given by:

$$\rho \sim \frac{1}{l_p^2 R_H^2},$$

which is of the order of the critical density as observed!

4 From $\Lambda$ to Modified Dark Matter (MDM)

The dynamical cosmological constant (originated from quantum fluctuations of spacetime) can now be shown to give rise to a critical acceleration parameter in galactic dynamics. The argument is based on a simple generalization of Verlinde’s recent proposal of entropic gravity for $\Lambda = 0$ to the case of de-Sitter space with positive $\Lambda$. Let us first review Verlinde’s derivation of Newton’s second law $\vec{F} = m\vec{a}$. Consider a particle with mass $m$ approaching a holographic screen at temperature $T$. Using the first law of thermodynamics to introduce the concept of entropic force $F = T \Delta S/T$, and invoking Bekenstein’s original arguments concerning the entropy $S$ of black holes, $\Delta S = 2\pi k_B \hbar \Delta x$, Verlinde gets $F = 2\pi k_B \frac{\Delta x}{\hbar} T$. With the aid of the formula for the Unruh
temperature, $k_B T = \frac{\hbar}{2\pi}$, associated with a uniformly accelerating (Rindler) observer, Verlinde obtains $\vec{F} = m\vec{a}$. Now in a de-Sitter space with positive cosmological constant $\Lambda$ for an accelerating universe like ours, the net Unruh-Hawking temperature, as measured by a non-inertial observer with acceleration $a$ relative to an inertial observer, is $\tilde{T} = \frac{\hbar}{2\pi \sqrt{a^2 + \tilde{a}_0^2}}$, where $a_0 \equiv \sqrt{\Lambda}/3$. Hence the entropic force (in de-Sitter space) is given by the replacement of $T$ and $a$ by $\tilde{T}$ and $\tilde{a}$ respectively, leading to $F = m\sqrt{a^2 + \tilde{a}_0^2 - a_0}$. For $a \gg a_0$, we have $F/m \approx a$ which gives $a = a_N \equiv GM/r^2$. But for $a \ll a_0$, $F \approx m\frac{a^2}{2\pi c} = mv^2/r$ for circular motions, so the observed flat galactic rotation curves ($v$ being independent of $r$) now require $a \approx (2a_N a_0^3/\pi)^{\frac{1}{4}}$. But that means $F \approx m\sqrt{a^2 + \tilde{a}_0^2 - a_0}$, the modified Newtonian dynamics (MoND) scaling, proposed by Milgrom. Thus, we have recovered MoND with the correct magnitude for the critical galactic acceleration parameter $a_c = a_0/(2\pi) \approx cH/(2\pi) \sim 10^{-8}\text{cm/s}^2$ (where we recall $H$ is the Hubble parameter). As a bonus, we have also recovered the observed Tully-Fisher relation ($v^4 \propto M$).

Next we can follow the second half of Verlinde’s argument to generalize Newton’s law of gravity $a = GM/r^2$. The end result is given by $\tilde{a} = G\tilde{M}/r^2$, where $\tilde{M} = M + M_d$ represents the total mass enclosed within the volume $V = 4\pi r^3/3$, with $M_d$ being some unknown mass, i.e., dark matter. For $a \gg a_0$, consistency with the Newtonian force law $a \approx a_N$ implies $M_d \approx 0$. But for $a \ll a_0$, consistency with the condition $a \approx (2a_N a_0^3/\pi)^{\frac{1}{4}}$ requires $M_d \approx \frac{1}{2} \left( \frac{a}{a_0} \right)^2 M \sim (\sqrt{\Lambda}/G)^{1/2} M^{1/2}/r$. (Note the curious connections among $M_d$, $\Lambda$ and $M$.) Thus dark matter indeed exists. And the MoND force law derived above, at the galactic scale, is simply a manifestation of dark matter!

5 Observational Tests of MDM

In order to test MDM with galactic rotation curves, we fit computed rotation curves to a selected sample of Ursa Major galaxies given in, using the mass-to-light ratio $M/L$ as our only fitting parameter. For the CDM fits, we use the Navarro, Frenk & White density profile, employing three free parameters (one of which is the mass-to-light ratio.) We find that both models fit the data well (and more or less equally well)! But while the MDM fits use only 1 free parameter, for the CDM fits one needs 3 free parameters. Thus the MDM model is a more economical model than CDM in fitting data at the galactic
scale. As for dark matter density, the profiles predicted by MDM and CDM agree well in the asymptotic (large $R$) regime. See Ref. 10) for details.

To test MDM with astronomical observations at a larger scale, we compare dynamical and observed masses in a large sample of galactic clusters studied by Sanders 16) using the compilation by White, Jones, and Forman. Sanders 16) studied the virial discrepancy (i.e., the discrepancy between the observed mass and the dynamical mass) in the contexts of Newtonian dynamics and MoND. He found the well-known discrepancy between the Newtonian dynamical mass ($M_N$) and the observed mass ($M_{\text{obs}}$):

$$\langle M_N/M_{\text{obs}} \rangle \approx 4.4.$$ And for the sample clusters, he found

$$\langle M_{\text{MoND}}/M_{\text{obs}} \rangle \approx 2.1.$$ We 10) have adapted Sanders’ approach to the case of MDM. Noting that the argument used in Section 4 does allow $M_d$ to include a term of the form $\xi \left( a_0 a \right) M$ with an undetermined universal parameter $\xi$, we (in some unpublished work) have decided to use a more general profile of the form $M_d = \left[ \xi \left( a_0 a \right) + \frac{1}{3} \left( a_0 a \right)^2 \right] M$. For $\xi \approx 0.5$, we get

$$\langle M_{\text{MDM}}/M_{\text{obs}} \rangle \approx 1.0.$$ (As an aside, we have refit the galaxy rotation curves using $\xi = 0.5$ and have found equally good fits.) Thus the virial discrepancy is eliminated in the context of MDM! At the cluster scale, MDM is superior to MoND.

6 The Dark Sector and Infinite Statistics

What is the essential difference between ordinary matter and dark energy from our perspective? To find that out, let us recall our discussion in Section 2, and liken the quanta of dark energy to a perfect gas of $N$ particles obeying Boltzmann statistics at temperature $T$ in a volume $V$. For the problem at hand, as the lowest-order approximation, we can neglect the contributions from matter and radiation to the cosmic energy density for the recent and present eras. Thus let us take $V \sim R^3_H^3$, $T \sim R_H^3$, and $N \sim (R_H/l_P)^2$.

A standard calculation (for the relativistic case) yields the partition function

$$Z_N = (N!)^{-1} (V/\lambda^3)^N,$$ where $\lambda = (\pi)^{2/3}/T$, and we get, for the entropy of the system,

$$S = -\left( \frac{\partial (-T n Z_N)}{\partial T} \right)_{V, N} = N \left( \ln(V/N\lambda^3) + 5/2 \right).$$

The important point to note is that, since $V \sim \lambda^3$, the entropy $S$ becomes nonsensically negative unless $N \sim 1$ which is equally nonsensical because $N$ should not be too different from $(R_H/l_P)^2 \gg 1$. But the solution 12) is obvious: the $N$ inside the log of $S$ somehow must be absent. That is the case if
the Gibbs $1/N!$ factor is absent from the partition function $Z_N$, implying that the “particles” are distinguishable and nonidentical!

Now the only known consistent statistics in greater than two space dimensions without the Gibbs factor is infinite statistics (sometimes called “quantum Boltzmann statistics”) \textsuperscript{11}). Thus the “particles” constituting dark energy obey infinite statistics, instead of the familiar Fermi or Bose statistics. \textsuperscript{12})

To show that the quanta of MDM also obey this exotic statistics, we first reformulate MoND via an effective gravitational dielectric medium, motivated by the analogy \textsuperscript{17}) between Coulomb’s law in a dielectric medium and Milgrom’s law for MoND. Ho, Minic and I then find that MoNDian force law is recovered if the quanta of MDM obey infinite statistics.

What is infinite statistics? Succinctly, a Fock realization of infinite statistics is provided by the commutation relations of the oscillators: $a_k a_l^\dagger = \delta_{kl}$. Curiously a theory of particles obeying infinite statistics cannot be local \textsuperscript{11}). But the TCP theorem and cluster decomposition have been shown to hold despite the lack of locality \textsuperscript{11}). Actually this lack of locality is not unexpected. After all, non-locality is also present in holographic theories, and the holographic principle is an important ingredient in the formulation of quantum gravity. Infinite statistics and quantum gravity appear to fit together nicely, and non-locality seems to be a common feature of both of them. \textsuperscript{12}) Perhaps it is the extended nature of the dark quanta that connects them to such global aspects of space-time as the Hubble parameter and the cosmological constant.

Acknowledgment

This talk is partly based on work done in collaboration with H. van Dam, J.J. van der Bij, S. Lloyd, M. Arzano, T. Kephart, C. M. Ho, D. Minic, D. Edmonds, D. Farrah, and T. Takeuchi. I thank them all. The work reported here was supported in part by the US Department of Energy, the Bahnson Fund and the Kenan Professorship Research Fund of UNC-CH.

References


The CALET Mission on the International Space Station

Pier Simone Marrocchesi *
Dept. of Physical Sciences, Earth and Environment, Univ. of Siena, Italy
and INFN Sezione di Pisa

for the CALET collaboration

* corresponding author: marrocchesi@pi.infn.it

Abstract

After a launch from the Tanegashima Space Center in Japan and successful completion of on-orbit commissioning, data calibration and verification, the CALorimetric Electron Telescope (CALET) moved to regular observation mode in mid-October 2015. To date, more than a hundred million events generated by high-energy charged particles and photons coming from space with energies above 10 GeV have been recorded on the JEM-EF external platform aboard the International Space Station. CALET is a mission of the Japanese Aerospace Agency (JAXA) in collaboration with the Italian Space Agency (ASI) and NASA. Its main science objectives include the exploration of the electron (+positron) spectrum above 1 TeV whose shape might reveal the presence of sources of acceleration just a few kpc away from Earth. With excellent energy resolution, proton rejection capability ($>10^5$) and low background contamination, CALET will search for possible signatures of dark matter in the spectra of both electrons and gamma rays. High precision measurements of the energy spectra, relative abundances and secondary-to-primary ratios of cosmic nuclei from proton to iron will be carried out as well as the detection of trans-iron elements. Deviations from a simple power-law, as reported by CREAM,
PAMELA and AMS-02 in proton and He spectra, will be studied with high accuracy in the region of a few hundred GeV and extended to the multi-TeV region and to heavier nuclei. Gamma-ray transients are detected by a dedicated Gamma-ray Burst Monitor (GBM).

1 Introduction

The CALorimetric Electron Telescope (CALET) is a space-based experiment for long term observations of high energy cosmic radiation on the International Space Station (ISS). The CALET mission of the Japanese Aerospace Agency (JAXA), in collaboration with the Italian Space Agency (ASI) and NASA, carried out preliminary phase studies in 2007-09 followed by the construction and commissioning of the payload and leading to a successful launch of the instrument on August 19, 2015 from the Tanegashima Space Center (Japan). CALET reached the ISS on August 24 on board of the transfer vehicle HTV5 (Kounotori) and was emplaced on the Exposure Facility of the Japanese Experimental Module (JEM-EF). In mid October the preliminary phase of on-orbit check-out and calibrations were accomplished. Since then the instrument is operating in science mode and transmitting data to the ground stations. At the time of writing, more than a hundred million events generated by high-energy charged particles and photons coming from space with energies above 10 GeV have been recorded.

Figure 1: (a) left: CALET layout. From top to bottom: CHD hodoscope, IMC imaging calorimeter and TASC total-absorption calorimeter; (b) right: event display of a candidate electron event with energy 4.2 TeV.

CALET is an all-calorimetric instrument designed to achieve a large proton rejection capability (>10⁶) with a fine grain imaging calorimeter (IMC) followed by a total absorption calorimeter (TASC). The overall thickness of CALET at normal incidence is 30 X₀ and ∼1.3 proton interaction length (λ₁). The charge identification of individual nuclear species is performed by a

1 dedicated trigger modes allow to lower the energy threshold to ∼ 1 GeV
two-layered hodoscope of plastic scintillators (CHD) at the top of the apparatus (Fig. 1a), providing a measurement of the charge $Z$ of the incident particle over a wide dynamic range ($Z = 1$ to $\sim 40$) with sufficient charge resolution to resolve individual elements $^3$, $^4$ and complemented by a redundant charge determination via multiple $dE/dx$ measurements in the IMC. The IMC is a sampling calorimeter longitudinally segmented into 16 layers of scintillating fibers (with $1 \text{ mm}^2$ square cross-section) interspaced with thin tungsten absorbers. Alternate planes of fibers are arranged along orthogonal directions. It can image the early shower profile in the first $3 X_0$ and reconstruct the incident direction of cosmic rays with good angular resolution. The TASC is a $27 X_0$ thick homogeneous calorimeter with 12 alternate X-Y layers of lead-tungstate (PWO) logs. It measures the total energy of the incident particle and discriminates electrons from hadrons with the help of the information from the CHD and IMC. The instrument is described in more detail $^1$, $^2$, $^5$ elsewhere.

2 Main science goals
The CALET telescope will perform precise measurements of high energy cosmic rays over a target period of five years, with an extensive physics program that includes the detection of possible nearby sources of high energy electrons; searches for signatures of dark matter in the spectra of electrons and $\gamma$-rays; long exposure observations of cosmic nuclei from proton to iron and trans-iron elements; measurements of the CR relative abundances and secondary-to-primary ratios; monitoring of gamma-ray transients and studies of solar modulation.

3 The electron spectrum
The primary science goal of CALET $^6$ is to perform high precision measurements of the electron spectrum from 1 GeV to 20 TeV. CALET will scan very accurately the energy region already covered by previous experiments with an excellent energy resolution and a low background contamination. By integrating a sufficient exposure on the ISS, CALET will be able to explore the energy region above 1 TeV where the presence of nearby sources of acceleration is expected to shape the high end of the electron spectrum and leave faint, but detectable, footprints in the anisotropy. In order to meet this experimental goal, CALET has been designed to achieve a large proton rejection capability ($>10^5$) thanks to a full containment of electromagnetic showers in the calorimeter and a fine-grained imaging of their early development in a $3 X_0$ thick pre-shower (IMC). A preliminary candidate electron event with energy 4.2 TeV is shown in Fig. 1b.
The TeV region. An exciting possibility is that the observation of the electron spectrum in the TeV region may result in a direct detection of nearby astrophysical sources of high energy electrons. In fact, the most energetic galactic cosmic-ray (GCR) electrons that can be observed from Earth are likely to originate from sources younger than $\sim 10^5$ years and located at a distance less than 1 kpc from the Solar System. This is due to the radiative energy losses that limit the propagation lifetime of high energy electrons and consequently the distance they can diffuse away from their source(s). Since the number of potential sources satisfying the above constraints are very limited, the energy spectrum of electrons might have a distinctive structure and the arrival directions are expected to show a detectable anisotropy. There are at least 9 candidate Supernova Remnants (SNR) with ages $< 10^5$ years and distances less than 1 kpc from the solar system. Possible contributions to the observed GCR electron spectrum from both distant and nearby sources were calculated. Known candidates that may give a contribution in the TeV region include Vela, Cygnus loop and Monogem, in order of strength. Among these, the relatively young Vela ($\sim 10^4$ years) at a distance of ($\sim 0.25$ kpc) is a very promising candidate.

Figure 2: Expected electron spectrum (red data points) from CALET in 5 years of observations in the hypothetical case of a prominent spectral contribution from Vela.

The TeV region might as well conceal a completely different scenario where “nearby” acceleration sources would not be detected and the spectrum found to roll off at a characteristic cutoff energy. In this case, the measurement of the “end point” of the electron spectrum could be used to constrain the cosmic-ray diffusion coefficient.

The sub-TeV region. The electron energy spectrum from 10 GeV to 1 TeV could be the result of the contribution of several unresolved sources. In
this energy region CALET energy resolution and long exposure will allow to significantly improve the knowledge of the detailed spectral shape and angular distribution of the inclusive electron spectrum. This will provide information on the average features of the source spectrum, the diffusion time, the density of sources and possibly their nature, either as astrophysical objects (e.g. a nearby pulsar) or the result of the annihilation/decay of dark matter particles \(^8\). Both possibilities have been proposed to interpret the recent measurements suggesting a hardening of the inclusive spectrum in the range 200 GeV - 1 TeV. The presence of an additional spectral component is also required to explain the now established rise of the positron fraction above \(\sim\)10 GeV as measured by PAMELA \(^9\) and extended to the hundreds GeV region by AMS-02 \(^10\).

4 Dark Matter searches.

Dark Matter (DM) candidates include WIMPs (Weakly Interacting Massive Particles) from supersymmetric theories, such as the LSP neutralino, that may annihilate and produce gamma rays and positrons as a signature. CALET will perform a sensitive search of DM candidates in the inclusive electron spectrum, as discussed above, and in gamma-ray spectra. According to a class of models, the annihilation/decay of dark-matter particles in the galactic halo could produce sharp gamma-ray lines in the sub-TeV to TeV energy region, superimposed to a diffuse photon background. CALET will be capable of investigating such a distinctive signature, thanks to a gamma-ray energy resolution of 3\% above 100 GeV, that can be improved to 1\% with a reduced on-axis effective area (fiducial volume acceptance cuts to require a total lateral containment of the shower). The precise determination of the line shape of any spectral feature is expected to play a crucial role to discriminate among different models of dark matter, or it might suggest an alternative astrophysical interpretation.

Another class of DM candidates, as suggested in \(^11\), are Kaluza-Klein (KK) particles resulting from theories involving compactified extra-dimensions. They may annihilate in the galactic halo and produce an excess of positrons observable at Earth. Unlike neutralinos, direct annihilation of KK particles to leptons is not suppressed. A sharp cutoff close to the KK mass might produce a detectable feature in the inclusive electron energy spectrum. They can also decay into gamma rays and a difference in the line shape between a neutralino and a KK candidate has to be expected \(^12\). If a high energy "gamma-ray line" will be observed, CALET might be able to resolve the nature of dark matter by studying its spectral shape thanks to an excellent energy resolution (better than Fermi-LAT above 10 GeV).
5 Cosmic-ray spectra

CALET will perform long term observations of light and heavy cosmic nuclei from proton to iron and will also detect trans-iron elements up to $Z \sim 40$. It will be able to identify the most abundant CR elements with individual element resolution and measure their spectral shape and relative abundance in the energy range from a few tens of GeV to several hundreds of TeV. CALET will first investigate - with very high accuracy - the intermediate energy region from 200 GeV/n to 800 GeV/n where a deviation from a single power-law has been reported for both proton and helium spectra by CREAM and PAMELA and recently confirmed with high statistics measurements by AMS-02. In a relatively short time, CALET will be able to close the gap between the AMS-02 highest energy points and CREAM lowest points for proton and He and will extend the energy reach to the multi-TeV region. CALET will carry out an accurate scan of this energy region to verify the presence of a spectral break and/or a progressive hardening of the spectrum by measuring its curvature and break point position.

Figure 3: (a) left: CALET (red filled circles) expected proton measurements after 1 year of observations in a restricted ($\sim 1/3$) fiducial acceptance (statistical errors only). (a) left: Proton rigidity spectrum from 50 GeV to 10 TeV with AMS-02 data points (diamonds); PAMELA (open triangles) from lowered by 3.2% as prescribed in 19)) and CREAM-I data points (filled squares) below 10 TeV; (b) right: A partial compilation of B/C data including CALET expected data (red points) after 5 years. The dashed (dot-dashed) lines are drawn for a Leaky Box Model with $-\delta = 0.33, 0.50$, respectively.

An example is given in Fig.3a where the expected proton data points (filled red circles) from CALET after 1 year of data taking are calculated, assuming the respective AMS-02 spectral parametrizations and taking into account the expected efficiencies. The errors are statistical only and refer to a restricted fiducial acceptance corresponding to a geometric factor of $\sim 0.04$ m$^2$sr (about 1/3 of the whole acceptance).
CALET will also provide precision measurements of the He spectrum to verify a possible violation of the universality of spectral indices with atomic number, whereby the He spectrum is harder than proton’s at high energy. With AMS-02 momentum measurements limited\(^2\) to a few TV by its MDR, precise observations of the proton and helium fluxes in the multi-TeV region are likely to come from purely calorimetric experiments already in orbit like CALET and DAMPE \(^{17}\) or missions scheduled for a launch in the near future, like ISS-CREAM \(^{18}\). On a longer observation time scale of 5 years, CALET is expected to explore the proton energy spectrum up to \(\sim 900\) TeV, the He spectrum to \(\sim 400\) TeV/amu and to measure the energy spectra of the most abundant heavy nuclei with sufficient statistical precision up to \(\sim 20\) TeV/amu for C and O and \(\sim 10\) TeV/amu for Ne, Mg, Si and Fe \(^{6,21}\).

Secondary-to-primary flux ratios. Direct measurements of the energy dependence of the flux ratio of secondary-to-primary elements (e.g.: B/C, sub-Fe/Fe) can discriminate among different models of CR propagation in the galaxy. Above 10 GeV/amu, the energy dependence of the propagation path-length is often parametrized in the form \(E^{-\delta}\). An accurate measurement of the spectral index parameter \(\delta\) is crucial to derive the spectrum at the source by correcting the observed spectral shape for the energy dependence of the propagation term. These measurements have been pushed to the highest energies with Long Duration Balloon (LDB) experiments. However, they remain at present statistics limited to a few hundred GeV/amu and suffer from systematic uncertainties due to the production of secondary nuclei in the residual atmospheric grammage at balloon altitude that may become dominant in the TeV/amu region. With a long exposure and in the absence of atmosphere, CALET can provide new data \(^{6,21}\) to improve the accuracy of the present measurements of the B/C ratio (Fig.3b) above 100 GeV/amu and extend them beyond 1 TeV/amu.

6 Gamma-ray astrophysics.

Observation of gamma-ray sources is not a primary objective for CALET. However, its excellent energy resolution and good angular resolution (better than 0.4°, including pointing uncertainty) will allow for accurate measurements of diffuse gamma-ray emission and detection of more than 100 bright sources at high latitude from the Fermi-LAT catalogue. Given the on-axis effective area

\(^2\)AMS-02 energy range can be extended for nuclei with \(Z > 2\) by using the TRD, while for protons and He the calorimeter is limited by a thickness of only \(\sim 0.5\) interaction length that significantly reduces the expected number of proton interactions.
of $\sim 600 \text{ cm}^2$ for energies above 10 GeV (reduced by $\sim 50\%$ at 4 GeV) and field of view of 45° from the vertical direction, CALET is expected to detect $\sim 2.5 \times 10^4$ ($\sim 7000$) photons from the galactic (extra-galactic) background with $E > 4 \text{ GeV}$ and $\sim 300$ photons from the Vela pulsar with $E > 5 \text{ GeV}$.

**Gamma-ray Transients.** CALET will also monitor X-ray/gamma-ray transients in the energy region 7 keV to 20 MeV with a dedicated Gamma-ray Burst Monitor (CGBM). It will extend GRB studies carried out by other experiments (e.g. Swift and Fermi/LAT) and provide added exposure when the other instruments are not available or pointing to a different direction. Furthermore, high energy photons possibly associated with a burst event can be recorded over the entire CALET energy range down to 1 GeV where the CALET main telescope has still (limited) sensitivity, albeit with low resolution. Upon the detection of a GRB, an alert will be transmitted to a network of ground "antennas" (including LIGO and VIRGO) for the possible simultaneous detection of gravitational waves and their electromagnetic counterparts. At the time of writing more than 20 GRBs were recorded by the CGBM.

7 Conclusions

CALET reached the ISS on August 2015 and was emplaced on the Exposure Facility JEM-EF. At the time of writing, CALET is successfully operating in science data mode and has recorded more than 100 million events. A 2 year period of observations has started with a target of 5 years.

References

21. P.S. Marrocchesi et al., IEEE NSS (Seoul), N31.3 (2013)
COSMIC-RAY DIRECT DETECTION

Piergiorgio Picozza
INFN and University of Rome Tor Vergata

Laura Marcelli
INFN, Sezione di Roma Tor Vergata

Abstract

Direct research on cosmic-rays in the energy region between tens MeV/n and few TeV/n has been extensively undertaken by experiments on board stratospheric balloons, satellites, space stations, since the sixties. The main goals are the search for primordial antimatter, dark matter annihilation signals and exotic particles and the study of the mechanisms of production, acceleration and propagation of cosmic-rays. The monitoring of the cosmic-ray solar modulation, the detection of solar flares and studies on the radiation belts around the Earth complete the research program. A review of the major results up to an energy of few TeV obtained by the current experiments, mainly PAMELA and AMS-02, will be presented in this paper.
1 Introduction

Cosmic-rays are a sample of solar, galactic and extragalactic matter which includes all known nuclei and their isotopes, as well as electrons, positrons, and antiprotons. The cosmic-ray particles, at least up to about $10^{15}$ eV, are considered of galactic origin and shock waves of expanding supernovae remnants are the main candidates to supply the power for their acceleration. The accelerated particles are injected into the interstellar space, where they remain about $10^7$ years before escaping into the intergalactic space. New particles and spallation products are obtained by interaction of cosmic-rays with the interstellar matter. Solar modulation affects the low energy part of the cosmic-rays and plays an important role in the precise determination of their interstellar energy spectrum. The experimental methods divide the cosmic-ray energy spectrum in two large intervals. The first, below some hundreds TeV, is explored by direct measurements, carried out by experiments on stratospheric balloons or in space on board satellites or the International Space Station (ISS), with single particle identification and energy denition. The second one, at higher energy, where the low particle flux makes feasible only indirect observations, is measured by very large on ground detectors and the particle nature is inferred, on basis of the statistics and with considerable systematic uncertainties, studying their interaction with the atmosphere. In this paper we deal with direct measurements of cosmic-rays up to 2 TeV with a focus on the results on antimatter particle and on light nuclei energy spectra measured by PAMELA satellite experiment and AMS-02 on board the ISS.

2 PAMELA and AMS-02

Primary scientific objectives of PAMELA and AMS-02 are precise measurements of the antiparticle energy spectra to unravel possible contributions from exotic sources. They can also search for heavy antimatter, in particular antihelium (primordial antimatter), for new matter in the Universe (strangelets?) and perform accurate studies of nuclei and their isotopes to test cosmic-ray propagation models. Concomitant goals include the study of solar physics, solar modulation and radiation belts. The PAMELA experiment is performed from an international collaboration including Italy, Russia, Germany and Sweden. The core of the instrument, shown in fig. 1, left panel, is a magnetic
spectrometer for particle momentum and charge sign determination. The MDR is of the order of 1.2 TeV. A Time of Flight system provides timing, dE/dx measurements and the primary PAMELA trigger. The separation between hadronic and leptonic components is made, at the order of $10^5$ of proton rejection power, by an imaging silicon-tungsten detector, 16 radiation length deep, with high segmentation and able to measure electron energy up to 300 GeV with a resolution of a few percent. A further improvement to the rejection power is given from a neutron counter that detects neutrons produced in the showers that incident particles create in the calorimeter; neutrons are more numerous in the hadronic shower compared with the electromagnetic one. A thick scintillator above the calorimeter and an anticoincidence system complete the apparatus. More technical details can be found in 1). PAMELA has been inserted in a pressurized vessel and installed on board of the Russian satellite Resurs DK-1 dedicated to Earth observation. It was launched on June 15, 2006, in an elliptical orbit, ranging between 350 and 610 km and with an inclination of 70°. In September 2010 the orbit was changed to a nearly circular one with

Figure 1: PAMELA (left) and AMS-02 (right) apparatuses.
diameter of 570 km and it has not changed since then. In fig. 2 an overview of the most interesting PAMELA results is displayed.

On May 16th, 2011, the AMS-02 instrument, constructed by a worldwide collaboration, was placed outside the ISS. The detector, shown in fig. 1, right panel, has a size of \(3\times3\times3\) m\(^3\) and a weight of 7 tons. The main instrument of AMS-02 is a magnetic spectrometer composed of a permanent magnet and a tracking system. The MDR ranges from 0.23 TV for an acceptance of 0.41 m\(^2\)sr to 2.21 TV for an acceptance of 0.01 m\(^2\) sr. Complementary detectors are a Transition Radiation Detector, a trigger and Time-of-Flight system, a Ring Imaging Detector and, at the bottom of the apparatus, an electromagnetic calorimeter of 16 radiation lengths made of lead scintillating fibers.

3 Primary cosmic-rays

Hydrogen and helium are the most abundant components of the galactic cosmic-rays, constituting about 98% of the total flux, and are believed to be of primary origin. Many experiments have measured the energy spectra of hydrogen and helium over many years, but the first high statistics and high precision measurements, shown in fig. 3, left panel, have been carried out by PAMELA \(^2\). The data show interesting spectra features that figure significant and intriguing implications for the understanding of high energy galactic processes. Both
Figure 3: Proton (top data set) and helium (bottom data set) fluxes measured by PAMELA (left panel) and ratio of the fluxes between proton and helium (right panel) in the rigidity range 1 GV to 1.2 TV. The pink shaded areas represent the estimated systematic uncertainty.

spectra show a deviation from a single power law, with energy spectra that gradually soften in the rigidity range 20-230 GV, then exhibiting a spectral hardening at a rigidity of about 230-240 GV. This behavior challenged the previous paradigm of acceleration of cosmic-rays by a single supernovae remnant and has been interpreted as an indication of different populations of cosmic-ray sources, as novae stars and explosions in superbubbles. Furthermore, for long time the possible uniqueness of the index spectrum for all nuclei, including protons, has been a debated issue. The precise PAMELA data, shown in fig. 3, right panel, as ratio of the proton to helium fluxes versus rigidity, in order to cancel systematic instrumental effects, clearly evidence a difference between the proton and helium slopes. The ratio shows a continuous and smooth decrease and it is well described by a power law down to 5 GV with a spectral index of 0.1.

These results have been confirmed four years later by AMS-02 that released data between 1 GV - 1.8 GV for protons (fig. 4, top) and 1.9 GeV - 3 TeV for helium (fig. 4, bottom). The highly precise data show a smoothly variation of the spectral indexes as a function of rigidity.

4 Secondary-Primary cosmic-rays

Many are the physical processes that cosmic-rays undergo during propagation, as diffusion, spallation, emission of synchrotron radiation, etc. and shape the
injection spectra and chemical composition into the observed values. A detailed knowledge of these processes is therefore needed in order to interpret the experimental data in terms of source parameters, or in estimating the expected background when searching for contributions from new sources. A powerful tool to constrain the parameters of the galactic propagation models is the measurement of the secondary to primary flux ratios, as B/C, $^2$H/$^4$H, $^3$H/$^4$H. The B/C flux ratio is a clean and direct probe of propagation mechanisms, and it is considered as the ”standard tool” for studying propagation models. Indeed, boron is produced in negligible quantities by stellar nucleosynthesis processes and almost all of the observed boron is believed to be from spallation reactions of CNO primaries on atomic and molecular H and He present in the ISM. The B/C ratios measured by PAMELA $^3$ and AMS-02 $^4$, respectively in the energy range 2.02 - 260 GV and 2 GV - 1.8 TV, are shown in fig. 5, left side. In fig. 6, right side, are shown the absolute fluxes of boron and carbon measured by PAMELA $^3$). They are in good agreement with the previous measurements, except at low energy where solar modulation plays an important role depending from the different solar activity period in which experimental data have been collected.

PAMELA studied also the rare isotopes $^2$H and $^3$He in cosmic-rays, re-
Figure 5: Left: boron to carbon ratio in rigidity units as measured by PAMELA and AMS-02 experiments 4). Right: Absolute boron and carbon fluxes multiplied by $E^{2.7}$ as measured by PAMELA together with results from other experiments 3).

sulting mainly from the nuclear interactions of primary cosmic-ray protons and $^4$He with the interstellar medium. The isotopic composition of hydrogen and helium was measured, respectively, between 100 - 1100 MeV/n and 100 - 1400 MeV/n, over the 23rd solar minimum from 2006 July to 2007 December. Absolute fluxes and ratios are shown in fig. 6 5).

5 Antiparticles

The first historical discovery of antiprotons on the top of the atmosphere was done by the balloon-borne experiments carried out from Robert Golden and Edward Bogomolov in 1979 6, 7). They detected an amount of antiprotons much higher than expected from interactions of cosmic-rays with the interstellar matter. These data were interpreted in terms of primary antimatter coming from antimatter domains in a baryonic symmetric Universe, evaporation for Hawking effect of primordial mini black holes, exotic particles annihilation. Many balloon-borne experiments followed these pioneer ones, but the antiprotons excess had not been confirmed. However, concerning the positron to electron ratio, a hint of an increase or a flatness above 7 GeV seems to appear in the data. This ambiguous feature was interpreted in terms of a possible annihilation in pairs of WIMP particles, in the framework of supersimmetry, that produces as final state pairs of $e^+ - e^-$, $\bar{p} - p$, etc. However, no clear assump-
tion was possible to obtain from these data, so that experiments at higher energies, better knowledge of the astrophysics background, higher statistics and continuous monitoring of solar modulation were considered mandatory to extract conceivable exotic components from standard production.

PAMELA first, and four years later AMS-02, completely clarified the experimental situation opening a new fascinating and intriguing scientific case. The antiproton flux measured by the PAMELA experiment in the energy range 60 Mev - 180 GeV \(^8\) is shown in fig. 7 in the left side, while the antiproton to proton ratios measured by PAMELA in the same range and AMS-02 from 0.5 Gev up to 450 Gev \(^9\) are shown in the right side. The data at the energies of PAMELA do not present features or structures expected from exotic sources, so they place strong limits to dark matter annihilation models and set tight constraints on parameters relevant for secondary production. However, the recent data of AMS-02 \(^9\) show a flatness in the antiproton to proton ratio at high energy that leaves room for an antiproton contribution from an exotic source.

The positron to all electron ratio measured by the PAMELA experiment
Figure 7: Left: antiproton flux measured by the PAMELA. Right: antiproton to proton ratio measured by PAMELA and AMS-02.

and published in Nature\textsuperscript{10}, fig. 8, aroused great interest both in the astrophysics and particle physics communities. The data, covering the energy range 1.5 - 100 GeV, show two clear features. At low energies, below 5 GeV, the PAMELA results are systematically lower than data collected during the 1990’s because of the different periods of solar activity when data have been collected\textsuperscript{11}. At energies above 10 GeV they show a positron fraction increasing significantly with energy, contrary with the standard calculations that

Figure 8: PAMELA positron fraction with other experimental data and with secondary production model.
foresee a continuous decrease. Many theoretical models explained the origin of this observed excess as annihilation or decaying of dark matter. The most problematic challenge posed by the PAMELA results is the asymmetry between leptonic (positron fraction) and hadronic ($\bar{p} - p$ ratio), difficult to explain in the framework in which the neutralino is the dominant dark matter component. The best explanation is obtained in terms of a direct leptonic annihilation channel for a wide range of the WIMP mass, but in this case a large boost factor is required. Another explanation relates to a contribution from nearby and young pulsars, objects well known as particle accelerators. Other models consider this excess of positrons as due to a standard production in some inhomogeneity in the SNR density in our Galaxy \(^{12}\) or in the same site where protons are accelerated \(^{13}\). More recent PAMELA data \(^{14}\) and the first AMS-02 results \(^{15}\) are displayed in fig. 9 showing a very good agreement in the energy region not affected by solar modulation. It is worth to note that the last data published by AMS-02 on the positron fraction \(^{16}\) between 0.5 - 500 GeV show a flatness above $\sim$200 GeV that suggests the possibility of a decrease at higher energy as expected from exotic source contribution.

Figure 9: Left: PAMELA \(^{14}\) and AMS-02 positron fractions \(^{15}\). Right: High energy AMS-02 results \(^{16}\).

References

REVIEW ON GALACTIC COSMIC RAY DETECTION

Chiavassa Andrea
Università degli Studi di Torino & INFN

Abstract
Galactic cosmic rays cover a range from $10^{10} eV$ (when primary particles are not affected by the solar wind) up to $10^{17} - 10^{18} eV$ (where extra-galactic primaries are expected to dominate the measured flux). This energy range is covered by very different experiments: until $\sim 10^{14} eV$ primaries can be measured directly with detectors operating on balloons or on satellites. While greater energies must be studied by indirect experiments sampling the EAS (Extensive Air Showers) secondary particles.

Indirect EAS experiments are mainly limited by the systematic errors due to their energy calibration. I will discuss the main sources of these errors: the choice of the hadronic interaction model and of the mass of the primary particle (that cannot be measured on an event by event basis).

I will summarize some recent measurements of the all particle spectrum, showing that, keeping into account the differences due to the energy calibration, they agree on the spectral shape. Then I describe the measurements of the light and heavy primaries mass groups spectra, discussing the claimed features.
1 Introduction

Primary cosmic rays with energy greater than $10^{14}\text{eV}$ cannot be studied by direct experiments operating on balloon or on satellites and their detection is only possible by means of indirect EAS experiments. All the main characteristics of the primary particle (i.e. mass, energy and arrival direction) must therefore be derived measuring the Extensive Air Showers (EAS) generated by the interaction of primary cosmic rays with atmospheric nuclei.

The experiments operating in the $10^{14} - 10^{18}\text{eV}$ energy range can be divided in two groups:

- **Surface arrays**: sampling the EAS at fixed atmospheric depth. Detecting the particle density and arrival times at different distances from the shower core these arrays derive the arrival direction of the primary cosmic ray, the number of charged particles ($N_{\text{ch}}$) and the number of muons ($N_{\mu}$) in the EAS at observation level. These detectors operate with a 100\% duty cycle.

- **Cherenkov arrays**: detecting the cherenkov light emitted by particles during EAS development. These arrays perform an almost calorimetric measurement of the primary energy. The atmospheric depth of the shower maximum can be estimated comparing the cherenkov photon densities measured at two different distances from the shower core. These arrays operate only during clear moonless nights, therefore their duty cycle is limited to $\sim 10 - 15\%$.

The main problem of these experiments is that energy and mass calibrations have to be performed by means of EAS simulations that are based on hadronic interaction models extrapolated from lower energy collider measurements.

In spite of all the mentioned calibration problems the results of last generation experiments brought to a deeper understanding of the knee of the primary cosmic ray spectrum. The scenario resulting from such measurements favors an astrophysical interpretation of the knee, even if the resulting picture is not complete and, to further improve our knowledge, high precision experiments are needed.
2 Energy and Mass Calibration of Indirect Experiments

2.1 Energy Calibration

Surface arrays derive the energy of the primary particle either from the measured value of the number of charged particles ($N_{ch}$) either from the measured value of the number of muons ($N_\mu$) or from a combination of both. These shower parameters depend not only on the energy and mass of the primary cosmic ray but also on the atmospheric depth crossed by the EAS, i.e. the zenith angle $\theta$. The shower evolution in atmosphere is usually treated by two different approaches: a first one, based on the experimental data, is the so called ”constant intensity cut” 1), while a second one describes the EAS atmospheric evolution by means of a complete simulation.

The main sources of systematic errors in the energy calibration of indirect experiments are related to the two main hypotheses that must be done when running the EAS simulations: the mass of the primary particle generating the shower and the hadronic interaction model describing high energy interactions.

For the choice of the mass of the primary particle three different strategies are used by experiments:

- the primary energy is calculated for the two extreme values of the primary mass (Hydrogen and Iron). The spectra that are obtained for these two cases represent the upper and lower limits bracketing the ”true” spectrum.

- Starting from the primary chemical composition measured at lower energy by direct experiments and assuming its evolution with energy the mean value of the primary mass is calculated as a function of the primary energy. This mean value of the primary mass is then used to convert the experimental observable to primary energy.

- Combining the $N_{ch}$ and $N_\mu$ values measured for each event a parameter correlated with the primary mass is evaluated. Then this parameter is used in the conversion from the experimental observable to the primary energy 2).

The second main source of systematic error in the energy calibration of indirect experiments is the choice of the hadronic interaction model used in
the EAS simulation. The more often used hadronic interaction models used by current EAS experiments are: QGSJetII-02 \(^3\), Sibyll2.1 \(^4\) and EPOS1.99 \(^5\). The KASCADE-Grande experiment has evaluated the differences in the energy assignment due to the hadronic interaction model choice \(^6\), finding that they are smaller than 20%.

2.2 Mass Calibration

Due to EAS development fluctuations indirect experiments cannot determine the mass of every single event. Studies of the primary chemical composition can be performed using statistical methods like, for instance, the unfolding analysis introduced by the KASCADE experiment \(^7\), or separating events in two mass groups, i.e. light and heavy primaries.

The shower observables sizable to such purpose are: the measurement of the shower maximum development, the ratio between observables representing the charged particles and the muon numbers at observation level, a correlation between the number of particles at observation level and a parameter reflecting the shape of the particle lateral distribution. In order to assign a value of the primary mass to any of the previously indicated experimental observables a complete EAS simulation is needed, therefore all these analyses depend on the high energy hadronic interaction model used.

New versions of the previously mentioned hadronic interaction models have been released after the publication of the results obtained by LHC experiments studying the interactions of the 7\( TeV \) proton beams. At the time of writing this contribution are known the results obtained running the EPOS-LHC \(^8\) and QGSJetII-04 \(^9\) hadronic interaction models. These new codes do not introduce major differences in the energy calibration, the major changes concerns the mass calibration because both models produce more muons with respect to the pre-LHC ones. Therefore we expect a lighter chemical composition analyzing experimental data with post-LHC hadronic interaction models.

3 Experimental Results

Having discussed the main sources of systematic errors affecting the energy and mass measurements performed by indirect experiments, in this section I will discuss some recent results obtained in the \( 10^{14} - 10^{18}eV \) energy range.
3.1 All Particle Energy Spectrum Measurements

Figure 1 shows a compilation of the all particle energy spectrum measured by different experiments using different techniques, operating at different height above sea level and calibrated by different hadronic interaction models.

Looking at figure 1 we can have the impression of a confused situation, with measurements showing big differences between each other, but simply shifting, as already proposed in 19), the single experiment energies by an amount smaller than the previously discussed systematic errors (i.e. $\Delta E/E \leq 20\%$) all the results agree much better, as can seen in figure 2. The differences shown in figure 1 can thus be attributed to the energy calibration of the experiments and there is a general agreement about the structure of the spectrum showing, in addition to the well known feature of the knee, two faint but significant structures at $\sim 10^{16}eV$ and at $\sim 10^{17}eV$.

3.2 Mass Groups Energy Spectra

Up to now indirect experiments have separated events in two mass groups following two different approaches.

A first analysis technique separates the events according to the measured ratio between the muon and charged particles numbers, converted, by means...
Figure 2: All particle energy spectrum measured in the $10^{14} - 10^{18}$ eV energy range. The energy calibration of the experiments are shifted by an arbitrary amount smaller than the calibration systematic errors. References can be found in the caption of figure 1, the energy shifts applied are reported in the legend.

of the constant intensity cut, to a reference zenith angle. Events with a high value of this ratio are attributed to heavy primaries while those with a low value to light ones.\textsuperscript{20, 21, 22)}

The ARGO-YBJ experiment\textsuperscript{11)} being a full coverage detector operating at high altitude (i.e. almost at the level of shower maximum) and not having a muon detector, selected the H and He events using two parameters obtained operating the prototypes of the LHAASO telescopes together with the ARGO-YBJ RPC carpet.

Figure 3 shows a summary of the, energy calibrated, KASCADE-Grande and ARGO-YBJ measurements of the light primaries spectra. In the plot also the direct measurements of the H and He spectra performed by the CREAM collaboration\textsuperscript{23)} are shown, their agreement with the low energies measurement of the ARGO-YBJ experiments confirms the validity of the energy calibration of this experiment and suggests also the good reliability (at least around $10^{13}$ eV) of the hadronic interaction models included in the CORSIKA code.

The KASCADE experiment published\textsuperscript{20)} the spectra of the "electron rich" (i.e. light primaries) and the "electron poor" (i.e. heavy primaries) events. These spectra are not calibrated in energy (they are shown as a function
of the muon density at fixed core distance), but using the integral flux above the change of slope we can identify this feature with the knee of the all particle energy spectrum.

Figure 3: Spectra of the light primaries measured by ARGO-YBJ (the different analysis shown by the collaboration are described in 11)) and KASCADE-Grande. Lower energy direct measurements of the CREAM experiment are shown as a reference. The former measurement obtained by the KASCADE collaboration is not shown in this plot as it was displayed in a non energy calibrated way 20)

The ARGO-YBJ collaboration published the spectrum of the H and He primaries showing a very sharp change of slope at $730 \pm 230 \pm 70 TeV$, this results is difficult to coincide with the previously cited KASCADE light primaries spectrum, the two change of slope are at different energies.

The KASCADE-Grande experiment published the spectra of the light and heavy components in two slightly different energy ranges with selection cuts defined in order to enhance the spectral features of these two components:

- an hardening of the light component spectrum $^{22)}$; at $10^{17.08 \pm 0.08} eV$ the spectral slope changes from $\gamma = -3.25 \pm 0.05$ to $\gamma = -2.79 \pm 0.08$.
- A steepening of the heavy primaries spectrum $^{21)}$; at $log(E/eV) = 16.92 \pm 0.04$ the spectral slope changes from $\gamma = -2.76 \pm 0.02$ to $\gamma = -3.24 \pm 0.05$. 

4 Concluding Remarks

The energy range dominated by galactic cosmic rays has been studied in the last years with high precision and high statistics experiments, both at the lower energies with satellites experiments (i.e. CREAM, PAMELA and AMS-02) and at the higher ones with EAS experiments (i.e. ARGO-YBJ, KASCADE, ICETOP, KASCADE-Grande and more). This big amount of data brought us to a deeper insight by which we learned that the spectrum is more complicated than it was thought. Beside the well known structure called the knee (the steepening observed at $2 - 4 \times 10^{15} \text{eV}$) more structures have been observed: a hardening of the H and He spectrum around $\sim 200 \text{GeV/nucleon}$; a hardening of the all particle spectrum at $\sim 10^{16} \text{eV}$ and a faint steepening around $\sim 8 \times 10^{16} \text{eV}$.

The studies of the primary chemical composition performed with EAS indirect experiments has moved from studies of the mean values of the experimental observables to the more informative studies of the spectra of primary mass groups (the single element separation is very difficult due to shower development fluctuations). The results obtained, separating on a event by event basis, the light and heavy primaries show a steepening of the light component at different energies: below $10^{15} \text{eV}$ for the ARGO-YBJ collaboration $^{11}$ and above $10^{15} \text{eV}$ for the KASCADE collaboration $^{20}$. These two spectral features cannot be coinciled and these different observations have to be clarified by future experiments. A hardening of the light primaries spectrum has been observed at $10^{17} \text{eV}$ by the KASCADE-Grande experiment $^{22}$. Concerning the heavy primaries spectrum a steepening has been measured by the KASCADE-Grande experiment $^{21}$.

The analysis technique of separating different mass groups on event by event basis is very powerful and must be further investigated to reach the goal of obtaining more than two groups. Moreover a larger statistics experiment, separating at least two mass groups, will allow the study of the arrival direction anisotropy for different mass groups.

References


PROSPECTS OF SPECTRUM MEASUREMENTS FOR INDIVIDUAL SPECIES OF COMIC RAY PARTICLES ABOVE 100 TEV USING LHAASO ARRAY

B. Bi, Z. Cao, Y. Liu, L. Ma, H. Wu, L. Yin, S. Zhang for LHAASO Collaboration

Institute of High Energy Physics, 19B, Shijingshan Yuquan St. Beijing 1100049

Abstract

Recently the knee of the spectrum of a mixed cosmic ray proton and Helium nuclei was found below 1 PeV by the hybrid experiment with ARGO-YBJ and a prototype Cherenkov telescope of LHAASO project. A similar measurements about the knees will be carried out with significantly improved performance of identification of primary particles and enormous statistics using the LHAASO instruments. The first 1/4 LHAASO array including 6 Cherenkov telescopes covering a solid angle of $\sim 0.2$ sr, 22,500 m$^2$ water Cherenkov detector and an array of muon detectors covering 250,000 m$^2$ using about 10,000 m$^2$ muon-sensitive area will measure air showers with great details. Here we present a preliminary investigation on the capability of particle identification, in particular for pure proton samples. More sophisticated analysis is under development.
1 Introduction

The knee of the spectrum of a mixed cosmic ray proton and Helium nuclei was found around 0.7 PeV 1) 2) by the hybrid experiment with ARGO-YBJ and a prototype Cherenkov telescope of LHAASO project 3) 4). The result “gives fundamental inputs to galactic cosmic ray acceleration models”, as the paper concluded 1). ARGO-YBJ fully covered RPC detector measured the air shower cores and nearby particle distributions precisely, thus the telescope was enabled to measure the shower energy with a Gaussian resolution of 25% nearly unbiased over the range from 100 TeV to 3 PeV. Combining the shape of lateral distribution near the cores with the shower Cherenkov image Hillas parameters together, the primary protons and Helium nuclei (H&He) are identified out of the well selected good samples of showers which hit in the RPC array. The purity of the selected sample reached 93% below 0.7 PeV 5). Even if the contamination of heavier nuclei increases with shower energy from 2.5% at 125 TeV to 13% at 1 PeV, a clear deviation from the single index power law spectrum was observed above 0.7 PeV, indicating the knee of the H&He spectrum being around 0.7 PeV. Limited by the aperture of the telescope, only 94 events were collected above 0.8 PeV, thus a large uncertainty of the knee energy and the spectral index of the H&He spectrum above the knee still remained 1). Moreover, establishing a pure H sample and observing the knee of the proton spectrum are still a dream so far. LHAASO will enable the measurements with greatly improved performance due to the scale of the experiment. LHAASO is an air shower array at high altitude of 4410 m above sea level. The coverage of the array is 1.3 km$^2$ with 5195 scintillator detectors with a spacing of 15 meters and 1171 water Cherenkov muon detectors with a spacing of 30 meters, a big water pool in the center of the array with 3000 Cherenkov detector cells of 5m x 5m and 12 air Cherenkov/fluorescence telescopes. In 2018, the first 1/4 of the experiment will start operating providing a large amount of data, so that relevant results on cosmic ray physics will be already achievable in only one year of data taking. The 1/4 array will include 6 Cherenkov telescopes covering a solid angle of ~0.2 sr, 22,500 m$^2$ water Cherenkov detector and an array of muon detectors covering 250,000 m$^2$ using about 10,000 m$^2$ muon-sensitive area. The operation of such an instrument for one year will yield a sample of more than 100k well reconstructed air shower events above 0.1 PeV with the cores, arrival directions and energies being measured at the resolution of 3 m,
0.3° and 20%, respectively. The lateral distributions of energy fluxes in the area of 25 m² near the shower cores, Hillas parameters of shower Cherenkov images, remaining energies of the showers as they hit into the pond and muon content are measured simultaneously. Combining the shower energy together, those parameters will be useful to identify the composition of the primary particles with high purity. The goal is to establish a pure proton sample of at least 20,000 events in one year, and measure the knee of the proton spectrum using LHAASO. In this paper, we are exploring the way of reaching the specific goal.

2 Detectors, Measurements and Uncertainties

The atmosphere above the LHAASO Water Cherenkov Detector Array (WCDA) watched by the Wide Field-of-View Cherenkov Telescope Array (WFCTA) and the surface detector arrays, WCDA and Muon Detector Array (MDA), form a complete shower detector complex.

WCDA is a water pond with a depth of 4.5 m divided into 900 cells. Each cell has two PhotoMultiplier Tubes (PMT) anchored on the bottom at the center of the cell of 5m×5m. The photocathodes of the PMTs are 1 inch and 8 inch (the ratio of the cathode area is 101:28353), respectively. The larger one is for small signals from 1 photoelectron (PE) to 3000 PEs, and the smaller one catches large signals equivalently from 2800 to 28,000,000 PEs. Secondary electrons and photons falling in the cell will induce cascade processes which produce Cherenkov light in water. The light intensity is proportional to the total energy carried by the air shower electrons and photons in the cell. The larger PMT will time the arrival of the secondary particles with a resolution of 2 ns. The energy flux, which falls rather rapidly with the distance from the cell to the shower core, will be measured by the PMTs with the resolution of 50% about 10 PE on the smaller PMTs and 1% about 100,000 PEs, respectively. Combining the temporal and spatial distribution of the secondary particles, one will reconstruct the shower arrival direction at the resolution of 0.3° and core position at the resolution of 3 m. The total amount of the energy, measured by the total number of equivalent PEs on PMTs, in the cell hit by the core mainly carried by high energy photons and electrons produced in last few generations of the air shower without suffering many Coulomb scattering. Therefore, it is a good measure of the hadrons produced in the last few generations, thus a useful parameter sensitive to the identity of the primary particle, because the
number of hadrons in a shower reduces with the elongation of the shower quite
sensitively 7).

Each WFCT in the array of 6 telescopes has a light collecting area of 5 m² and a camera of 1024 pixels, each of them covering a patch of the sky of 30° in angular diameter, watching a FoV of 14°×16° 8). Pixel detector is a SiPM followed by a flash A/D convetor-based (FADC) front end electronics (FEE) 9). Air showers in the FoV will be imaged by at least 10 registered pixels. Given the accurately measured shower geometry by WCDA, the telescopes will be useful in measurements of total energy and image shape of the shower. Light intensity weighted length (L) and width (W) of the image are sensitive to the identity of the primary particle. The ratio of L/W is an optimized parameter i.e. separation between the distributions of the ratios due to the light and heavy primaries is found rather significant 1).

MDA is an array of underground water Cherenkov detectors. In this paper only 300 MDs partially surrounding WCDA is relevant. The layout of the array is illustrated in Figure 1. Each MD is burred 2.5 m below the surface to screen electrons and photons in air showers. The detector is a 36 m² cylinder filled with pure water with a depth of 1.2 m and having a single 8” PMT installed on the top at the central position. Cherenkov photons produced inside the detector by muons will eventually reach to the PMT through a complex path of reflections on the inner surface of the detector. The efficiency of muon detection is better than 95% for every detector 10).

Using all detector arrays in LHAASO, therefore, one is able to measure a shower with at least 5 relevant parameters in addition to the basic shower geometrical parameters, i.e. shower energy by WFCT, shower image shape by also WFCT, energy flux near the shower core by the 5m×5m cell hit by the shower core in WCDA, remaining shower energy by also WCDA and muon content outside the core region of the shower by MDA. They are rather independent parameters since image shape is an indicator of the depth of shower maximum, energy flux near the core measures the number of hadrons in the latest generations in the cascade process, the remaining energy illustrates the age of the shower as it hits on the ground and the number of muons measures the large transverse momentum hadron production in the whole cascading history. The shower energy is determined by using the total number of photoelectrons in the shower image measured by WFCT, which is in fact dependent of the
primary composition \cite{1}). In the hybrid measurement, however, the remaining shower energy measured by WCDA could help the energy deposit in the air measured by WFCTs to improve the energy resolution and reduce the systematic dependence of the primary composition. Following the iteration procedure suggested in ref. \cite{1}, one may eventually reconstruct the shower energy after their primary composition is well determined. In the iterating procedure, the entanglements between the parameters can be greatly weakened by grouping showers according to their energies. Thus, the shower composition is able to be also determined with a minimized systematic uncertainty, except for that due to the interaction models.

Figure 1: The layout of the central part LHAASO. The red dotted circle includes the components of the 1/4 LHAASO array considered in this paper: a) the 22,500 m² pool of WCDA (indicated by the shaded square), b) ~ 200 MDs (indicated by blue dots), c) 6 telescopes of WFCTA near the pool (indicated by filled black squares).

The uncertainty due to the interaction models can be partially estimated by means of LHC data. For both regions of pseudo rapidity from 8.81 to 8.91 and greater than 10, the models such as EPOS-LHC, SIBYLL 2.1 and QGSJet-II 04 reproduce the distributions of secondary gamma-like and neutron-like particles with energy from 200 GeV to 3 TeV with reasonable accuracy. For instance, the predictions of those models are all deviating from the data taken by LHCf experiment \cite{11}, but less than ±50% for gammas and neutrons below 2 TeV.
Above 2 TeV, deviation is still significant. Muon content in air shower has even larger uncertainty in predictions of those models. As an example, if the experiment NA22 data about the $\pi^0$ and $\rho^0$ production were take into account in SIBYLL, the content of muon with energy greater than 1 GeV in a proton shower at 100 EeV would be boosted by a factor of 2\textsuperscript{12}) comparing with the current version SIBYLL 2.1. Those uncertainties have to be put in mind as they are used in the composition analysis described below.

3 Identification of Primary Particles

With all those uncertainties in mind, we have studied the shower parameters suitable for cosmic ray composition measurements. We carefully evaluate the parameters for their sensitivity to the composition of the primary particle and possible correlation between them. In the following we will describe the most relevant five parameters and corresponding mass separation capabilities, determined by simulating primary particles of different species and energy, using the QGSJetII hadronic interaction model.

As mentioned, the total number of photoelectrons in the Cherenkov image $N_{pe}$ is a good shower energy estimator, if it is normalized to $R_p = 0$ and $\alpha = 0$ where $R_p$ is the impact parameter of the shower axis to the Cherenkov telescope and $\alpha$ is the space angle between the shower direction and the optical axis of the telescope. The normalized parameters is denoted as $N_{pe}^0 = \log_{10}N_{pe} + 0.0092(R_p/1m) + 1.05\tan\alpha$.

The energy flux near the core as parameter $p_F = \log_{10}W_{max} - 1.39\log_{10}N_{pe}^0$, where $W_{max}$ is the total number of equivalent photoelectrons recorded by the small PMT in the water Cherenkov detector cell hit by the shower core. The energy dependence is reduced in the definition. The separation of the $p_F$ distributions is 36% in average between proton and iron. The width of the $p_F$ distribution is typically 32% for proton and 14% for iron, respectively.

The WFCTs watches shower longitudinal development from a distance of $R_p$ and take it as the shower image. Because the shower geometry is precisely measured by WCDA, the distribution of number of photoelectrons in the image along the shower axis describes the shower profile at certain precision. The centroid of the image indicates the shower maximum position in sky while the direction of the shower points the start of the shower. The angular distance $\Delta \theta$ between the shower arrival direction and the centroid could be used to measure
the atmospheric depth for shower maximum. However, the image is stretched longer for farther showers due to pure geometric effect. Defining the parameter $p_X = \Delta \theta - 0.0097 R_p - 0.47 \log_{10} N_{pe}^0$, one can reduce the geometrical effect, and the shower elongating effect. The separation of the $p_X$ distributions is about 23% in average between proton and iron. The width of the $p_X$ distribution is typically 28% for proton and 21% for iron, respectively.

The ratio of length $L$ and width $W$ of the shower Cherenkov image taken by WFCTs, as mentioned above, can be used to define a dimensionless parameter, $p_C = L/W - 0.0139 R_p + 0.267 \log_{10} N_{pe}^0$. Here, the pure geometrical elongation effect of the image due to the spatial distance of the shower axis from the telescope and the energy dependence are reduced in the definition of $p_C$. The separation of the $p_C$ distributions is 42% in average between proton and iron. The width of the $p_C$ distribution is typically 20% for proton and 24% for iron, respectively.

The $\mu$-content is measured using the MDs in the array surrounding WCDA. For showers well contained in WCDA, the registered MDs distribute at least 30 m away from the shower core and spread out a large area in MDA. Simulation shows that the average number of muons recorded by all MDs is about 18 for showers between 100 TeV and 130 TeV. Due to the rather rapidly falling lateral distribution of muons in a shower and the uneven distribution of the detectors respect to WCDA, the number of muons recorded by MDs varies very much depending on the shower core position in WCDA. A fitting procedure is developed to obtain the total muon content $N_{\mu}$ in the shower. Below 10 PeV, the energy dependence of the muon content is nearly universal for all species. The separation of the distributions of parameter describing the muon content, defined as $p_\mu = \log_{10} N_{\mu} - 0.982 \log_{10} N_{pe}^0$ to reduce the shower size dependence, is about 3.7% between proton and iron, with typical widths of 3.4% for proton and 2.8% for iron, respectively.

Figure 2 shows the one-to-one correlation between the described parameters for different species of well reconstructed events generated assuming a primary composition of equal-weighted 5 mass groups ($H$, $He$, $CNO$, $AlMgSi$, $Fe$) and a spectral index being -2.7 below the knee and -3.1 above, respectively. The knee for each spectrum is assumed to be at $700Z$ TeV where $Z$ is the average charge of the mass group. Further analysis with different assumptions on mass composition and spectral indices will estimate the corresponding
uncertainty. Selections for samples of either pure protons or pure iron are rather straightforward using the two-parameter analysis by setting cuts on the correlation maps. Using an assumption proposed by Horandel\textsuperscript{13)} for heavier compositions as the background of the pure proton sample, the purity of 90% for the proton sample can be reached with sufficiently high selecting efficiency. For a mixed sample of proton plus Helium nuclei, $H&He$, it is possible to reach an even higher purity such as 95%. An example of using the particle density near the shower core instead of the energy flux $p_F$ and the parameter $p_C$ of the Cherenkov image, has been the analysis of the data produced by the combined experiment with one LHAASO prototype telescope and the ARGO-YBJ RPC carpet detector. With this technique, the important discovery of the knee at energy below 1 PeV in the $H&He$ spectrum has been made (ref. \textsuperscript{1}). An enhancement of a factor of $\sim 18$ in statistics with the $1/4$ LHAASO array is foreseen in one year of measurement. The expected number of events per year for proton and $H&He$ samples with the corresponding purities are shown in Figure 3.

![Figure 2](image_url)  

**Figure 2:** One-to-one correlation between the parameters $p_F$, $p_X$, $p_C$ and $p_\mu$ for different species of well reconstructed events (see the text for the assumption on composition and spectral index). The color indicates the particle type (black for proton, red for helium, green for CNO, blue for MgAlSi and pink for Fe). The last figure shows the $p_\mu$ distributions for proton and iron showers, as an example of how the two species can be separated using this parameter.
To separate other species, such as Helium or CNO, out from all reconstructed events, more sophisticated analysis techniques in the multi-parameter measurements have been under development, for instance the Artificial Neuron Network, Boosted Decision Tree or other methods will be used in the analysis.

Figure 3: The integrated distribution of number of events per year for 1/4 LHAASO array. The open dots represent pure proton samples with a purity of 90% and the filled dots represent H&He samples with a purity of 95%, respectively. The duty cycle is assumed to be 15%.

4 Summary

Separation between proton or H&He showers from well measured air shower samples by 1/4 LHAASO array is briefly discussed in this paper. With multiple parameters, i.e. shower energy, \( p_F, p_X, p_C \) and \( N_\mu \), being measured, one can select pure proton or H&He samples with high purity by applying simple cuts on the one-to-one correlation maps. More than 2500 proton events above the knee of 700 TeV would be collected in one year operation with an assumption of 15% duty cycle. The heavier nuclei can be also separated out by more sophisticated analyses which are under development. A pure proton samples with well determined shower energy are much more useful than simply measuring the spectrum. For instance, they will help to understand many details of interaction models, particularly for the muon production. It is well known that there is still big uncertainty in the models. One of the good feature of such a sample is that the energy of protons is determined independent of the
muon content of showers.

5 Acknowledgements

This work is supported in China by the Chinese Academy of Sciences (CAS) (0529110S13), the Key Laboratory of Particle Astrophysics, Institute of High Energy Physics, CAS. Projects No. 11475190 and 11505190 of NSFC also provide support to this study.

References

HIGH-ENERGY COSMIC RAYS: GALACTIC OR EXTRAGALACTIC?

Jörg R. Hörandel\textsuperscript{1,2,0}, A. Achterberg\textsuperscript{1}, S. Buitink\textsuperscript{1,3}, A. Corstanje\textsuperscript{1}, H. Falcke\textsuperscript{1,2,4,5}, A. Nelles\textsuperscript{1,2,7} J.P. Rachen\textsuperscript{1}, L. Rossetto\textsuperscript{1}, P. Schellart\textsuperscript{1,8}, O. Scholten\textsuperscript{6}, S. ter Veen\textsuperscript{4}, S. Thoudam\textsuperscript{1,9} T.N.G. Trinh\textsuperscript{6}, A. van Vliet\textsuperscript{1}

\textsuperscript{1}Department of Astrophysics, IMAPP, Radboud University Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands
\textsuperscript{2}NIKHEF, Science Park Amsterdam, Amsterdam 1098 XG, NL
\textsuperscript{3}Astrophysical Institute, Vrije Universiteit Brussel, Pleinlaan 2, Brussels 1050, Belgium
\textsuperscript{4}ASTRON, Postbus 2, Dwingeloo 7990 AA, The Netherlands
\textsuperscript{5}Max-Planck-Institut für Radio Astronomie, Bonn, Germany
\textsuperscript{6}KVI-CART, Groningen University, P.O. Box 72, Groningen 9700 AB, NL
\textsuperscript{7}now at: Department of Physics and Astronomy, University of California Irvine, Irvine, CA 92697-4575, USA
\textsuperscript{8}now at: Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA
\textsuperscript{9}now at: Department of Physics and Electrical Engineering, Linneuniversitetet, 35195 Växjö, Sweden

Abstract

The bulk of cosmic rays is assumed to be accelerated in our Milky Way, while the highest-energy particles in the Universe are attributed to extra-galactic sources. A transition is expected at energies around $10^{17}$ to $10^{18}$ eV. A new method is presented to measure the properties of cosmic rays in this energy region: the radio detection of air showers. And a comprehensive model is discussed to describe consistently the Galactic and extra-galactic components of cosmic rays from GeV energies up to $\sim 10^{20}$ eV.

1 Introduction

Cosmic rays (ionized atomic nuclei) impinge on the Earth with (kinetic) energies covering a wide range from MeV energies up to beyond $10^{20}$ eV. At
energies below $\sim 100$ MeV they are accelerated in energetic outbursts of the Sun. At higher energies, the are assumed to originate in our Milky Way, being accelerated in Supernova remnants (e.g. $^1$, $^2$). At energies exceeding $10^{18}$ eV it becomes increasingly difficult to magnetically bind the particles to our Galaxy. Thus, particles with energies above $\sim 10^{18}$ eV are usually considered to be of extra-galactic origin. A transition from a Galactic to an extra-galactic origin of cosmic rays is expected at energies around $10^{17}$ to $10^{18}$ eV $^3$, $^4$).

In this paper we will shed new light on the understanding of the origin of cosmic rays in the transition region ($10^{17} - 10^{18}$ eV). This necessitates a precise measurement of the properties of cosmic rays, namely their arrival direction (on the sky), their (kinetic) energy, and their particle type (atomic mass $A$).

The flux of cosmic rays is steeply falling, approximately following a power law $\propto E^{-3}$. In our region of interest, cosmic rays are only measured indirectly, using large ground-based detector installations. High-energy cosmic rays impinging on the atmosphere, initiate cascades of secondary particles, the extensive air showers. The challenge of the indirect measurements is to derive the properties of the incoming cosmic rays from air-shower observations. Most challenging is the measurement of the particle type, since the sensitivity of air shower measurements is only proportional to $\ln A$. Intrinsic shower fluctuations allow to divide the measured cosmic rays in up to five mass groups for the best experiments $^5$).

A (new) method to measure the properties of cosmic rays via the radio detection of air showers is described in Sect. 2. These measurements yield one of the key observables, the evolution of the mass composition of cosmic rays as a function of energy. These observations are the basis for a consistent model for the origin of Galactic and extra-galactic cosmic rays as outlined in Sect. 3, with particular emphasis on the transition region $10^{17} - 10^{18}$ eV.

2 Radio detection of air showers

Many secondary particles in extensive air showers are electrons and positrons. They emit radiation with frequencies of tens of MHz mainly due to interaction with the magnetic field of the Earth. Radio detection of air showers is suitable to measure the properties of cosmic rays with nearly 100% duty cycle $^6$, $^7$). The LOFAR radio telescope $^8$, $^9$) is one of the leading installations for the radio measurements of air showers.
In the last years the radio technique has been established as a precise method to measure the mass composition of cosmic rays. The LOFAR measurements together with the predictions of the CoREAS simulation package result in a complete understanding of the emission mechanisms. With LOFAR the properties of the radio emission have been measured with high accuracy \(11, 12, 13\) in the frequency range \(30 - 80\) MHz, which allows us to establish key features, such as the lateral density distribution of the radio signals \(14, 15\), the shape of the shower front \(16\) – important to reconstruct the arrival direction of the incoming cosmic ray, and the polarization of the radio signal \(17\). These measurements help to understand the emission processes in the atmosphere and to quantify the contributions of the two mechanisms, being responsible for the radio emission of air showers – namely the geomagnetic effect (i.e. charge separation in the geomagnetic field) and the Askaryan effect (charge excess in the shower front). We obtained the first quantitative measurements in the frequency range \(120 - 240\) MHz \(18\). We also recorded air showers during thunderstorm conditions \(19, 20\) and measured the structure of electric fields in the atmosphere.

The good agreement between the measurements and the predictions of the CoREAS code is essential to identify the type of incoming cosmic ray. This is inferred from the (atmospheric) depth of the shower maximum \(X_{\text{max}}\), one of the standard measures to estimate \(\ln A\). To measure \(X_{\text{max}} \) \(22, 23\) we analyse simultaneously measurements of the radio emission and the particle detectors, to determine \(X_{\text{max}}\) with an accuracy of \(\sim 20\) g/cm\(^2\) with the dense LOFAR core, thus, reaching the state of the art – the uncertainty of the Pierre Auger Observatory fluorescence detector. The measured values for the depth of the shower maximum \(X_{\text{max}}\) are used to derive the mean logarithmic mass of cosmic rays

\[
\langle \ln A \rangle = \left( \frac{X_{\text{max}} - X_{\text{p max}}}{X_{\text{Fe max}} - X_{\text{p max}}} \right) \times \ln A_{\text{Fe}}.
\]

This necessitates predictions for the depth of the shower maximum for impinging protons and iron nuclei, \(X_{\text{p max}}\) and \(X_{\text{Fe max}}\), respectively. The resulting mean mass is depicted in Fig. 1 as a function of energy for the LOFAR results together with the world data set \(21\). Two hadronic interaction models are used (EPOS and QGSJET) to interpret the data.
Figure 1: Mean logarithmic mass $\ln A$ as measured by various experiments, interpreted with two hadronic interaction models. In addition, predictions are shown for three different models of the additional Galactic component: cosmic rays from Wolf-Rayet stars ($\text{C}/\text{He} = 0.1$ and $\text{C}/\text{He} = 0.4$), and cosmic rays being re-accelerated by the Galactic wind. See $21$ for further details.

3 The transition from Galactic to extra-galactic cosmic rays

To understand the implications of the LOFAR measurements and the available world data set from direct and indirect measurements a model has been developed to consistently describe the observed energy spectrum and mass composition of cosmic rays with energies up to about $10^{18}$ eV $21$. We assume that the bulk of Galactic cosmic rays is accelerated by strong Supernova remnant shock waves $24$. After acceleration, cosmic rays undergo diffusive propagation through the Galaxy. During the propagation, cosmic rays may again encounter expanding Supernova remnant shock waves, and get re-accelerated. As the probability of encountering old Supernova remnants is expected to be larger than the younger remnants because of their bigger sizes, re-acceleration is expected to be produced mainly by weaker shocks. Since weaker shocks generate a softer particle spectrum, the resulting re-accelerated component will have a spectrum steeper than the initial cosmic-ray source spectrum produced by strong shocks. For a reasonable set of model parameters, it is shown that the re-accelerated component can dominate the GeV energy region while the non-
re-accelerated component dominates at higher energies, thereby explaining the (recently) observed GeV-TeV spectral anomaly.

We assume a source spectrum for the individual cosmic-ray components at the sources proportional to a power law in total momentum $p$ with an exponential cut-off, which can be written in terms of momentum/nucleon as

$$Q(p) = A Q_0 (a p)^{-q} \exp \left( - \frac{A p}{Z_{pc}} \right),$$

where $Q_0$ is a normalization constant, $q$ is the spectral index, and $p_c$ is the cut-off momentum for protons. We assume that the maximum energy of cosmic rays attained during the acceleration process is proportional to the nuclear charge number $Z$: $E_c = Z \cdot 4.5 \cdot 10^6 \text{ GeV}$. With these assumptions the energy spectra for individual elements in cosmic rays are perfectly described from the lowest energies (direct measurements at $\sim 1 \text{ GeV}$) up to about $10^{16} \text{ eV}$.

Our study shows that a single Galactic component with rigidity-dependent energy cut-offs in the individual spectra of different elements cannot explain the observed all-particle spectrum at energies exceeding $\sim 2 \cdot 10^{16} \text{ eV}$. Similar findings have already been obtained earlier \cite{25}. We discuss two approaches for a second component of Galactic cosmic rays: re-acceleration at a Galactic wind termination shock and Supernova explosions of Wolf-Rayet stars.

 Galactic winds can lead to the production of an additional component of cosmic rays which can dominate at high energies. Galactic winds, which start at a typical velocity of about few km/s near the disk, reach supersonic speeds at distances of a few tens of kpc away from the disk. At about a hundred kpc distance, the wind flow terminates resulting into the formation of termination shocks. These shocks can encounter cosmic rays escaping from the disk into the Galactic halo, and re-accelerate them via the diffusive shock acceleration process. The re-accelerated cosmic rays can return to the disk through diffusive propagation against the Galactic wind outflow. For an energy dependent diffusion process, only the high-energy particles may be effectively able to reach the disk. In order to describe the observed all-particle spectrum around $10^{16}$ to $10^{18} \text{ eV}$ we assume an injection efficiency of 14.5% and a cut-off energy for protons of $9.5 \cdot 10^7 \text{ GeV}$.

While the majority of the Supernova explosions in the Galaxy occur in the interstellar medium, a small fraction is expected to occur in the winds of massive progenitors like Wolf-Rayet stars. Magnetic fields in the winds of
Wolf-Rayet stars can reach of the order of 100 G, and it has been argued that a strong Supernova shock in such a field can lead to particle acceleration up to energies of $\sim 10^9$ eV $^{27}$. We estimate a frequency of $\sim 1$ Wolf-Rayet explosion every 210 years. This corresponds to $\sim 1$ Wolf-Rayet explosion for every 7 Supernova explosions occurring in the Galaxy. The source indices of the different cosmic-ray species and the propagation parameters for the Wolf-Rayet cosmic rays are taken to be the same as for the ‘regular’ component from Supernova remnants. Different elemental compositions of the Wolf-Rayet winds are discussed in the literature: a carbon-to-helium (C/He) ratio of 0.1 and 0.4. The latter scenario can explain almost all observed features in the all-particle spectrum and the mass composition of cosmic rays up to $\sim 10^{18}$ eV, when combined with a canonical extra-galactic spectrum as expected from strong radio galaxies or a source population with similar cosmological evolution. The resulting spectrum is shown in Fig. 2. In this two-component Galactic cosmic-
ray model, the 'knee' at $\sim 4 \cdot 10^{15} \text{ eV}$ and the 'second knee' at $\sim 4 \cdot 10^{17} \text{ eV}$ in the all-particle spectrum are due to the cut-offs of the first and second Galactic cosmic-ray components, respectively.

Finally, at energies above $10^{18} \text{ eV}$ several assumptions for an extra-galactic component have been investigated: from a minimal contribution to scenarios with a significant component below the 'ankle' (at $\sim 4 \times 10^{18} \text{ eV}$). It has been found that extra-galactic contributions in excess of regular source evolution are neither indicated nor in conflict with the existing data. We find that an extra-galactic contribution is unlikely to dominate at or below the second knee. The main result is that the second Galactic component predicts a composition of Galactic cosmic rays at and above the second knee that largely consists of helium or a mixture of helium and CNO nuclei, with a weak or essentially vanishing iron fraction, in contrast to most common assumptions. This prediction is in agreement with new measurements from LOFAR and the Pierre Auger Observatory which indicate a strong light component and a rather low iron fraction between $\sim 10^{17}$ and $\sim 10^{18} \text{ eV}$.

4 Summary

The radio detection of extensive air showers enables us to measure the properties of cosmic rays (arrival direction, energy, and particle type) above energies exceeding $10^{17} \text{ eV}$ with high precision.

We developed a model to consistently describe the observed energy spectrum and mass composition of cosmic rays from GeV energies up to $10^{20} \text{ eV}$. We adopt a three component model: 'regular' cosmic rays being accelerated in Supernova remnants up to $\sim 10^{17} \text{ eV}$, a second Galactic component, dominating the all-particle flux between $\sim 10^{17}$ and $\sim 10^{18} \text{ eV}$ from cosmic rays being accelerated by exploding Wolf-Rayet stars, yielding a strong contribution of He and CNO elements, and, finally, an extra-galactic contribution at energies above $\sim 10^{18} \text{ eV}$.

5 Acknowledgements

JRH is grateful to the organizers of the Vulcano workshop for their kind invitation and the great hospitality on the Vulcano island.
References

THE NATURE AND ORIGIN OF ULTRA-HIGH ENERGY COSMIC RAY PARTICLES

Peter L. Biermann$^{1,2,3,4}$, Laurentiu I. Caramete$^{5,1}$, Federico Fraschetti$^{6}$, László Á. Gergely$^{7}$, Benjamin C. Harms$^{3}$, Emma Kun$^{7}$, Jon Paul Lundquist$^{8}$, Athina Meli$^{9}$, Biman B. Nath$^{10}$, Eun-Suk Seo$^{11}$, Todor Stanev$^{12}$, and Julia Becker Tjus$^{13}$.

Abstract

We outline two concepts to explain Ultra High Energy Cosmic Rays (UHECRs), one based on radio galaxies and their relativistic jets and terminal hot spots, and one based on relativistic Super-Novae (SNe) or Gamma Ray Bursts (GRBs) in starburst galaxies, one matching the arrival direction data in the South (the radio galaxy Cen A) and one in the North (the starburst galaxy M82). The most likely identification of the origin of observed Gravitational Wave (GW) events is stellar binary black hole (BH) mergers in starburst galaxies such as M82 with the highest rate of star formation, so the highest far-infrared (FIR) luminosity, at the edge of the universe visible in 10 - 300 Hz GWs; at low heavy element abundance $Z_{ch}$ the formation of stellar BHs extends to a larger mass range. A radio galaxy such as Cen A sequence of events involves first the merger of two Super-Massive Black Holes (SMBHs), with the associated ejection of low frequency GWs, then the formation of a new relativistic jet aiming into a new direction: ubiquitous neutrino emission follows accompanied by compact TeV photon emission, detectable more easily if the direction is
towards Earth. The ejection of UHECRs is last. Both these sites are the perfect high energy physics laboratory: We have observed particles up to ZeV, neutrinos up to PeV, photons up to TeV, 30 - 300 Hz GW events, and hope to detect soon of order $\mu$Hz to mHz GW events. Energy turnover in single low frequency GW events may be of order $\sim 10^{53}$ erg. How can we further test these concepts? First of all by associating individual UHECR events, or directional groups of events, with chemical composition in both the Telescope Array (TA) Coll. and the Auger Coll. data. Second by identifying more TeV to PeV neutrinos with recent SMBH mergers. Third by detecting the order $< \mu$Hz GW events of SMBH binaries, and identifying the galaxies host to the stellar BH mergers and their GW events in the range up to 300 Hz. Fourth by finally detecting the formation of the first generation of SMBHs and their mergers, surely a spectacular discovery.

1 Introduction: Challenges of High Energy Events

Today we have an abundance of riches, with almost certainly more to come: We have a very well defined spectrum of ultra high energy cosmic ray (UHECR) particles, with chemical composition information, and a kink up in the overall spectrum, near $3 \cdot 10^{18}$ eV. We have directional information, with a weak directional hot spot in the South, and a much better defined directional hot spot in the North, both suggesting specific galaxies as sources, that have long been ranked as the leading candidates of their activity in the local universe: The radio galaxy Cen A in the South, with a recent SMBH binary merger, and the starburst galaxy M82 in the North with relativistic Super-Novae (SNe) or Gamma Ray Bursts (GRBs). Starburst galaxies just as radio galaxies can

---

produce both energetic particle populations up to order $10^{21} Z \text{ eV}$, where $Z$ is the charge of the nucleus.

We also have very high energy neutrino events, that may originate in recent mergers of Super-Massive Black Holes (SMBHs): How can we recognize such events? In a merger between two galaxies, both with a super-massive black hole at the center, orbital angular momentum wins, and so leads to a new direction of the final spin. The final spin direction defines the axis of the new relativistic jet, which has to plow a new channel through the dense material near the newly defined center of the merged galaxy. This plowing leads to powerful injection, acceleration and particle interaction, therefore giving rise to lots of high energy neutrinos, detectable in the case the new jet points at Earth, and then recognizable via a flat spectrum to near THz radio frequencies. Thus, these mergers give rise to a low frequency GW background, yet to be discovered, in the range from order 1 mHz on down.

Furthermore, we now have the detections of GW Events from the merging of stellar black holes. This reminds us that massive stars almost all are in comparable mass binary star systems, and so the GW event rate ought to scale directly with the supernova rate of massive stars detectable in their radio emission. The supernova rate in turn scales with the FIR luminosity. Some of all supernovae also explode as relativistic SNe or GRBs, allowing yet higher energy particles to be produced. And so any starburst galaxy produces in parallel SNe, GRBs, and as a consequence particles to $10^{17.3} Z \text{ eV}$ from SNe, up to order $10^{21} Z \text{ eV}$ from relativistic SNe or GRBs, the corresponding high energy neutrinos, and GW events in the 10 - 300 Hz range.

2 The spectrum of cosmic rays

In combination of Auger Coll. 1) and the Telescope Array Coll. (TA) 5) data with other experiments we now have a consistent spectrum of cosmic rays, that readily finds an interpretation 131); for recent reviews see 79, 73), and fundamental books 124, 51). This explanation requires a contribution from massive stars such as Wolf-Rayet stars, which have powerful stellar winds. In this context it is of interest to note that we have good radio data on supernova explosions racing through a wind 34, 50, 117, 118, 119, 120) 74, 93, 70, 140). These early data, covering six explosions that probably were Wolf-Rayet stars, so Blue Supergiant stars, and two explosions that probably were a Red Super-
giant star, all suggest the same: i) The speed of the shock is initially about 0.1 c, ii) the magnetic field in the post-shock region is about 1 Gauss at a radial distance of $10^{16}$ cm, iii) the run of the magnetic field is close to $r^{-1}$ with $r$ being the radial coordinate. The numbers allow the two limits of the energy that can be reached to be worked out, i) in the limit that the magnetic field is mainly parallel to the shock normal, the Bohm limit \textsuperscript{41}); and conversely ii) in the limit that the magnetic field is mainly parallel to the shock surface, the Jokipii limit \textsuperscript{67}). These numbers suggest $E_{\text{Bohm}} = 10^{15.3 \pm 0.3} Z$ eV; and $E_{\text{Jokipii}} = 10^{17.3 \pm 0.2} Z$ eV, where $Z$ is the charge of the particle. It is striking that the first energy corresponds well to the knee, and the second energy corresponds well to the ankle, suggesting that both scattering regimes are important, Bohm and Jokipii. The very fact, that for the two types of stars, with very different wind velocities, and so different wind densities, give rise to the same magnetic field, precludes the Bell-Lucek mechanism \textsuperscript{82, 20}) as the key path to attain these magnetic fields in the observed shock. The Bell-Lucek mechanism may have worked prior to the SN-explosion, when very often strong eruptive events are detected \textsuperscript{128, 52, 98, 127, 130}), as also observed for $\eta$ Carinae in the 19th century. We note that the main sequence stages of these stars do have a magnetic field, but usually of order only a few hundred Gauss \textsuperscript{65, 69, 86, 137, 138}). However, as these stars had an inner convective zone with magnetic fields attaining the $10^6$ or even $10^7$ Gauss range \textsuperscript{23}) and the surface of the star due to the strong winds is reaching deep down, it is also possible that these magnetic fields seen in the explosion were derived from the former inner magnetic fields. That would be consistent with their properties being so similar in the cases well observed sofar.

Relativistic SNe or GRBs would obviously enhance the particle energies further, probably reaching $10^{21}$ Z eV; however, the IceCube Coll. evidence \textsuperscript{4}) speaks against a dominant GRB contribution of the UHECR flux. On the other hand, the TA Coll. evidence suggests a proton contribution from the starburst galaxy M82, see below.

3 Distribution of arrival directions: all sky

The Auger Coll. finds a weak correlation in arrival direction of events with the radio galaxy Cen A, consistent with a very old prediction \textsuperscript{56}). The TA Coll. finds a much more pronounced directional “hot spot” in
arrival directions not far from the starburst galaxy M82, consistent with the expectation that relativistic SNe or GRBs are a special form of massive star explosions, that lead to the acceleration of protons and nuclei to energies near $10^{21}$ eV (e.g. 24, 25); for a quantitative approach see 92, 136, 139), with many later papers and reviews, e.g. 101, 102, 66, 90). We note that close to 100 percent of all massive stars are in binary systems of comparable mass 39), allowing mass transfer between the two stars and also the formation of two stellar BHs 62).

The spread of arrival directions (of order 10 degrees) as well as the central shift in direction (near 0 degrees for Cen A, and near 20 degrees for M82) can be understood as scattering and orbit bending in the Galactic wind of our Galaxy, subject to a $k^{-2}$ turbulence, with $k$ the wavenumber, indicative of what can be described as compressible turbulence, super-sonic and/or super-Alfvénic turbulence, or shock-dominated irregularities 71, 72, 48, 31). A detailed exploration of the implications (see, e.g., the Faraday sky in 99)), and the evidence for a Galactic wind 43, 44), and 134) has yet to be done.

How can we check on such interpretations? In the case of a starburst galaxy such as M82 the magnetic field of a Galactic wind, even when highly irregular, ought to impose a certain geometric pattern on the arrival directions, strongly depending on the charge of the particles, since all bending runs with energy over charge $E/Z$; there is now some evidence of this geometry. Additionally, along with massive star SNe there ought to be occasional events of GWs, such as observed already 7, 8). The rate of massive star SNe as well as GW events ought to scale with the rate of massive star formation, so the luminosity of a starburst in the FIR 21, 75, 111, 77). The higher the rate, the more likely it is to detect an event within a certain time window; on the other hand, with a luminosity function decreasing with luminosity 77) we have the highest such rate near the limit of sensitivity, when the lowest detectable flux corresponds to the most powerful starburst within the survey volume. Lagache et al. state “Luminosity function evolution is such that the power output is dominated by LIRGs at $z \simeq 0.7$ (although they represent only 3% of the galaxies) and by ULIRGs at $z \simeq 2.5$ (although they represent only 1% of the galaxies).”, so that the brightest galaxies will dominate in sampling GW events; here LIRG means “Luminous Infra-Red Galaxy”, with a star formation rate of about $10 - 100 M_\odot yr^{-1}$, and ULIRG means “Ultra-Luminous Infra-Red
Galaxy”, with a yet higher star formation rate and possibly additional feeding from an Active Galactic Nucleus (AGN). However, if relativistic SNe or GRBs and GW events depend on the formation of Black Holes, for which the mass range depends on the heavy element abundance $Z_{ch}$, then all this may get more subtle, and will require more exploration. To be compatible with the IceCube Coll. limits, it appears as a testable concept that the directional hot spot due to M82 is only a localized patch of very energetic protons, on top of a $4\pi$ spread of nuclei from the radio galaxy Cen A; this may explain why the TA Coll. flux tends to be a tad above the Auger Coll. flux.

In the case of radio galaxies it is of interest to focus on merger events of super-massive binary black holes, which generally give rise to a new spin direction of the SMBH after the merger: in that case the jet has to plow a new channel, maximizing injection, acceleration, and interaction, so providing a prime site for the acceleration of UHECRs: Kun et al. argue that in the case where after the merger the relativistic jet points at Earth this stage can be identified by observing a flat radio spectrum from GHz frequencies all the way to the FIR, so near THz radio frequencies, such as detected by IRAS, WMAP or PLANCK. A first search identified two neutrino track events (with about a degree in directional uncertainty: 2, 3)) with such radio sources, with a very low combined probability that the identification is random. The task remains to identify more track events. This interpretation implies that recent SMBH mergers are prodigious producers of UHECRs. It also implies that there ought to be a strong background of gravitational waves in the range order $1 \, \mu \text{Hz}$ to $1 \, \text{mHz}$, allowing for the main part of the spectrum of SMBHs, between $\sim 10^6$ and $\sim 10^8 M_\odot$, and a possibly relevant redshift range between 10 and 100. The observation of such slow GW events will remain a task for the future, while the Lisa Path Finder mission allows grounds for optimism in the long term.

4 The spectrum below $3 \cdot 10^{18}$ eV, the ankle

LOFAR and Kaskade-Grande have demonstrated that the long expected extragalactic proton component may be there at these lower energies. One unanswered question here is whether this flux results from a long time integral of the locally produced high energy cosmic ray particles or is the long distance accumulation from various active sources. This distin-
tion depends on the probably extremely inhomogeneous structure of the large scale intergalactic magnetic field. This magnetic field is embedded into the accretion flow towards filaments and sheets, and so there ought to be a specific energy beyond which particles can no longer escape this accretion flow and just propagate along filaments and sheets. This has focussing as consequence, along filaments the flux of particles caught does not diminish with distance except by scattering to higher energy and ensuing escape, and along sheets the flux weakens with inverse distance \(110\). This could produce a strong spectral turn-down feature, which is not seen except at a characteristic energy of about \(6 \cdot 10^{19}\) eV. This turn-down could also be due to spatial limits to acceleration in the source, or to interaction with the microwave of far-infrared background. We see no such feature at lower energy within the current data.

5 Jets and terminal hot spots in relativistic jets

Shocks in relativistic flow appear as prime candidates to inject and accelerate protons and heavier nuclei to ultra high energy; remember, that GRBs are commonly interpreted as relativistic jets with initially very high Lorentz factor (to several hundred), while relativistic jets emanating from SMBHs may range from Lorentz factor barely above unity to about 100 (see, e.g., \(59\)); relativistic SNe are one other possibility \(66\). These shocks may plow through a starburst region with the standard galactic cosmic ray spectrum and use them for injection to ultra high energy \(59, 29, 132\).

Here we briefly review the relevant arguments based on the series of papers started by \(45, 46\), and \(84\), based on some ideas in \(22\), and used again in \(28, 30, 31\).

First some comments on jets: Jet flow suffers dramatically from adiabatic losses, and yet keeps going from near the central BH to \(3,000 r_s\) to order \(10^{24}\) cm, sometimes to \(10^{25}\) cm or even more, where \(r_s\) is the Schwarzschild radius of the central BH. Jets can be expected to start at a near-relativistic speed of sound, but cool down rapidly. Each shock system consumes only a minute fraction of kinetic energy (remember, entropy is increased in any shock). The observational evidence suggests a spiral (almost DNA-like geometry) pattern of highly oblique shocks: so we have continuous re-acceleration (see, also, \(144\)) of the particle population \(88\). However, each shock is strong suggesting that the internal sound-speed is sub- or only weakly relativistic.
The ubiquitous cutoff spectrum of non-thermal emission near $3 \cdot 10^{14}$ Hz observed $10^5, 10^6, 10^8, 10^7, 32, 125, 116) \ 33, 109, 87, 100$ since the mid-seventies in jets, terminal hot spots and compact unresolved Active Galactic Nuclei can be explained as the combined effect of first protons (or nuclei) getting accelerated to the synchrotron loss limit, and giving rise to a $E^{-2}$ spectrum. Such a spectrum results in a $k^{-5/3}$ spectrum of magnetic irregularities (the same spectrum as $71, 72$), but here via excitation at all wavelengths, $18, 19$).

Note that nearly every shock is preceded by another shock with injection of both turbulence and energetic particles further upstream $88$). Electrons then get accelerated in that same spectrum of irregularities again to their loss limit, and at their maximal energy give a maximal synchrotron frequency independent of all parameters, $\nu_e^* \lesssim 3 \cdot 10^{14}$ Hz. Generalizing for synchrotron losses and photon interaction losses is straight-forward and does not modify the numbers substantially. This translates to a maximal energy for protons of $E_{p,max} \simeq 1.4 \cdot 10^{20} \text{eV} \ (\nu_e^*/3 \cdot 10^{14} \text{Hz})^{1/2} B^{-1/2}$, where $B$ is the magnetic field in Gauss, typically observed to be of order mGauss in compact nuclei. There is also a spatial limit $10^{21} L_{46}^{1/2}$ eV, where $L_{46}$ is the jet power in units of $10^{46}$ erg/s $81, 46$; we note that the maximal magnetic field in jets is given by a Poynting flux jet which scales with the square of the magnetic field. Therefore we have in combination $E_{p,max} \simeq 1.4 \cdot 10^{21}$ eV (no boosting assumed here). Thus UHECR particles are required to explain why the feature is so ubiquitous.

Combining then the maximal emission frequencies of protons and electrons we find $\nu_{\text{syn,p,max}} / \nu_{\text{syn,e,max}} = \{m_p/m_e\}^3$, matching the first characteristic of the double-bump spectrum of blazars $28$, the spectral distance between the two bumps. Integrating then downstream (see also $22, 114, 143$) from each shock along a stream-line following $68$) gives

$$\{L_p/L_e\} \sim \{n_p/m_p n_e/m_e\} \{\gamma_{p,max}/\ln(\gamma_{e,max}/\gamma_{e,min})\} \{m_e/m_p\}^{3/3} \sim 1$$

matching the the second characteristic of the double-bump spectrum of blazars, the crudely equal luminosity of the two bumps (e.g. $57, 121$). In this approach the blazar sequence arises from the dependencies on SMBH mass and boosting factor $28$. We propose this is the basic explanation of this observation. Quite obviously, many interactions such as Inverse Compton of these two photon-bumps ensue and modify what we observe. This requires proton (or nuclei) acceleration in the sources, in radio galaxies: after all, radio galaxies, perchance pointing at us, are blazars.
One important test is the variability time-scale $\tau$: 
$$\tau \simeq \frac{r}{(2 \Gamma^2 c)} \times 10^{1}\text{ s},$$
where $r$ is the radial distance, and $\Gamma$ is the Lorentz factor of the jet; minutes seen in TeV photons $^{9,10}$ imply $10^{16.6}$ cm, $\simeq 3000 \, r_S$ of a $10^8 \, M_\odot$ SMBH for $\Gamma \simeq 100$. This is near where jets turn into a conical outflow $^{141,85,84}$.

This is an old prediction, but a new argument: Radiogalaxies are sources at energies $> 10^{20}$ eV, with an energy that may approach $\sim 10^{63}$ erg, possibly on occasion even more. Lower energy budget CR-sources are intergalactic shocks $^{64}$, Gamma Ray Bursts $^{102}$, micro-quasars $^{63,94}$, jet-supernovae, pulsar wind nebulae $^{47}$, powerful supernovae $^{26,66}$, and probably yet other activities we do not understand yet. Ginzburg & Syrovatskij $^{56}$ identified the nearby candidates before the discovery of UHECRs: the radio galaxies i) Cen A ($= $ NGC 5128), a recent SMBH merger; ii) Vir A ($= M87 = $ NGC 4486), a recent SMBH merger; and iii) For A ($= $ NGC 1316), perhaps also a SMBH merger, but in this case the radio morphology is ambiguous, while for Cen A and M87 the old jet directions are clearly visible, so a re-orientation of the dominant jet is recognizable, a direct consequence of the merger of two SMBHs. Considering the energetics of the observed compact jets in radio galaxies, a hierarchical ranking can be derived of what radio galaxies may contribute $^{38}$, and the first three radio galaxies are just those already identified by $^{56}$; the detailed statistics show, that Cen A is expected to dominate the entire integral of the UHECR contribution from radio galaxies lower down in the ranks.

5.1 How do SMBHs start?

The observational evidence suggests that the SMBH mass function $^{36}$ starts around $3 \cdot 10^6 \, M_\odot$, and its shape can be fully explained by merging in the gravitational focussing limit $^{115,55}$, minimizing any electromagnetic output $^{42,36}$.

Why would star formation pick such a mass? First massive stars can form in dense groups in the gravitation of the Dark Matter potential well of a dwarf galaxy $^{126,40}$; then stars agglomerate $^{123,112}$ to form a more massive star. Massive stars also have winds, driven by radiation interaction with heavy elements ($^{83}$ and many later papers): So their maximum mass attainable is several hundred $M_\odot$ at most $^{142}$. However, at zero heavy element abundance massive stars can grow to much higher mass, close to $10^8 \, M_\odot$. At that point massive stars hit an instability, combining radiation pressure with subtle effects...
of General Relativity 10, 11). As a consequence they explode: So with infall their BH mass may reach about $3 \cdot 10^6 \, M_\odot$ explaining the observations. A key prediction of this specific model is that the initial formation of the first generation of SMBHs is only allowed at metal abundance $Z_{ch}$ near zero, a test which we hopefully can make in the future.

There are alternate pictures (e.g. 95, 96) using a specific model of Dark Matter. Other models just using massive stars and accretion would produce a large range in redshifts of when these SMBHs begin (e.g. 89), as would a concept in which gas collapses directly to a compact object 122), possibly a black hole.

The sky distribution of SMBHs allows further constraints on the origin to be set: it can be shown in a graph, where colors are distance: Black, Blue, Green, Orange, Red, for the redshifts intervals in steps of 0.005 to 0.025, i.e. distance intervals of 20, to 100 Mpc.

![Figure 1: The sky in super-massive black holes $> 3 \cdot 10^7 \, M_\odot$, where colors are distance: Black, Blue, Green, Orange, Red, for the five redshifts intervals in steps of 0.005 to 0.025. The coordinate system is with the Galactic Plane across the center, and Galactic Center (GC) at the right/left edge (See 36, 37).](image)

The striking feature in this sky distribution is the semi-circular feature of
hundreds of SMBHs, many of which are above $10^8 M_\odot$: This can be understood as a consequence of the freeze-out of the expansion of a spherical disturbance, as discussed elsewhere (lectures by P.L. Biermann at the Chalonge meetings in Paris 2015 and 2016, and ensuing discussions). Geometrically this could be due to the cut of an expanding baryonic spherical shell through a Dark Matter Zeldovich pancake, using the run-away cooling mechanism proposed in shocks due to supersonic flow; in this specific case the arc would correspond to the third bump in the Micro-Wave Back-Ground (MWBG) power spectrum. This may require a very high redshift to get started.

A test of any such picture would be the common detection of such partial circular arcs of SMBHs; considering the skymap shown this may indeed be a common occurrence, sometimes corresponding to the first peak in the MWBG power spectrum.

5.2 SMBH energetics

As shown in work by P.P. Kronberg (lecture at DRAO Nov 2015) allowing for $PdV$-work in understanding the scale of energy of giant radio galaxies the total energy may reach rather close to a good fraction of $M_{SMBH} c^2$, allowing for other channels than just radio emission, possibly as close as $\sim 1/2$ \cite{61}. The two observed GW stellar BH merger events correspond to about 0.05 of $M c^2$ in GWs emitted. One can speculate that stellar BHs ought to merge starting from a small spin, but that SMBHs in radio galaxies may start from near maximum spin, as it is derived from the orbital angular momentum of the merging SMBHs (mass and spin enter the maximally allowed efficiency). It follows that there ought to be a powerful GW background due to the formation and merging of SMBHs, in the range between order 1 $\mu$Hz and order 1 mHz, depending on the exact mass range and redshift of the first generation of formation of SMBHs and their merging history \cite{91,113,78,135}. As shown in \cite{27} this redshift could be quite high.

We note that the existing observational limits of any GW background all pertain to either today, the recombination redshift \cite{6}, or earlier epochs even. For any redshift $< 100$, the maximum redshift allowed by the mechanism in \cite{27} there is no limit at all today at the frequency range given by SMBHs other than the observed energy density of Dark Energy (DE).
6 Summary

We have outlined two concepts to explain UHECRs, one based on radio galaxies and their relativistic jets and terminal hot spots, and one based on relativistic SNe or GRBs in starburst galaxies, one matching the arrival direction data in the South (the radio galaxy Cen A) and one in the North (the starburst galaxy M82). The most likely identification of the origin of observed GW events is starburst galaxies such as M82 with the highest rate of star formation, so the highest FIR luminosity, at the edge of the universe visible in 10 - 300 Hz GWs; the value of the heavy element abundance $Z_{ch}$ restricts the mass range for stellar BHs \(^{62}\). The radio galaxy sequence of events involves first the merger of two SMBHs, with the associated burst of low frequency GWs, then the formation of a new jet aiming into a new direction: ubiquitous neutrino emission follows (detectable more easily if the direction is towards Earth) accompanied by compact TeV photon emission. The ejection of UHECRs is last.

So these sites are the perfect high energy physics laboratory: We have particles up to ZeV, neutrinos up to PeV, photons up to TeV, and of order $\mu$Hz to 300 Hz GW events; inside the source the energies may go higher. Energy turnover in GW single events may approach $\sim 10^{63}$ erg, possibly on occasion even more.

How can we further test these concepts? First of all by associating individual UHECR events, or directional groups of events, with chemical composition in both the TA Coll. and the Auger Coll. data. Second by identifying more TeV to PeV neutrinos with recent SMBH mergers. Third by detecting the order $\mu$Hz GW events and identifying the galaxies host to the stellar BH mergers and their GW events in the range up to $\sim 300$ Hz. Fourth by finally detecting the formation of the first generation of SMBHs and their mergers, surely a spectacular discovery.

7 Acknowledgement:

G. Smoot, P. Sokolsky, M. Teshima, H. de Vega, and A. Zdziarski for intense discussions on these and related topics.

References

10. I. Appenzeller et al., Astron. & Astroph. 18, 10 (1972a).
54. L.Á. Gergely et al., Class. and Qu. Grav. 27, ms.194009 (2010).
64. S.E. Hong et al., Astrophys. J. 785, id. 133 (2014).
Abstract

The Pierre Auger Observatory is collecting ultra high energy cosmic rays (above $10^{17}$ eV) to study their origin and nature. A review of selected analyses is presented, with emphasis given to the measure of the energy spectrum and mass composition, and to the contributions to the study of hadronic interactions in this extreme energy range. The new perspectives opened by the current results call for an upgrade of the Observatory, the main characteristics of which are presented together with the foreseen performances.

1 Introduction

Ultra high energy cosmic rays have been studied at the Pierre Auger Observatory for more than 10 years by recording the associated extensive air showers (EAS). The Observatory \cite{1} comprises a surface detector (SD) made up of a grid of 1600 water-Cherenkov stations covering an area of about 3000 km$^2$ and an air-fluorescence detector (FD) with a total of 24 telescopes in four sites on...
the perimeter of the array. The SD samples the particle components of extensive air showers with a duty cycle of nearly 100%, while the FD measures the longitudinal development of showers along their path in the atmosphere, with a duty cycle of $\approx 15\%$ (clear moonless nights). An infill array, with 61 water Cherenkov detectors on a denser grid of 750 m, allows an extension of the energy range of the SD-1500 m (fully efficient above $3 \times 10^{18}$ eV) down to $3 \times 10^{17}$ eV. Energies as low as $10^{17}$ eV are measured by means of 3 additional high elevation telescopes (HEAT). A sub-array of 124 radio sensors (AERA) working in the MHz range is employed to study radio emission from EAS and to identify mass-sensitive radio parameters.

The collected data provide information on the nature and origin of the primary cosmic rays and their astrophysical interpretation. From the point of view of particle physics, the measured observables allow us to set constraints on hadronic interactions and test their modelling in an energetic and kinematic region not reachable at accelerators. A selection of the most important results is presented in this paper; for most of them, we refer the reader to the latest updates summarized in 2, 3).

2 Selected results

The energy spectrum above $3 \times 10^{17}$ eV has been measured with unprecedented precision and statistics using 4 different data sets: the SD vertical and horizontal ones (within and beyond $\theta = 60^\circ$ respectively), the infill and the hybrid data. Exploiting the hybrid data set (consisting of all events detected by both the SD and FD), the energy calibration can be directly obtained from the data 4). The resulting spectrum, shown in Fig.1, shows a flattening above the ankle ($E_{\text{ankle}} = 4.8 \times 10^{18}$ eV) where it can be described by a power-law distribution with spectral index 2.6 and clearly shows a suppression of the flux above $\approx 4.2 \times 10^{19}$ eV with a significance of more than 20 $\sigma$. The dominant systematic uncertainty of the spectrum stems from the 14% overall uncertainty in the energy scale.

Different observables can be used to obtain information on the primary composition, the most direct of which is the depth of maximum development of the longitudinal shower profile, measured by the FD. $\langle X_{\text{max}} \rangle$ is related to the depth of the first interaction of the primary and to the subsequent development of the shower; for this reason, the interpretation in terms of composition is com-
Figure 1: The combined Auger energy spectrum.

Complicated by the large uncertainties in the hadronic interaction models used in the simulations. Having been corrected for the detector resolution, \( \langle X_{\text{max}} \rangle \) and its RMS can be directly compared to the predictions of air shower simulations using recent post-LHC hadronic interaction models, as shown in Fig. 2. Our measurements are clearly at variance with model predictions for pure composition; assuming no change in hadronic interactions at these energies, they point to a composition getting heavier above the ankle.

Figure 2: First two moments of the \( X_{\text{max}} \) distribution, compared to the model predictions, assuming either an all-proton or all-iron composition.

A deeper insight can be obtained by studying the shape of the \( X_{\text{max}} \) distributions: comparing them with those expected for different mass fractions.
(for diverse hadronic interaction models), the best fitting mixture of nuclei can be derived. This study shows that the data can be best reproduced with the inclusion of intermediate nuclei, the proton fraction strongly decreasing above $10^{19}$ eV; a 10-15% proton contribution seems to appear again above $\approx 2.5 \times 10^{19}$ eV. The high fraction of protons in the ankle region, together with the anisotropy limits derived from our data \(^7\) suggests that already below $10^{18.5}$ eV protons are mainly extragalactic and could point to an interpretation of the ankle as being due to the energy loss of extragalactic protons through electron-positron pair production during propagation in the cosmic microwave background (CMB) \(^6\). However, exploiting the correlation between $X_{\text{max}}$ and the number of muons produced in the EAS at energies around $10^{19}$ eV, which is sensitive to the mixture of primary masses, we showed that the composition around the ankle is actually mixed, thus disfavouring that hypothesis.

An attempt to understand the origin of the suppression was made by simultaneously fitting both the spectrum and the evolution of $X_{\text{max}}$ above $10^{18.7}$ eV. A simple astrophysical model was used assuming identical sources, homogeneously distributed in a comoving volume, injecting H, He, N and Fe nuclei. The spectrum at the source was described as a broken power law with a rigidity-dependent exponential cutoff. The best fit to the spectrum was obtained by subsequent cutoffs of the different groups of elements, with $R_{\text{cut}} = 10^{18.67}$ V and a very hard source spectrum with slope $\gamma = 0.94$, thus pointing to a flux suppression partly due to the reach to the maximum energy within the source. A second local minimum, with $\gamma = 2$ and larger maximum rigidity, similar to that expected for energy loss effects due to propagation, can fit the spectrum, but the $X_{\text{max}}$ distributions are too wide to agree with those measured. The best fit position strongly depends on the details of propagation and of the air shower development, the uncertainties of which are much larger than the statistical uncertainty of measured data.

The Pierre Auger detectors are sensitive to both the electromagnetic and the muonic components of EAS, the first one probing high energy interactions in the first part of the shower development, while muons are produced in low energy pion decays. Different analyses can thus be performed to study the characteristics of the hadronic interactions in the ultra high energy domain. Examples are the measurement of the $\sigma_{p-\text{Air}}$ and corresponding $\sigma_{p-p}$ at $\sqrt{s_{pp}} = 38.7$ and $55.5$ TeV \(^3,\ ^9\) and the finding that current air shower simu-
lations fail to describe the relationship between the longitudinal shower profile and the lateral particle densities at ground level.  

Thanks to the possibility to measure both the particle densities and their temporal distribution in each station, different methods to derive information on the muon component can be exploited. Inclined showers, muon dominated, are analysed to obtain the relative muon number $R_{\mu}$ with respect to the expectations for proton primaries and the QGSJetII-04 model. The data point to a clear underestimation of the muon number in the models (Fig.3, top-left). The mean depth of shower maximum can be converted into a prediction of the mean logarithmic muon content $<\ln R_{\mu}>$ at $\theta = 67^\circ$ for each hadronic interaction model (Fig.3, top-right). We observe a muon deficit from 30 to 80% in the simulations, depending on the models. The estimated deficit takes the mass composition of cosmic rays into account. The time distribution of muons in each SD station can be correlated to the production depth.

Figure 3: $R_{\mu}$ vs energy (top-left), $<\ln R_{\mu}>$ vs $X_{\text{max}}$ (top-right) and evolution of $<X_{\mu \text{max}}>$ with energy (bottom).
The elongation rate of $<X_{\mu_{\text{max}}}^\text{max}>$ is shown in the bottom panel of Fig.3. Besides disfavouring a constant composition, this result shows again the discrepancy with models, more severely with EPOS-LHC.

3 The AugerPrime upgrade

In the past 10 years, the Auger results have led to major breakthroughs in the study of cosmic rays. Different models have been built trying to reproduce our results\textsuperscript{8),} but the many unknowns about source distribution, composition, galactic and extragalactic magnetic fields, etc. prevent the emergence of a uniquely consistent picture. New information on the nature of the primaries is mandatory to address the problem of the origin of ultra high energy particles.

As discussed above, the origin of the flux suppression is still unknown, whether it be due to propagation effects or to exhaustion of the sources. We need mass composition information above $40 \text{ EeV}$, currently not available due to the intrinsic duty cycle of the FD. Furthermore, the direct detection of cosmogenic photons or neutrinos would be direct evidence of the GZK effect. Studies of the arrival directions of UHECRs with composition related selections will be most important to understand the reasons for the lack of small-scale anisotropy at the highest energies. The evaluation of the proton fraction above a few times $10^{19} \text{ eV}$ is the decisive ingredient for estimating the physics potential of existing and future cosmic ray, neutrino, and $\gamma$-ray detectors. From the particle physics point of view, direct measurements of the muon component of EAS will allow the study of hadronic interactions in an energy and kinematic region not explorable by terrestrial accelerators.

The AugerPrime upgrade of the Observatory has been specifically designed to improve the composition-sensitive information\textsuperscript{13).} Along the line of a hybrid design, each SD will be equipped with a top scintillator layer (Fig.4(a)). Shower particles will be sampled by two detectors (scintillators and water-Cherenkov stations) having different responses to the muonic and electromagnetic components, thus allowing us to reconstruct each of them separately. The muonic component will be derived in each station by subtracting the signal observed in the scintillator from that seen in the water Cherenkov tank. By fitting the muon lateral distribution, the muon signal at 800 m from the core $S(800)$ can be used as a composition related observable. More sophisticated methods, based on multivariate analyses or on shower universality\textsuperscript{14)}
will allow us to correlate the detector signals at different lateral distances and exploit the information of the arrival time of the EAS and of the temporal structure of the measured signals.

A preliminary demonstration of the potentials of AugerPrime can be obtained by taking two extreme and opposite assumptions fitted to the Auger flux and composition data: a maximum-rigidity (scenario 1) and a photodisintegration one (scenario 2). The muon number relative to that expected for an equal mix of p-He-CNO-Fe as primary particles, the mean $X_{\text{max}}$ and its RMS are shown in Fig.4(b-d).

![Figure 4](image)

Figure 4: (a) One of the AugerPrime upgraded surface detectors; (b-d) reconstructed relative muon number $R_\mu$, $X_{\text{max}}$ and RMS($X_{\text{max}}$) for the 2 considered scenarios (see text).

Their values are quite similar in the region below $10^{19.2}$ eV, covered by
data of the FD, but the two scenarios can be distinguished with high significance and statistics in the GZK suppression region, where the models predict significantly different extrapolations.

The currently available Auger data already underline the inefficiency of models in producing the observed number of muons. Different modifications of hadronic interaction models can be built in order to increase the number of produced muons without modifying the $X_{\text{max}}$, as shown in Fig. 5 (left).

As an example of the potential of AugerPrime in studying hadronic interactions, the mean shower-by-shower correlation of the muon density with $X_{\text{max}}$ is shown in Fig. 5 (right) for different exotic interaction model scenarios.

![Figure 5: Discrimination power of the event-by-event correlation between the muonic signal at ground and the $X_{\text{max}}$.](image)

The upgrade of the SD will also include newer electronics, with faster FADCs (120 MHz sampling compared to the current 40 MHz) and an increased dynamic range, allowing us to extend the measure to the larger signals closer to the shower core. To complement the SD upgrade, a network of underground muon detectors, each of 30 m$^2$ area, is now being deployed in the Infill area, for mass composition studies in the sub-ankle region and direct verification of the extraction of the muonic signals from the combination of top scintillators and water Cherenkov tanks. An upgrade of the FD is also foreseen: the operation mode of the FD will be changed to extend measurements into night periods with a higher light background, in order to reach a 50% increase of the on-time.

The AugerPrime upgrade is now undergoing its Engineering Array phase. Its full operation is foreseen from 2018 until 2025, when event statistics will more than double compared with the existing Auger data set, adding event-by-event mass information.
References


COSMIC-RAY STUDIES USING THE ALICE DETECTOR
AT LHC

Mario Sitta, for the ALICE Collaboration
Università del Piemonte Orientale and INFN, Alessandria, Italy

Abstract

ALICE, a general purpose experiment designed to investigate nucleus-nucleus collisions at the CERN LHC, has also been used to detect atmospheric muons produced by cosmic-ray interactions in the atmosphere. In this contribution the analysis of the multiplicity distribution of the atmospheric muons detected by ALICE between 2010 and 2013 is presented, along with the comparison with Monte Carlo simulations. Special emphasis is given to the study of high multiplicity events containing more than 100 reconstructed muons. It is shown that such high multiplicity events demand primary cosmic rays with energy above $10^{19}$ eV, and that the frequency of these events can be successfully described by assuming a heavy mass composition of primary cosmic rays in this energy range, and using the most recent interaction models to describe the development of the air shower resulting from the primary interaction.
1 Introduction

ALICE (A Large Ion Collider Experiment)\textsuperscript{1) is a general-purpose, heavy-ion detector at the CERN LHC. It focuses on the study of the properties of the Quark-Gluon Plasma (QGP) created in strongly interacting matter at extreme energy densities in high energy nucleus-nucleus collisions. This study is performed by means of various probes emerging from the collisions; in particular rare probes like charmed and beauty mesons and baryons can give valuable information about the formation and evolution of the QGP state.

Besides the Heavy-Ion physics program, ALICE developed also a cosmic-ray physics program, which exploits the excellent tracking capabilities of the Time Projection Chamber (TPC) to detect and reconstruct the muons produced by the cosmic radiation with the atmosphere.

The use of collider detectors to study the atmospheric muons was pioneered by the LEP experiments ALEPH, DELPHI and L3. Despite being small apparatus with respect to other underground cosmic-ray experiments, they had a high tracking performance and could make use of their magnetic field. All results reported by the LEP experiments were consistent with the standard hadronic interaction models, except the observation of high multiplicity muon-bundle events \textsuperscript{2), 3)}, even under the assumption of the highest measured flux and a pure Iron spectrum, the Monte Carlo models of that time failed to describe the rate of those high multiplicity events.

A development of this program is possible at LHC, where the experiments are expected to operate for many years with the possibility of collecting a very large sample of cosmic-ray data. In this context ALICE began a cosmic-ray physics program since 2009. Between 2010 and 2013 around 22.6 million events with at least 1 cosmic muon were recorded in 30.8 days of live time. Additionally in 2012 a special trigger configuration allowing the detection of high-multiplicity events during proton-proton collisions was tested.

2 Experimental setup

The ALICE experiment is located at Point 2 of the LHC tunnel, 52 m underground and with 28 m of overburden rock. This depth of rock completely absorbs the electromagnetic and hadronic components of the cosmic-ray induced air shower, and poses a threshold of 16 GeV for vertical muons\textsuperscript{4).}
ALICE is a typical collider experiment with a cylindrical symmetry around the beam pipe. A solenoid magnet houses the central barrel detectors, where different techniques are exploited to detect and reconstruct all particles coming out from the primary interaction vertex. A forward muon arm, consisting of an absorber, a large dipole magnet and planes of gaseous detectors, is located on one side outside the solenoid magnet. A detailed description of the ALICE detector is given in 1). For the cosmic-ray data taking ACORDE, TOF and SPD are used as triggering detectors while the TPC is used for reconstructing the muon tracks.

The Alice COsmic Ray DEtector (ACORDE) is an array of 60 scintillator modules located on the three top octants of the ALICE magnet, covering 10% of their surface. Each module consists of two superimposed plastic scintillators. The trigger is given by the coincidence of the signal from \( n \) different modules in a 100 ns window. For the analysis presented here the \( n = 2 \) coincidence was used, whose rate is about 1 Hz.

The Silicon Pixel Detector (SPD) forms the two innermost coaxial cylinders of the Inner Tracking System (ITS) around the beam pipe, centered at the nominal position of the interaction point and with a radius of 39 and 76 mm respectively. Composed by 10M pixels segmented into 120 modules, it can provide both, triggers and particle position. The trigger configuration used for this work requires a coincidence between signals coming from the top and bottom halves of the outermost layer; its rate is around 0.16 Hz.

The Time Of Flight (TOF) detector is a cylindrical array of 1638 Multi-gap Resistive Plate Chambers pads arranged in 18 sectors, completely surrounding the TPC. The trigger configuration used for the present analysis requires a coincidence between a signal in a sector of the upper part and a signal in either a sector in the opposite lower part forming a back-to-back alignment with respect to the central axis, or in one of the three sectors contiguous to the opposite lower sector. The rate of this trigger configuration is about 80 Hz.

The ALICE Time Projection Chamber (TPC) is the largest detector of this type, and is the main ALICE tracking device with excellent capabilities for high-track densities. The TPC has an inner radius of 80 cm, an outer radius of 280 cm and a total length of 500 cm along the beam axis. It is filled with a mixture of Ne-CO\(_2\)-N\(_2\) and it is read out by multi-wire proportional chambers at both end caps. The total area for cosmic muon detection is about 26 m\(^2\),
however after applying a minimum length cut to the reconstructed muon tracks
the effective area reduces to about 17 m².

3 Data selection and reconstruction

The data used for the cosmic-ray analysis presented here were collected between 2010 and 2013 during periods without circulating beams in LHC, both with and without magnetic field (the maximum strength of the ALICE solenoid field is 0.5 T). Cosmic-ray data were recorded with a logical OR of at least two out of the three aforementioned trigger configurations, depending on the run period. The integrated live time amounts to 30.8 days, during which ~22.6M events with at least 1 reconstructed muon in the TPC were accumulated. Most events were classified as either single or multi-muon events, with a small percentage of “interaction” events where high energy muons have interacted with the iron yoke of the magnet producing a shower of particles passing through the TPC. A multi-muon event is defined as an event with at least 5 muons; the data sample contained 7487 multi-muons events.

TPC tracks were required to consist of at least 50 clusters (out of a maximum of 159) and, in events where the magnetic field was on, to have a minimum momentum of 0.5 GeV/c, to eliminate all possible background from $e^\pm$. For multi-muon events a parallelism cut was also applied, which requires the angular difference between two tracks to be $\cos \Delta \psi > 0.99$. Finally, to match up and down tracks a maximum distance of closest approach of 3 cm in the TPC middle plane was imposed. A muon reconstructed with two TPC tracks (up and down) is called a matched muon; a track satisfying all cuts but the maximum distance is still accepted as muon candidate but flagged as single-track muon. Most single-track muons are particles crossing the TPC near the
edges where part of their trajectory may fall outside the sensitive volume.

4 Muon multiplicity distribution of atmospheric muons

The ability of the ALICE TPC in separating a high density of muons together with the measurement of other observables like momentum, charge and direction, unimaginable with a standard cosmic-ray apparatus, permits a new approach to the analysis of cosmic-ray events.

The main topic related to the cosmic-ray physics investigated by ALICE is the study of the muon-multiplicity distribution (MMD). The MMD obtained from the whole data sample and corrected for trigger efficiency is shown in fig. 1. The systematic uncertainties were estimated by varying the parameters of the track reconstruction and matching algorithms.

The data show a smooth distribution up to a muon multiplicity of $\sim 70$ and then 5 events with a multiplicity greater than 100. Events with $N_\mu > 100$ are defined High Muon Multiplicity (HMM) events.

![Muon multiplicity distribution](image)

Figure 1: Atmospheric muon multiplicity distribution of the whole sample of data (2010–2013) corresponding to 30.8 days of data taking. (Figure taken from 5)

In order to understand the MMD, simulated events equivalent to 30.8 days of live time were generated using CORSIKA 6 as event generator
and QGSJET \(^7\) for the hadronic interaction model. CORSIKA 6990 with QGSJET II-03 was used to study the MMD and HMM events, CORSIKA 7350 with QGSJET II-04 was used to further check and confirm the HMM events. Two samples, pure \(p\) (representing a light composition) and pure Fe (representing an extremely heavy compositions), were generated. The primary cosmic-ray energy was restricted in the interval \(10^{14} < E < 10^{18} \text{eV}\) following the usual power law energy spectrum \(E^{-\gamma}\), with a spectral index \(\gamma = 2.7\) below the knee \((E_k = 3 \times 10^{15} \text{eV})\) and \(\gamma = 3.0\) above. The total all-particle absolute flux was extracted from \(^8\). For each shower the core was randomly scattered at surface level over an area \(205\times205 \text{m}^2\) centered above the nominal LHC interaction point in ALICE.

The comparison between the MMD in the range \(7 < N_\mu < 70\) and the simulated distribution fitted with a power-law function is shown in fig. 2 \(^5\), where errors are shown separately (statistical and systematic) for data, and summed in quadrature for Monte Carlo. Below \(N_\mu \sim 30\) the data, as expected, are between the pure \(p\) composition (approaching it at low multiplicity) and the pure Fe composition (at higher multiplicity). Above \(\sim 30\) the low statistics does not allow to draw any firm conclusion, though the experimental points, considering their errors, are inside the region limited by the \(p\) and Fe curves.

**Figure 2:** Measured MMD compared with values and fits obtained from simulations with \(p\) and Fe primaries. (Figure taken from \(^5\))
5 High muon multiplicity events

In 30.8 live days 5 HMM events were recorded, corresponding to a rate of 1.9 × 10^{-6} Hz. The highest multiplicity cosmic muon event reconstructed in the TPC was found to contain 276 muons, corresponding to a muon areal density of 18.1 m^{-2}. To estimate the rate of these events, while limiting the fluctuations in the number of HMM simulated events, a live time equivalent of 1 year was simulated. A simplified Monte Carlo (which does not simulate the rock overburden and the detector response, but simply extrapolates all muon tracks down to the ALICE level) demonstrated that only primaries with $E > 10^{16}$ eV contribute to these events. Therefore the full simulation was restricted to primaries with $10^{16} < E < 10^{18}$ eV, still with two extreme primary compositions, pure $p$ and pure Fe. To further reduce the statistical fluctuations, four additional simulations were performed, reusing the same EAS sample and randomly varying the shower core in the 205×205 m² area. Given that the TPC acceptance is some 3000 times smaller, this ensures that the samples are statistically independent. By averaging the 5 samples the number of HMM events in 1 year is estimated while reducing the statistical fluctuations. The uncertainties are dominated by statistical errors on real data, and by systematic errors on Monte Carlo. There are two sources of systematic errors in simulation, the uncertainties in the generation parameters and the muon reconstruction algorithm. Both were carefully estimated and found to amount to ~ 20%.

In tab. 1 the results from Monte Carlo simulations are compared with data. The rate of HMM events can be well reproduced by the latest interaction models and a primary flux extrapolated from the direct measurements at 1 TeV. Pure Fe primary composition seems in closer agreement with measured rate, though the large uncertainty of the latter prevents a definite conclusion about the origin of these events. This is consistent with the fact that HMM events stem from primaries with energy > $10^{16}$ eV, where the composition is expected to be dominated by heavier elements.

6 Conclusions

In 2010–2013 the ALICE experiment collected 30.8 live days of cosmic-ray data. The MMD distribution at low and intermediate multiplicity is well reproduced by Monte Carlo simulations using CORSIKA 6990 with QGSJET II-03 model.
The measurements by ALICE presented here suggest a mixed ion primary cosmic-ray composition with an average mass increasing with energy. In the same period 5 HMM events were recorded. The observed rate is consistent with the predictions of CORSIKA 7350 with QGSJET II-04 model using a pure Fe primary composition and energies > $10^{16}$ eV. For the first time the rate of HMM events has been well reproduced using conventional hadronic interaction models and reasonable primary fluxes.

**References**

5. J. Adam *et al.* (ALICE Coll.), JCAP *01*, 032 (2016)
Review on Cosmic-Ray Radio Detection

Frank G. Schröder
Institut für Kernphysik, Karlsruhe Institute of Technology (KIT), Germany

Abstract
Extensive air showers still are our only access to the highest-energy particles in the universe, namely cosmic-ray nuclei with energies up to several 100 EeV. Studying open questions in cosmic-ray physics, like their yet unknown origin requires the reconstruction of the energy and mass of the primary particles from the air-shower measurements. Great progress has been achieved lately in the development of the radio detection technique for this purpose. There now is a consistent picture of the mechanisms behind the radio emission, which is in agreement with measurements. Several second-generation, digital antenna arrays are operating in different parts of the world not only aiming at the further development of the technique, but also contributing to cosmic-ray physics at energies above 100 PeV. Recently it has been demonstrated experimentally that radio detection can compete in precision with established techniques for air showers, like the measurement of secondary particles on ground, or fluorescence and Cherenkov light emitted by air showers. Consequently, cosmic-ray observatories can benefit from radio extensions to maximize their total measurement accuracy.
1 Introduction

Radio detection of particle cascades is one of several measurement techniques for astroparticle physics in the energy range above $10^{16}\text{ eV}$\(^{1,2}\). These cascades are primarily cosmic-ray air showers, but radio detection is also used for the search for neutrino-initiated cascades in ice and in the lunar regolith\(^3\). This proceeding focuses on the application of the radio technique on cosmic-ray physics via the measurement of air showers. In any case, there is no doubt that radio detection will work also in other media and a practical demonstration will be just a question of time. Compared to optical techniques like the detection of Cherenkov or fluorescence light, radio detection is available around the clock and not limited by light or weather conditions, except for thunderstorms directly above the radio antennas\(^4\). The first radio measurements of air showers took place already in the 1960's, though with limited accuracy due to the analog electronics available at that time\(^5\). Current antenna arrays have reached measurement accuracies similar to the established optical techniques in the energy range above $10^{17}\text{ eV}$, and in a few years the SKA can achieve an even higher precision and a lower threshold around $10^{16}\text{ eV}$\(^6\).

For air showers the dominant mechanism of radio emission is the deflection of the electrons and positrons in the geomagnetic field, which induces a transverse current in the shower front. This leads to linearly-polarized radio emission whose strength increases with the local size of the geomagnetic field, and with $\sin \alpha$, i.e., the angle between the shower axis and the geomagnetic field\(^8\). The strength of the geomagnetic emission also depends on the density of the medium: In inclined showers developing higher up in the air the emission region around the shower maximum is more extended. Thus, there is more energy emitted in radio waves than for vertical showers\(^9\). In dense media like ice, to the contrary, the showers are so compact that geomagnetic emission is negligible.

As a second mechanism the Askaryan effect contributes, i.e., radially polarized emission due to the time-variation of the electron excess in the shower front. The Askaryan effect has similar strength in all media, which makes it the dominant mechanism in dense media, where the geomagnetic emission is negligible. The strength in air depends on the shower inclination and dis-
Figure 1: CoREAS simulations of the radio emission by two air showers, one initiated by a proton, and one by an iron nucleus. The height and the color code indicate the amplitude at ground level. The steepness of the footprint depends on the distance to the shower maximum; the small asymmetry is caused by the interference of the geomagnetic and Askaryan effects.

tance to the shower axis. Typically the amplitude of the Askaryan effect is only $10^{-20}$% of the geomagnetic amplitude. The radio emission of air showers seems to be understood to at least this level of $10^{-20}$%, and current simulation programs for the radio emission, such as CoREAS, agree with measurements of the absolute radio amplitude to better than $20\%$.

For both emission mechanisms the radio emission is coherent when the wavelength is larger than the optical pathlength of the shower front. For the frequency band of $30-80$ MHz chosen by many experiments, this is the case up to a few $100\,\text{m}$ distance from the shower axis. This means that radio emission is forward beamed into a narrow cone with an opening angle of the order of $3^{\circ}$ (see figure 1). As a consequence, antenna arrays have to be either relatively dense with antenna spacings on the order of $100\,\text{m}$, or have to target inclined showers, since the illuminated area on the ground increases with the distance to the shower maximum. Recently, the Auger Engineering Radio Array (AERA) has measured that for inclined showers the radio footprint has a size similar to the particle footprint of several $\text{km}^2$ at $10^{18}\,\text{eV}$.

For a given shower direction the radio amplitude is proportional to the number of electrons, and radio detection provides a calorimetric measurement of the shower energy similarly to air-fluorescence or air-Cherenkov light detection. This makes radio detection complementary to particle detectors, which can measure only muons for inclined showers, since the electromagnetic...
component is absorbed in the air (see figure 2). Since the radio measurement of the calorimetric shower energy in combination with the number of muons depends statistically on the mass of primary particle, this combination of radio and particle detectors is a very promising technique for future large-scale arrays \(^{18}\). Additionally the atmospheric depth of the shower maximum provides complementary information on the type of the primary particle. However, before applying the radio technique on a large scale for inclined showers, some investigation is still necessary on how accurately the shower energy and the position of the shower maximum can be measured. Nonetheless, these analysis techniques are already fairly advanced for air showers with zenith angles below 60°. Current antenna arrays have achieved reconstruction precisions for the energy and the shower maximum comparable to those of the leading optical techniques, i.e., about 15% energy precision and about 20 g/cm\(^2\) for the atmospheric depth of the shower maximum.

2 Precision of shower parameters

Radio detection is sensitive to the three shower parameters most important for cosmic-ray physics: the direction of the shower axis, which is equal to the arrival direction of the primary particle; the shower energy, which is an estimator for the energy of the primary particle; the atmospheric depth of the shower maximum, \(X_{\text{max}}\), which is an estimator for the mass composition of the primary cosmic rays.

The direction can be reconstructed with a precision of better than 0.7°, as
shown by LOPES featuring nanosecond-precise time calibration \(^{19}\) and using digital radio interferometry \(^{20}\). With a very dense and accurately synchronized array, such as LOFAR, a resolution of even 0.1° might be possible \(^{21}\), though 1° resolution usually is sufficient since charged cosmic rays are deflected anyway by magnetic fields on their way to Earth.

There are two ways to reconstruct the energy from measurements of the radio amplitude: First, the time and space integral of the signal power yields the total radiation energy of the shower at radio frequencies. This radiation energy increases quadratically with the shower energy due to the coherent nature of the radio emission. AERA has demonstrated a precision and a scale uncertainty for estimating the energy of the primary particle by this method of better than 20\% \(^{15, 22}\). Second, the radio amplitude at a detector specific distance is proportional to the shower energy. While this method is not as universal, it might be more robust for measurements close to the detection threshold. The precision demonstrated by this method is similar to the first one, e.g., about 20\% for LOPES \(^{23}\), and about 15\% for Tunka-Rex (see figure 3) \(^{24}\).

The position of the shower maximum is the parameter most difficult to reconstruct. Nevertheless, there are several methods for this, since several properties of the radio signal depend on the distance to the shower maximum. LOPES \(^{26, 23}\) and Tunka-Rex \(^{24, 25}\) have shown that the slope of the lateral

Figure 3: Radio measurements of the shower energy and of \(X_{\text{max}}\) by Tunka-Rex compared to the coincident air-Cherenkov measurements by Tunka-133 \(^{24}\).
Figure 4: Top-down reconstruction method for $X_{\text{max}}$ for an example event measured by AERA. Left: the simulated radio amplitude is matched with the measurements by adjusting the amplitude scale and the core position; crosses indicate sub-threshold stations. Right: the $X_{\text{max}}$ value of the best fitting simulation is assumed as the real $X_{\text{max}}$ of the measured air showers, which for this example event is confirmed by coincident fluorescence measurements (FD)\textsuperscript{14}).

distribution can be used, and Tunka-Rex has achieved a precision of $40 \text{ g/cm}^2$, which is twice the value achieved by the leading air-fluorescence technique. Moreover the steepness of the hyperbolic radio wavefront\textsuperscript{20}) and the slope of the frequency spectrum\textsuperscript{27}) are sensitive to $X_{\text{max}}$, but the precision achievable under practical conditions is not yet clear.

The most precise, but also most computationally intensive method for $X_{\text{max}}$ is a top-down approach introduced by LOFAR\textsuperscript{28}) and meanwhile also applied by AERA\textsuperscript{14}). Several Monte Carlo simulations with different distances to the shower maximum are produced for the shower geometry of an individual event. Then, the simulated amplitude is compared to the amplitude measured at the various antenna stations to check for which $X_{\text{max}}$ the simulations fit best (see figure 4). Hence, the method implicitly exploits all $X_{\text{max}}$ sensitive characteristics of the radio footprint, not just its slope. Featuring more than 100 antennas per event LOFAR demonstrated a precision of better than $20 \text{ g/cm}^2$ for $X_{\text{max}}$ by this method\textsuperscript{29}). This precision is already similar to that of air-fluorescence measurements and might be further improved by including the information of the wavefront, the frequency spectrum, and the polarization.
3 Conclusion

Due to significant progress in the development of the radio technique and in the understanding of the emission mechanism, radio measurements can now compete in precision with optical techniques for air showers, and this around the clock. Air-shower arrays made of particle detectors can especially profit from a radio extension providing more accurate information on the shower energy and mass composition. Moreover, there are at least two further applications of the radio technique for cosmic-ray science. By focusing on inclined showers huge radio arrays covering more than 100,000 km², such as GRAND 30), could acquire significant exposure for the highest-energy extragalactic cosmic rays at several 100 EeV, and simultaneously the search for ultra-high-energy neutrinos at EeV energies. Complementary to this, radio detection can increase our knowledge on the transition from galactic to extragalactic cosmic rays, assumed in the energy range above 10^{17} \text{ eV} 31). With several 10,000 antennas, i.e., a number similar to GRAND, but inside one square kilometer, the low-frequency core of the SKA 6) will measure air showers much more precisely than possible by the optical technique today.

Acknowledgements

Thanks go to my colleagues at KIT, to the LOPES, Pierre Auger and Tunka-Rex Collaborations for fruitful discussions, and to DFG for grant Schr 1480/1-1.

References

NEUTRINOLESS DOUBLE-BETA DECAY: WHERE WE ARE AND WHERE WE ARE GOING

Andrea Giuliani

CSNSM, Univ. Paris-Sud, CNRS, Université Paris-Saclay, Orsay, France
DISAT, Università dell’Insubria, Como, Italy

Abstract

This short review introduces neutrinoless double-beta decay and discusses the implications of this phenomenon on crucial aspects of particle physics. A critical comparison of the adopted technologies and of their physics reach is performed, illustrating the possible paths towards the next-generation searches that aim at fully covering the inverted-ordering region of the neutrino masses.

1 Introduction

Neutrinoless double-beta decay ($0\nu\beta\beta$) is a hypothetical rare nuclear transition (present half-life limits are $\sim 10^{26}$ y) which plays a unique role in understanding fundamental neutrino properties and exploring lepton number violation (LNV). It consists in the transformation of an even-even nucleus into a lighter isobar containing two more protons and accompanied by the emission of two electrons.
and no other particles, with a change of the total lepton number by two units: 
\((A, Z) \rightarrow (A, Z + 2) + 2e^−\). The standard process \((2\nu\beta\beta)\), which implies also the 
emission of two electron antineutrinos, is the rarest nuclear decay and has been 
observed in eleven nuclei with half-lives in the range \(10^{18} - 10^{21}\) y. The detection 
of the neutrinoless channel would be a major discovery, and would represent 
the observation of a new phenomenon beyond the Standard Model (SM) of 
elementary particles, establishing that neutrino is a Majorana particle, rather 
than a Dirac one as all the other fermions: it would be the only spin-\(\frac{1}{2}\) particle 
to coincide with its own antimatter partner, a possibility left naturally open by 
its neutrality. In this framework, a new mechanism of mass generation, besides 
the Higgs mechanism, could be in place for neutrinos explaining naturally the 
smallness of ordinary neutrino masses, and matter-antimatter asymmetry in the 
Universe could be accounted for through CP violation in the neutrino sector.

It is important to remark however that, in a beyond-SM perspective, \(0\nu\beta\beta\) 
is much more than a neutrino physics experiment. It is a powerful, inclusive test 
of LNV, which takes the form of a creation of electrons according to the process 
\(2n \rightarrow 2p + 2e^−\), implemented in nuclear matter. LNV is as important as baryon 
number violation and naturally incorporated by beyond-SM theories. In this 
respect, the experimental search for \(0\nu\beta\beta\) must be pursued with the highest 
possible sensitivity irrespectively of the related neutrino-physics scenario, as it 
is an essential element for a deep comprehension of the elementary constituents 
of matter and of fundamental interactions.

\(0\nu\beta\beta\) can be induced by a plethora of LNV mechanisms. Among them, the 
so-called mass mechanism – consisting in the exchange of virtual light Majorana 
neutrinos – occupies a special place, since it is mediated by the light massive 
neutrinos which undergo flavour oscillations. In this mechanism, the rate of 
the process is proportional – within an uncertainty due to the computation of 
the nuclear matrix elements – to the square of the effective Majorana neutrino 
mass \(M_{\beta\beta}\), related to the absolute neutrino mass scale and to the mass ordering. 

Present limits on \(M_{\beta\beta}\) from \(0\nu\beta\beta\) are in the range 60-600 meV (the experiment 
KamLAND-Zen \(^1\)) is leading the field), assuming that the axial charge \(g_A\) (the 
\(0\nu\beta\beta\) rate is proportional to \(g_A^4\)) is not quenched and equal to the free nucleon 
value of \(\sim 1.25\) (the most common approach in the literature). The possible 
quenching of \(g_A\) is an important open issue, since it could reduce even by a 
factor \(\sim 4\) the sensitivity to \(M_{\beta\beta}\). However, this quenching could have no
impact on other LNV mechanisms, and in any case it demands for even more powerful technologies and new experimental ideas.

Figure 1: The positions of the expected signals for the nine most favourable $0\nu\beta\beta$ isotopes are compared with background energy markers related to the maximum $\gamma$ energies of the $^{238}$U and $^{232}$Th chains and the maximum $\beta$ energy of the $^{238}$U chain.

2 Experimental concepts

In order to observe $0\nu\beta\beta$, experimentalists aim at the detection of the two emitted electrons, which share the total transition energy (the so-called Q-value of the process). The signature of $0\nu\beta\beta$ is therefore a peak at the Q-value in the sum-energy spectrum of the two electrons. The features of the $0\nu\beta\beta$ signal and the long expected life-times suggest immediately the desired properties of a powerful $0\nu\beta\beta$ experiment: large sources and high detection efficiency; high energy resolution; extremely low background (underground operation and high material radiopurity are basic features); viable isotopic enrichment in terms of price and throughput, as the natural isotopic abundance of appealing candidates is generally below 10% (with the exception of $^{130}$Te).
In current and future experiments, sources must contain at least tens or hundreds of kg of the isotope of interest. The next frontier is the tonne scale. This constraint makes particularly appealing the so-called calorimetric technique, in which the source is embedded in the detector. Zero background at large exposure scale is a big advantage, as it allows the experimentalists to exploit at best the costly enriched material and detector technology. The Q-value is a crucial criterion, as it affects both the phase space (which approximately scales as $\sim Q^5$) and the background. As a consequence, at the moment only nine isotopes – all with high Q-values – are or may become experimentally relevant. It is instructive to compare their Q-values with two important energy markers in terms of background sources: the 2615 keV marker, a $^{208}$Tl line in the $^{232}$Th chain, is the end-point of the natural $\gamma$ radioactivity; the 3270 keV marker is the Q-value of the $\beta$ decay of $^{214}$Bi, belonging to the radon progeny in the $^{238}$U chain. A graphical representation is provided in Fig. 1.

A first group of three isotopes ($^{76}$Ge, $^{130}$Te and $^{136}$Xe) have a Q-value above 2 MeV but below both markers, and therefore have to cope with the $\gamma$ background and with the radon-induced one. However, enrichment is viable and superb detection technologies can be employed for these nuclei: germanium diodes (GERDA 2), xenon liquid and gaseous detectors (EXO 3), NEXT 4), large liquid scintillator volumes incorporating the candidate nuclei (KamLAND-Zen 1), SNO+ 5), and TeO$_2$ bolometers (CUORE 6)). Thus, it is not surprising that the currently most sensitive experiments study these three nuclei (KamLAND-Zen 1), EXO-200 3), GERDA-I 2) and CUORE-0 7)), as shown in Fig. 2, which illustrates synthetically the experimental status. Conversely, the three candidates $^{48}$Ca, $^{96}$Zr and $^{150}$Nd are in the best position to carry out a background-free experiment, but they are ruled out in practice by having a very low isotopic abundance and, in addition, large-scale enrichment is impossible or prohibitively expensive. The remaining group of three candidates ($^{82}$Se, $^{100}$Mo and $^{116}$Cd) has a $0\nu\beta\beta$ signal out of the reach of the bulk of the $\gamma$ environmental background and furthermore can be effectively enriched. These three nuclides can be efficiently studied with bolometers. When used in a hybrid version adding a scintillation light readout, like in LUCIFER 8), LUMINEU 9, 10) and AMoRE 11), an almost background-free technology is available.
Figure 2: Sensitivities of the current $0\nu\beta\beta$ experiments to $M_{\beta\beta}$ (the span is due to different nuclear models). The reaches of present searches and of next-generation ones are indicatively shown. No $g_A$ quenching is assumed.

3 Critical comparison of the current technologies

The $0\nu\beta\beta$ community (which counts about 850 physicists) is starting to discuss about possible next-generation experiments capable to attack and in prospect fully cover the inverted ordering region, which can only marginally probed by current searches. This goal requires an isotope sensitive mass from hundreds of kg to several tonnes, depending on the adopted technology. Current experiments and ongoing R&D activities suggest that three main routes can allow achieving this objective, depending on the source configuration: we can distinguish searches adopting Fluid-embedded, Crystal-embedded and Externals Source approaches (FS, CS, ES respectively).

3.1 Fluid-embedded Source (FS)

This approach is adopted by the experiments KamLAND-Zen $^1$, EXO-200 $^3$, NEXT $^4$ and SNO+ $^5$. In these searches, either the isotope constitutes by itself the sensitive medium in form of gas or liquid (as $^{136}$Xe in EXO-200 and...
NEXT – in prospect nEXO and BEXT), or it is dissolved, typically at the level of few %, in a large scintillator volume using pre-existing infrastructures (as $^{136}$Xe in KamLAND-Zen and $^{130}$Te in SNO+). In both cases, the approach is calorimetric, enabling large efficiencies. The strongest point of this method is that it allows investigating large isotope masses (the 1 tonne scale is well within its reach) and accumulating large statistics. In addition, it is scalable (either by increasing the concentration in case of isotope solution – up to a maximum value dictated by physical limits or light collection efficiency – or by building larger and/or multiples structures in case of Xe as a sensitive medium). High radiopurity is achievable, as fluids are in general easier to purify than the solids adopted in the CS and ES options. However, the radiopurity of the containment structures remains an issue (volume fiducialization is in general necessary, reducing the efficiency in the isotope use), together with radon emanation. Unexpected contamination is always possible, requiring additional purification efforts (see the case of $^{110}$Ag in KamLAND-Zen). When the isotope is extremely diluted (as in the case of SNO+ where the isotope/scintillator ratio is a few $10^{-5}$), $^8$B solar neutrinos represent an ultimate irreducible background source. Isotopes used in this approach are the easiest to enrich ($^{130}$Te has in particular a record natural isotopic abundance of 34% and $^{136}$Xe – with its 5-10$/BBOM/g, has an enrichment cost $\sim$ 10 times lower than the average, even though it is to remark that 1 tonne of enriched Xe corresponds to 1/4 of the world annual Xe production) but have Q-values (2458 keV for $^{136}$Xe and 2530 keV for $^{130}$Te) below the limit of the natural $\gamma$ radioactivity (2615 keV of $^{208}$Tl). A drawback of $^{136}$Xe in particular is the proximity to the Q-value of a line of $^{214}$Bi (2448 keV), an isotope belonging to the radon progeny. This situation is aggravated by the general low energy resolution of the FS approach (which makes $2\nu\beta\beta$ a considerable background source): 250 keV FWHM for KamLAND-Zen and 90 keV FWHM for EXO-200. An exception in this scenario is NEXT, which has demonstrated energy resolution around 20 keV FWHM. In terms of background identification, the isotope solution approach can count on event location (and consequent fiducialization) and delayed coincidence. More powerful means are available for the liquid TPC of EXO-200 (multi-site versus single site-events) and especially for NEXT, which – with its high-pressure gaseous TPC – can use event topology as $0\nu\beta\beta$ signature. Finally, we remark that future evolutions of EXO-200 and NEXT could use atomic spectroscopy to identify the final nu-
clear state by detecting the presence of a $^{130}$Ba atom at the event location. In the frame of the FS approach, extensions to multi-tonne scale experiments are already under discussion (nEXO and KamLAND2-Zen).

3.2 Crystal-embedded Source (CS)

This approach is adopted by the experiments GERDA $^2$, MAJORANA $^1$, CUORE $^7$, AMoRE $^11$ and the demonstrators LUCIFER $^8$ and LUMINEU $^9, 10$ in the framework of CUPID $^13$, the proposed follow-up to the CUORE experiment. In these searches, the isotope is incorporated in high-purity single crystals with a very high mass fraction. As in the FS case, the approach is calorimetric. Here however the efficiency is much higher (80-90%) as no fiducialization is required for background control. In addition, energy resolution is much better than in the FS case (3 keV FWHM for the Ge diodes of GERDA, 5 keV for the TeO$_2$ bolometers of CUORE and in the range of 5-10 keV for the scintillating bolometers of LUCIFER, LUMINEU and AMoRE).

Since the $0
\nu\beta\beta$ signal is a peak, high energy resolution is of course welcome. In addition, $2
\nu\beta\beta$ is not an issue, with the exception of random coincidences in the $^{100}$Mo case (LUMINEU). Scalability is possible, even though achieving the tonne scale and beyond is not as easy as in the FS case. It can be accomplished however thanks to the intrinsic modularity of the CS approach. Crystals have masses of the order of a few kg in the Ge diode case and of 0.3 -1 kg in the bolometric case. Large sensitive masses are achievable by multiplication of the crystal number. Present infrastructures (GERDA and CUORE cryostats) allow housing a few hundreds of kg of isotope mass. The enrichment-purification-crystallization chain, especially in a large-scale context, represents however an important effort in these technologies. Low irrecoverable isotope losses in the crystal production processes are crucial, and have been demonstrated only for $^{76}$Ge, $^{100}$Mo and $^{116}$Cd up to now. The rather low Q-value of $^{76}$Ge (2039 keV) reduces the phase space for the $0\nu\beta\beta$ transition and makes background control harder, as several characteristic-energy photons of the natural $\gamma$ radioactivity contributes to the background. In the bolometric case, only the 2615 keV line of $^{208}$Tl is relevant for $^{130}$Te (CUORE case) and no major lines contributes in the scintillating bolometer case, where the involved isotopes have a Q-value above 2615 keV (3034 keV for $^{100}$Mo – LUMINEU and AMoRE – and 2998 keV for $^{82}$Se – LUCIFER). Specific technologies for background control are
available. In case of Ge diodes, pulse shape discrimination can reject multisite events generated by external $\gamma$'s. By exploiting this method and with the help of an active liquid-argon veto, outstanding results were recently obtained in terms of specific background by GERDA-II \cite{14}: a record value of the order of $7 \times 10^{-4}$ counts/(keV kg y). In case of bolometers, a simultaneous measurement of scintillation and heat can reject very efficiently the dominant $\alpha$ background in detectors based on ZnSe (LUCIFER), ZnMoO$_4$/Li$_2$MoO$_4$ (LUMINEU) and CaMoO$_4$ (AMoRE) crystals. This discrimination can be performed with higher difficulty in the non-scintillating TeO$_2$ crystals, using Cherenkov light. All the isotopes involved in the SC approach can be enriched by centrifugation, with costs which range from $\sim 20$ $\$/g of $^{130}$Te to 80-120 $\$/g in the other cases. In the frame of the CS approach, extensions to scales of several hundreds of kg or $\sim 1$ tonne are under discussion (GERDA upgrade in Gran Sasso, CUPID, and joint GERDA-MAJORANA experiment).

### 3.3 External Source (ES)

The only experiment beyond the R&D phase which plans to use the ES approach is SuperNEMO \cite{15}, which will be preceded by a small-scale demonstrator under commissioning. The enriched source, consisting of a thin foil (thickness $\sim 50$ mg/cm$^2$) containing 7 kg of $^{82}$Se in the demonstrator, is separated from the detecting section, which comprises a gas tracker and a plastic-scintillator calorimeter. The strong points of this technique are the compatibility with all the isotopes and the full topological reconstruction of the events, providing excellent background rejection. In addition, the sensitivity to the Majoron mode is unrivaled. Drawbacks are the low efficiency (30\%) and energy resolution (120 keV FWHM). Scalability is possible by replication of $\sim 5$ kg modules, but with high cost and space occupation. The low efficiency could be partially compensated by the use of the $^{150}$Nd, which has the highest phase space and potentially zero background because of the very high Q-value (3371 keV). Recently, the enrichment of this isotope by high-temperature centrifuges was demonstrated, but the cost remains very high.

### 4 Conclusions

We have shown in this review that the experimental search for $0\nu\beta\beta$ is a rich and living field. A healthy competition between different technologies is pushing
forward the reach of the current and future experiments. On a two-three year time scale, several searches will start to explore the inverted-ordering region of the neutrino mass pattern. On a longer time scale, we could have two – three experiments capable of fully covering this region and to approach the direct-ordering one.

References

UNDERSTANDING ICECUBE’S ASTROPHYSICAL NEUTRINO OBSERVATIONS

Ignacio Taboada, Georgia Institute of Technology
on behalf of the IceCube collaboration

Abstract

IceCube, a neutrino telescope, has observed an astrophysical neutrino flux in the 30 TeV - 5 PeV energy range. This flux has been detected with two channels and using several analyses techniques. It is expected, though not guaranteed, that the sources of these neutrinos would also be sources of cosmic rays. Electromagnetic counterparts have not been identified yet. Many usually assumed sources of neutrinos, such as blazars, GRBs, nearby starburst galaxies, etc. have been ruled out or severely constrained. The observed neutrino directions are consistent with isotropy, probably indicating an extragalactic origin. The flavor flux ratio of neutrinos is consistent with expectations of standard oscillations. IceCube has used it’s own data to constrain the possible sources. The consequences of missing gamma-ray counterparts is discussed.
1 IceCube

IceCube is a very-high-energy neutrino telescope in operation at the South Pole. Neutrinos interacting in or near the detector produce charged particles that radiate Cherenkov light which is detected by optical sensors (DOMs) in the highly transparent Antarctic ice. Using the time of the signals in the DOMs, the direction of the particles can be measured and the amount of light detected provides for energy. Over 5000 DOMs instrument a volume of $\sim 1 \text{ km}^3$. The gigantic volume is needed to compensate for the tiny neutrino-matter cross-section.

The two main methods of detection by IceCube are the muon channel, in which charged current interactions of $\nu_\mu$ result in a muon that travels several kilometers through rock or ice at the energies relevant to IceCube. With muons, directional uncertainty is 0.4 - 1 degree, but correlation of the muon inside of the detector is poor with the parent neutrino energy.

Through various channels, all neutrino flavors produce cascades or showers. In this case, all the energy is deposited in a volume that is very small compared to the detector dimensions. Lacking lever arm, angular resolution is 10-30 degrees, but energy is usually well correlated with the parent neutrino.

Most of the data collected by IceCube, 3 kHz, is due to atmospheric $\nu_\mu$ are observed at 3 mHz and astrophysical neutrinos are in the scale of $\sim 10$ per year depending on the detection method. Astrophysical neutrinos have a spectrum that is significantly harder than that of atmospheric muons or neutrinos, allowing for their identification.

2 The discovery of astrophysical very high energy neutrinos

The first evidence for astrophysical neutrinos was due to the serendipitous observation of two cascade events of PeV energy\footnote{\cite{1}} in a search for ultra-high-energy neutrinos due to the Berezinsky-Zatsepin effect, i.e cosmic ray protons of $\sim 5 \times 10^{19} \text{ eV}$ interacting with cosmic microwave photons.

Since that study was not optimized for PeV neutrino energies a follow up study was conducted. Two very simple selection criteria were used. First, events were required to be very energetic. The sum of light collected by the detector in the event was required to be at least 6000 photo-electrons. This roughly equivalent to 30 TeV. Second, the initial light deposition was required
to be inside a fiducial volume of \( \sim 40\% \) of the instrumented detector volume. The outer part of the detector was used as a veto. Very energetic through-going and down-going muons, the most relevant background, will almost always produce light first in the veto region than in the fiducial volume. But energetic neutrinos interacting in the fiducial volume will not have early light in the veto region. This method, called HESE, was initially applied to two years of data \(^2\). It has subsequently been expanded to include two additional years of data \(^3, 4\).

The best fit spectrum, with the 4 year study, is 
\[
E^2 dN/dE = 2.2 \times 10^{-8} (E/100 \text{ TeV})^{-2.58} \text{ GeV.cm}^{-2}.\text{s}^{-1}.\text{sr}^{-1}.
\]
It is interesting that this is a level similar to the Waxman-Bahcall bound \(^5\). Note that no conclusion can be unequivocally be drawn from this, as the neutrinos observed are in a different energy range that what is relevant for the bound. The data set is cascade rich, so directional information is degraded. However the observations are consistent with isotropy over \( 4\pi \) sr. In this data set there is no evidence for point sources and the events do not show correlation with the galactic plane.

3 Observation with through-going muons

Through-going muons are the traditional neutrino astronomy channel. Because Earth is used to filter background, the observation is limited to \( 2\pi \) sr. The flux has also been observed in an energy range 200 TeV to 4 PeV \(^6, 4\) using 6 years of data. The spectrum is harder than as measured with HESE, with a best fit spectral index of -2.1. The inconsistency is strong enough that it may be real and a consequence of the difference in energy range between the methods. Other methods, not explained in these proceedings, used by IceCube in the 30 TeV - 1 PeV range are consistent with index measured by HESE. The hardening of the spectrum may be an indication of, at least two populations.

An astrophysical source that results in one neutrino of >100 TeV may be result in many neutrinos above 1 TeV. So IceCube has used it’s own large data set of atmospheric neutrinos to search for point sources. They have not been found. Limits are stringent enough that the brightest neutrino source that is consistent with the astrophysical flux can be no brighter than \( \sim 1\% \) of the flux \(^3\). Point source limits have also been placed on correlations with candidate sources: Blazars reported by Fermi represent no more than 17% of the flux. Nearby starburst galaxies no more than 8%. The galactic plane, including
diffuse emission and sources can contribute no more than 14%.

Adding time allows for the study of GRBs. Bursts reported by satellites contribute no more than 1% of the astrophysical flux during the prompt phase \(^7\). And even if the correlation time is expanded to \(\pm 20\) hours, GRBs can represent no more than 12% of the astrophysical flux.

---

4 Consistency with standard neutrino oscillations

IceCube does not provide with perfect flavor identification. However a joint study of all detection methods, cascades and muons, has been used to constrain the flavor flux ratio of neutrinos \(^8\). Cosmic ray sources are expected to produce neutrinos in an astrophysical beam dump resulting in a flavor flux ratio of 1:2:0 for \(\nu_e,\nu_\mu,\nu_\tau\). In the beam dump the electron neutrino and one of muon neutrinos are due to muon decay and the other muon neutrino due to pion decay. Standard vacuum oscillations changes this ratio, at Earth to 1:1:1. Assuming that sources are such that muons loose energy before they decay (e.g. synchrotron losses), changes the source ratio to 0:1:0 and the Earth ratio to 0.2:0.4:0.4 \(^9\). Another possible source scenario is a cosmic ray source that traps all charged particles and only neutrons can escape. Neutrinos are
due to neutron decay in flight. In this case the source ratio is 1:0:0.

The joint study is consistent with standard oscillations for the case of source ratios 1:2:0 or 0:1:0 - but it is inconsistent at the 3 sigma level with the source ratio of 1:0:0. This joint study has also been used to provide a best fit spectral model with a resulting index of -2.5. This, as indicated before is in tension with the muon channel results. Figure 1 shows the joint spectral fit result. Figure 2 shows the allowed flavor ratios at Earth that are consistent with the joint study.

![Neutrino flavor triangle](image)

Figure 2: Neutrino flavor triangle. Each side of the triangle represents the fraction of the total flux for each flavor. Each marker corresponds to a source hypothesis after standard vacuum oscillations have been taken into account. The best fit value is shown in the white × sign - note however that the allowed contours are very wide.

5 Follow up with gamma-rays

Cosmic ray sources are expected to be sources of gamma rays too. Moreover given a neutrino spectrum, the gamma ray spectrum can be calculated in detail, either assuming p-p or p-γ interactions at the source. Follow up studies by VERITAS and HAWC have failed to find gamma rays in directional...
or temporal correlation with the neutrinos. The lack of observations may be interpreted several ways. It is possible that the sources are opaque to cosmic rays and gamma rays and only neutrinos escape. It is also possible that the sources are a distance that is large enough that the extragalactic background light attenuation of the gamma rays prevents their observation. Assuming that there is no attenuation, HAWC and VERITAS provide limits on the largest fraction a given source can be responsible for the total astrophysical neutrino source. The level of these limits is approximately 0.1% for each experiment.

![Figure 3: Limits on gamma ray source set by HAWC in the direction of likely astrophysical neutrinos reported by IceCube. The set of 28 events corresponds to events of 200 TeV found in 6 years of IceCube data described in section 3. The limits assume an index of -2.1, as consistent with IceCube’s study. The highest energy event, of more than 4 PeV, is indicated in blue. As a reference, the gamma ray spectrum matching a neutrino source of 1% ($f_s = 0.01$) of the total astrophysical flux is shown. Extragalactic background light attenuation for two redshifts are also shown. Also for reference, the Crab spectrum is shown.](image)

It has been observed that there are too many neutrinos when comparing the isotropic gamma ray background as measured by Fermi 12, 13). This has been used to rule out star burst galaxies at all redshifts - not only for nearby
ones as done with IceCube data. While Fermi gamma ray observations, in the
GeV scale are far less sensitive to extragalactic background light attenuation
than the TeV observations by HAWC and VERITAS, it requires to extrapolate
the spectrum from \( \sim 30 \text{ TeV} \) to GeV scale.

Multiple correlations between individual and set of neutrinos have been
claimed with several astrophysical objects, e.g. from PKS B1424-418 \(^{14}\). All
these correlations however have p-values that fall quite short of the 5 sigma
level that is usually required.

6 Conclusion and Outlook

The observation of very-high-energy astrophysical neutrinos by IceCube is an
history changing event for particle astrophysics. It shows the capabilities of
neutrinos to be used for astrophysical studies at the highest energies. The
identification of sources is a critical issue that will keep the community busy
for years to come.

A simple explanation for the observations is a large number of faint neu-
trino sources. On average each source produces a minuscule flux. But Poisson
fluctuations are enough for individual sources to result into a single neutrino
being observed by IceCube. This would explain isotropy, the lack of point
sources, no correlations with very-high-energy gamma rays, etc. It is however
far from proven that this explanation is correct. Evidence for 2 populations
of sources as explained in the text, may complicate identifying the sources
significantly.

The construction of KM3Net and a proposed upgrade of IceCube, Gen2,
will for certain improve the precision with which we can measure the astro-
physical spectrum and will provide with improved statistics on the detection
rate of neutrinos. In some cases, they additional statistics may prove enough to
identify a point source in neutrinos and allow the electromagnetic astronomers
to carefully study this area of the sky.

7 Acknowledgements

The author is supported by NSF grant PHY-1505230.


References


2. M.G. Aarsten et al. (IceCube collaboration), Science 342, 6161 (2013)


11. I. Taboada for the HAWC collaboration. This conference.


REMEmBERING GUIDO ALTARELLI

Mario Greco
Department of Mathematics and Physics and INFN, University of Roma Tre

Abstract

In memory of Guido Altarelli I present my personal recollections of the early times and his major role played in the development of QCD.

I have been asked to organize this session in memory of Guido Altarelli, who participated to the last edition of this astro-particle physics conference in 2014, and I have done it with great pleasure and sadness at the same time, because of my more than fifty-years-long friendship with Guido. We first met when students at the University of Rome in the early 1960s, and in the following many more years we got quite close friends together with our families. Often we discussed physical problems of common interest but we also shared together some nice vacation times. Our offices at the Physics Department of Roma Tre University have been just a few meters away.

In this talk I will recall the early times of his scientific career and his major role played in the development of QCD, in particular in those fields I
have been also involved, leaving to the following speakers to describe other aspects of his research activity and also the role he played in the development of phenomenology of particle physics, particularly at CERN. Guido himself has written his own recollections on the early days of QCD in Rome in the 70s and early 80s, in the occasion of my 70th anniversary \(^1\). A similar description of the contributions of Guido in the evolution of QCD has been presented by Keith Ellis last March at the 2016 La Thuile conference \(^2\).

Together with Guido Altarelli, I recall many other friends and future colleagues at the university times, as Franco Buccella, Giorgio Capon, Sergio Doplicher, Giovanni Gallavotti, Adalberto Giazotto, Luciano Maiani, Francesco Melchiorri, Piergiorgio Picozza, and Piero Spillantini. For sure many of them are well known to this audience of astro-particle physicists.

Many people know that Guido worked for his diploma thesis together with Franco Buccella, whom he knew since the primary schools, on the problem of single bremsstrahlung emission in electron-positron annihilation \(^3\), suggested to them by Raoul Gatto. Indeed in those years, after Bruno Touschek’s seminal idea of colliding beams, the construction at Frascati of AdA (in 1961, with an energy of 250 MeV) and later of ADONE (with an energy of 3 GeV) led much theoretical efforts in the roman area to investigate the physics of electron-positron annihilation. Actually Guido did his first attempt for his diploma thesis with Bruno Touschek, who was at the time very busy with the luminosity problems of AdA which, as it is well known, was finally moved to Orsay to improve the injection. In a cold day of autumn Guido finally succeeded in getting an appointment with him in his office, but because of a malfunctioning heater, Touschek’s jacket started burning, he got very nervous and the talk ended poorly.

After his degree, Guido went to Florence joining the new Gattos theoretical group, with other roman colleagues and friends, as Franco Buccella, Giovanni Gallavotti, Luciano Maiani and Giuliano Preparata, forming the group of the so called gattini (meaning kittens). They had been working in those years mainly on unitary symmetries and super-convergence relations in strong interactions. On the other hand, with the advent of ADONE, a sizeable theory
group developed at Frascati, whose members were Gianni De Franceschi, Paolo Di Vecchia, Etin Etim, Giulia Pancheri, Giancarlo Rossi and myself, under the direction of Bruno Touschek, who was quite convinced that a proper quantitative analysis of the experimental results from an $e^+e^-$ colliding beam needed a precise computation of QED radiative corrections re-summed to all orders. Some of the basic ideas and techniques developed there will be of great help in later years in the understanding and the description of resummation effects in QCD.

In 1968-70 Guido spent two academic years at NYU and Rockefeller University working on S-Matrix duality 4) and light-cone expansions 5) with R. Brandt and G. Preparata. The general physics context at that time was the approximate scaling properties shown by the SLAC deep inelastic scattering data, which had motivated the Feynman’s parton model 6) as well as the study of the properties of the commutator of the electromagnetic currents on the light cone 7). On the other hand Leon Ledermans group had presented their results 8) on the production of muon pairs in hadronic collisions and immediately later S. Drell and T. Yan 9) had proposed the quark-antiquark annihilation mechanism. At Frascati in the meantime ADONE was producing the first tests of QED in the GeV region and the early results on the large multi-hadron production 10).

Once back to Rome Guido started working with Nicola Cabibbo and Luciano Maiani on various subjects: sigma terms, chiral symmetry and scale invariance 11), deep inelastic phenomena 12), and after the discovery of the asymptotic freedom in gauge theories 13, 14), on the octet enhancement of non-leptonic weak interactions 15).

The discovery of the $J/\psi$ 16, 17) in November 1974 was a shock in Rome as elsewhere. At Frascati the ADONE accelerator team was able to raise the energy up to 3.1 GeV to produce the $J/\psi$ a couple of days after SLAC, and publish their results in the same PRL issue 18). Unfortunately, a private communication with the experimenters indicating an apparent forward-backward asymmetry observed in the muon pairs produced on resonance, gave rise to a paper from the Rome group 19), who wrongly concluded that the
particle discovered was the $Z_0$ boson. On the other hand in that period of time I had the terrific chance to be invited by Sid Drell to give a seminar on my previous works on duality, and arrived at SLAC just the day after the discovery. Those were days of great excitement, endless talks, and after some detailed discussions on the data with Burt Richter and Roy Schwitters a few days later I left to Mexico City which was my final destination. The careful study of the $J/\psi$ data and the subsequent discovery of the $\psi$, which I learnt from a local newspaper, immediately led Cesareo Dominguez and myself, using duality ideas, to interpret the new particles as a new series of bound states of charm quark-antiquark pairs and also predict the appropriate increase of the famous ratio $R$ after the open charm threshold $^{20}$, as independently suggested at the same time by Appelquist and Politzer $^{21}$ and De Rujula and Glashow $^{22}$.

Let's discuss now the great progress in the development of QCD realized with Altarelli Parisi (AP) equations $^{23}$. Before that, the applications of QCD to physical processes was quite complicated at least for two reasons. By using Guido's words, "QCD is a theory of quarks and gluons while only hadrons are observable. Moreover perturbation theory can only be applied in those particular domains of the strong interaction where approximate freedom, which is only asymptotic, can be reached" $^{23}$. In addition, "in spite of the relative simplicity of the final results, their derivation, although theoretically rigorous, is somewhat abstract and formal, being formulated in the language of renormalisation group equations for the coefficient functions of the local operators which appear in the light cone expansion for the product of two currents".

In 1976 Guido Altarelli and Giorgio Parisi were both on sabbatical leave in Paris, Guido at the Ecole Normale Superieure and Giorgio at the Institut des Hautes Etudes Scientifiques at Bures-sur-Yvette. Giorgio had presented at the Moriond conference a paper $^{24}$ containing an early form of the AP equation, namely a simple evolution equation for the electron and photon distributions after a bremsstrahlung emission, suggesting a similar treatment in QCD. The common paper appeared in 1977 $^{23}$ and the main virtue of their approach was to formulate the evolution of the parton densities as a branching process with probabilities determined (at leading order) by the splitting functions (proportio-
tional to the running coupling). A particular emphasis was devoted to prove that the splitting functions are a property of the theory and do not depend on the process, in particular the evolution does not apply only to deep inelastic scattering.

Within the general framework given by the AP equations, the Rome group contributed very much to the theory of Drell-Yan processes. In particular an important progress was made in 78-79 with the calculation of the next to the leading (NLO) corrections to Drell-Yan processes by Guido with Keith Ellis and Guido Martinelli 25). This was one of the first calculations of NLO corrections in QCD. They started by defining the quark parton densities beyond leading order in a precise way (for quarks they adopted the structure function $F_2$ as the defining quantity, gluons only enter at NLO in Drell-Yan processes). Then the calculation of NLO diagrams for both deep inelastic scattering and the Drell-Yan process allowed to derive the corrective terms for the Drell-Yan cross-section, as function of $Q^2$. The resulting corrections turned out to be surprisingly large. The ratio of corrected to uncorrected (Born) cross-sections was found to be rather constant in $Q^2$ and in rapidity. They denoted it as the K-factor. This large correction was clearly casting doubts on the convergence of perturbation theory.

Actually the origin of the main part of this correction could be traced back to effects that can be re-summed to all orders. Indeed in a series of works, the tools developed over the years for QED resummations at Frascati, were applied by us to the resummation of soft gluons in different QCD processes, in particular in collaboration with Giuseppe Curci and Yogi Srivastava 26, 27). This concerns the so called large $\pi^2$ terms and also the resummations near the phase space boundaries, both in deep inelastic scattering near $x = 1$ and in the Drell-Yan processes near $\tau = Q^2/s = 1$ 28). Giorgio Parisi also studied the $\pi^2$ problem at near the same time 29). In Figure 1 it is shown how we looked like in those years.

Another important theoretical problem for Drell-Yan processes that was attacked in those years is the evaluation of the transverse momentum ($p_T$) distribution of the produced virtual boson (a $\gamma$ or a $W^\pm$ or a $Z_0$). After the
LO perturbative result, valid for $p_T \approx Q$, completed in 78 by Guido with Parisi and Petronzio \cite{30}, the NLO perturbative calculation was obtained in 81-83 by K. Ellis, Martinelli and Petronzio \cite{31}. That followed the study of the resummation of the Sudakov double logarithms, paper by Dokshitser, Dyakonov and Troyan \cite{32} who however obtained an incorrect result. Then the correct re-summed answer was given in 79 by Curci, Srivastava and myself \cite{27} taking into account the conservation of the transverse momentum in the multi-gluon emission in the initial state. As soon as the data on the W and Z production from UA1 and UA2 at CERN proton-antiproton collider were available, an adequate theoretical prediction, adding the result of the Sudakov resummation to the complete one loop calculation was finally obtained in a paper by Guido, K. Ellis, G. Martinelli and myself \cite{33}. Using Guidos words \cite{1} this is an important paper, because it essentially contained all the crucial ingredients that describe the physics of this phenomenon. In the subsequent years the accuracy was much improved with the computation of sub-leading effects and with several different refinements, but the essential points were all present in our paper and the accuracy of our treatment was adequate for the quality of the first data. The same techniques are at present applied to the calculation of the $p_T$ distribution of the Higgs boson produced by gluon fusion” (see, for example, reference \cite{34}).
In Figure 2 all four of us are shown, participating into a conference in Crete in 1980, with thanks to Guido Martinelli.

I will stop here, not mentioning further research activity in QCD as, for example, in the small-x physics domain, where both of us have been involved in recent years [35, 36], leaving the floor to next speakers. As I already mentioned in the beginning, Guido was so kind to give a talk at La Thuile conference in 2011 on the occasion of my 70th birthday. Fortunately that talk had been recorded and, to conclude, I will show now a few pieces of his presentation in order to leave with us a last image of his impressive personality and kindness.
References

10. For a review, see for example V. Silvestrini, Procs. of the 16th International Conference on High-Energy Physics, Batavia, Illinois, 6-13 Sep. 1972.

NEUTRINO MASSES AND MIXING ANGLES:
A TRIBUTE TO GUIDO ALTARELLI

Ferruccio Feruglio
Dipartimento di Fisica e Astronomia G. Galilei, Università di Padova, Italy
Istituto Nazionale di Fisica Nucleare, Sezione di Padova, Italy

Abstract

I present a personal recollection of Guido Altarelli, focused on his contribution to the problem of neutrino masses and mixing angles. I recap the main ideas in model building, illustrating how the subject evolved from the study of continuous-abelian to discrete-nonabelian flavour symmetries, emphasising the point of view of Guido on the subject. I conclude by commenting the present status of the field.

1 Guido vision on neutrinos

I first met Guido in 1989, on the eve of LEP start, and this marked the beginning of an intense collaboration lasted 25 years. In 1989 I was a graduate student at the University of Geneva, while Guido was a leader of the CERN
theory group, fully committed to the LEP program and to its demanding activity. I was soon captured by Guido real passion for physics, for his deep perspective in many areas of our field and by his stunning ability in synthesising deep concepts with few well-chosen words. I remember the enthusiasm and energy with which he dragged me into the world of neutrinos when neutrino oscillations were discovered by Superkamiokande in 1998. He was eager to participate in the fascinating adventure launched by the new data and I had the fortune and the privilege of being at his side.

As for other aspects of particle physics, Guido had his own vision about neutrinos. Neutrino masses are very small, much smaller than the other fermion masses and this can be naturally explained by the violation of the total lepton number $L$ at a very large scale $M$. In Guido’s words\footnote{Sentences quoted here from works done in collaboration have all been written by Guido.}:

\begin{quote}
“Given that neutrino masses are certainly extremely small, it is really difficult from the theory point of view to avoid the conclusion that $L$ conservation must be violated. In fact, in terms of lepton number violation the smallness of neutrino masses can be explained as inversely proportional to the very large scale where $L$ is violated, of order $M_{GUT}$ or even $M_{Pl}$.”
\end{quote}

On dimensional grounds we have:

\begin{equation}
m_{\nu} \approx \sqrt{\Delta m_{\text{atm}}^2} \approx \left(\text{e.w. scale}\right)^2 M,
\end{equation}

leading to the estimate $M \approx 10^{14-15}$ GeV, not far from $M_{GUT}$. Guido considered this\footnote{Sentences quoted here from works done in collaboration have all been written by Guido.} “The most impressive numerology that comes out from neutrinos.” It is reasonable that the relation (1) arises from the seesaw mechanism, the simplest realisation of which requires a set of heavy right-handed neutrinos. To Guido these were strong indications in favour of a grand unified theory (GUT)\footnote{Sentences quoted here from works done in collaboration have all been written by Guido.}:

\begin{quote}
“We consider that the existence of right-handed neutrinos $\nu^c$ is quite plausible because all GUT groups larger than SU(5) require them. In particular the fact that $\nu^c$ completes the representation 16 of SO(10): $16 = \overline{5} + 10 + 1$, so that all fermions of each family are contained in a single representation of the unifying group, is too impressive not to be significant.”
\end{quote}

Guido believed that GUTs had to be incorporated in our picture of particle physics\footnote{Sentences quoted here from works done in collaboration have all been written by Guido.}:

\begin{quote}
“GUTs are the most attractive conjecture for the large scale picture of particle physics. GUT is not the Standard Model (SM), is beyond the SM, but is the most standard physics beyond the SM. Most of us think that there should be something like a
\end{quote}
Once the idea of heavy right-handed neutrinos is accepted, we get as a bonus an elegant explanation of the observed baryon asymmetry: “Another big plus of neutrinos is the elegant picture of baryogenesis through leptogenesis (after LEP has disfavoured baryogenesis at the weak scale).” Of course, once embedded in a GUT, neutrinos pose the problem of a consistent description of quarks and lepton masses and mixing angles in a less flexible setting. Guido regarded this challenge as a big opportunity. If dominated by the seesaw relation, $m_\nu = -m_\nu^T M^{-1}m_\nu$, neutrino masses are potentially linked to the other charged fermion masses. Back in 1998 the quark sector was already reasonably well-known, but a baseline model for quark masses and mixing angles was missing. Neutrino masses and the large atmospheric mixing angle indicated by the data were interesting new inputs which, especially in a constrained framework as the one provided by GUTs, could have brought a new insight into the flavour puzzle.

2 Lepton mixing angles and GUTs

In GUTs particle classification is greatly clarified. Quarks and leptons of the same generation belong to few multiplets of the grand unified group, a single representation being sufficient in the case of SO(10). Charge quantisation and gauge anomaly cancelation, which look miraculous within the SM, are thus neatly explained. Being members of the same gauge multiplets, quarks and leptons lose their fundamental distinction in GUTs and we should justify why in this context the lepton mixing angles are so different from the quark ones. An appealing explanation of this property can be found within SU(5) GUTs. In a minimal formulation of the SU(5) GUT, matter fields are described by three copies of the $10 = (q, u^c, e^c)$ and $\bar{5} = (l, d^c)$ representations, while the Higgs fields $\varphi$ and $\bar{\varphi}$ transform as 5 and $\bar{5}$, respectively. Fermion masses are described by the Yukawa interactions:

$$L_Y = 10 y_u 10 \varphi + 5 y_d 10 \bar{\varphi} + \frac{1}{M} 5 w 5 \varphi \bar{\varphi} + \ldots$$

where $y_{u,d}$ and $w$ are matrices in generation space. After electroweak symmetry breaking the first term describes up-quark masses, the last one is the grand unified version of the Weinberg operator and gives rise to neutrino masses of the type given in eq. (1). The second term describes at the same time down-quark masses and charged lepton masses, which are equal at the GUT scale.
in this approximations. Corrections to this relation are provided by additional contributions denoted by dots. It would be desirable that the matrices $y_{u,d}$ and $w$ had entries of the same order of magnitude, with no built-in structure. Starting from anarchical matrices $y_{u,d}$, we can easily produce the hierarchy observed in the charged fermion sector by a rescaling of the matter fields:

$$10 \rightarrow F_{10} \ 10 \ , \quad \bar{5} \rightarrow F_{\bar{5}} \ \bar{5} \ . \quad (3)$$

Here $F_{10,\bar{5}}$ are diagonal matrices of the type

$$F_X = \begin{pmatrix} \epsilon_X & 0 & 0 \\ 0 & \epsilon_X & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (1 \geq \epsilon_X \geq \epsilon'_X) \ . \quad (4)$$

For instance, after rescaling the 10 representations, the effective matrix of Yukawa couplings for the up quarks becomes

$$Y_u = F_{10} \ Y_u \ F_{10} \ , \quad (5)$$

which is hierarchical and nearly diagonal if $1 \gg \epsilon_{10} \gg \epsilon'_{10}$. By adjusting the suppression factors $\epsilon_{10}$ and $\epsilon'_{10}$ we can reproduce the quark masses and generate small contributions to the quark mixing angles. Such a mechanism is rather generic in model building. The rescaling matrices $F_X$ can arise in a variety of frameworks such as models with an abelian flavour symmetry, models with an extra dimension and models with partial compositeness or specific conformal dynamics 4).

Since the mass hierarchy in the down-quark and charged-lepton sectors is much less pronounced than in the up-quark sector, we need a milder rescaling from $F_{\bar{b}}$. As a useful reference we can choose

$$F_{\bar{b}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \ , \quad (6)$$

which corresponds to the so-called anarchy scenario 5). In this case we find

$$m_u : m_c : m_t \approx m_\nu^2 : m_e^2 : m_\mu^2 \approx m_\tau^2 : m_\tau^2 : m_\tau^2$$

which is approximately correct. Moreover, at the leading order we have

$$Y_e = Y_d^T \ , \quad (8)$$
where both $\mathcal{Y}_e$ and $\mathcal{Y}_d$ are lopsided matrices since only $F_{10}$ is operating. The relation (8) should be corrected since it leads to wrong mass equalities for the first two generations. The required corrections are sizeable, but not huge and (8) can still be valid at the level of orders of magnitude. In the limit where (8) is exact, it predicts a small contribution to the quark left-handed mixing and a large contribution to the lepton left-handed mixing, which is exactly what we observe. For the right-handed components a large (small) mixing for quarks (leptons) is predicted, which however is not observable at low energies.

The neutrino mass matrix is $m_\nu \propto F_5 w F_5 v^2/M$. When (6) holds neutrino mass ratios and mixing angles are generated from the random, order-one, matrix elements of $w$, which is consistent with the data to first approximation. Actually, the discovery of $\theta_{13} \approx 0.15$ and the first hints for a non-maximal atmospheric mixing angle have strengthened the case for anarchy. However within the extreme choice in eq. (6) there is no preference for the type of neutrino mass ordering and no explanation of the smallness of $\sin^2 \theta_{13}$ and $\Delta m_{\odot}^2/\Delta m_{\text{atm}}^2$. Guido thought that (6) could be replaced by a more generic possibility, such as

$$F_5 = \begin{pmatrix}
\lambda Q_1 & 0 & 0 \\
0 & \lambda Q_2 & 0 \\
0 & 0 & 1
\end{pmatrix}.$$  

(9)

Here $\lambda$ is an expansion parameter, typically smaller than 0.5 and $Q_{1,2}$ are two positive charges, $Q_1 \geq Q_2 \geq 0$. Anarchy is reproduced when $Q_{1,2} = 0$. It is not surprising that we found several examples with $Q_1$ non-vanishing where a small $\theta_{13}$ is more easily reproduced than in anarchy 6). In all the more successful examples the normal ordering of neutrino masses is preferred. First hints of such a preference are currently shown by several neutrino oscillation experiments.

Though rather appealing at first sight, this approach has clear limitations. The most severe one is that the entries of the matrices $y_{u,d}$ and $w$ are independent order-one parameters. Predictions for the various physical quantities can only be formulated in terms of distributions, assuming some statistical distribution for the unknown matrix elements of $y_{u,d}$ and $w$. Models in this class typically predict flat distributions for the CP violating phases. Thus features such as the closeness of the Dirac CP phase to the maximal value are purely accidental in this framework. It is not possible to go beyond order-of-magnitude estimates, whereas today we have precise data and we would like to
have models whose predictions can be tested at the level of accuracy reached by the present experiments.

3 More symmetry?

More predictive frameworks typically require more symmetries. Early model building has been largely influenced by some features of lepton mixing angles such as the smallness of $\theta_{13}$, the closeness of the atmospheric angle to the maximal value and, more recently, the indication of a maximal Dirac CP phase. Some form of quark-lepton complementarity has also been invoked. If one or more of these features are not accidental, they can help us in searching for some fundamental principle that rules the flavour sector. Several symmetric patterns of lepton mixing angles have been suggested in the past, such as the tribimaximal (TB) mixing or the bimaximal (BM) mixing:

$$U_{TB} = \begin{pmatrix} \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{3}} & 0 \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad U_{BM} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}. \quad (10)$$

They incorporate some of the above-mentioned features. Today we know that these patterns need sizeable corrections, but they can still be adopted as first order approximations to the true lepton mixing matrix $U_{PMNS}$. Following this approach we can regard $U_{PMNS}$ as an expansion around a leading order matrix $U_{PMNS}^0$, which can coincide with $U_{TB}, U_{BM}$ or some other symmetrical form:

$$U_{PMNS} = U_{PMNS}^0 + ... \quad (11)$$

where dots stand for corrections. It is not difficult to identify flavour symmetries leading to $U_{PMNS}^0$. For example discrete flavour symmetries showed very efficient in reproducing $U_{TB}, U_{BM}$ or other leading order patterns. These constructions require small non-abelian permutation groups, such as $A_4$ and $S_4$. In the so-called direct approach we can predict the three mixing angles and the CP violating phase, while neutrino masses are only constrained within extended ranges and are fitted by adjusting the free parameters $^7$.

Before a non-vanishing $\theta_{13}$ was established, the TB mixing nicely agreed with the data. In most of the models based on discrete symmetries, deviations from $U_{TB}$ were expected to be small, not to spoil the agreement between the
predicted and measured value of the solar mixing angle, which was known to a good accuracy. In particular the angle $\theta_{13}$ was expected not to exceed few degrees, a prediction that turn out to be wrong. While working with Guido on this topic, I was very excited about the neat prediction offered by our $A_4$ model. Guido, much wiser and forward-looking than me, wanted to add the following comment in our paper 8): “Special models are those where some symmetry or dynamical feature assures in a natural way the near vanishing of $\theta_{13}$ and/or of $\theta_{23} - \pi/4$. Normal models are conceptually more economical and much simpler to construct. We expect that experiment will eventually find that $\theta_{13}$ is not too small and that $\theta_{23}$ is sizably not maximal.”

Indeed $\theta_{13} \approx 90^\circ$ is much larger than the value predicted by the simplest schemes leading to TB mixing at the leading order. On the one side this supports models based on anarchy and its variants, which were expecting $\theta_{13}$ of about that size. On the other hand this result does not rule out models based on non-abelian symmetries, in particular discrete symmetries. For instance through a rotation in the 23 or 13 neutrino sectors, the TB mixing is modified into a pattern with non-vanishing $\theta_{13}$ and non-maximal $\theta_{23}$, while the solar angle is unchanged, to first approximation. Similarly the BM mixing can be altered by a rotation in the 12 charged lepton sector, bringing the solar angle to the experimentally allowed range and generating a non-vanishing $\theta_{13}$. These modifications can be obtained by relaxing the symmetry requirements and lead to testable sum rules among the three mixing angles and the Dirac CP violating phase. Alternatively we can look for other leading order mixing patterns, closer to the data, by scanning the set of discrete groups. There are infinitely many discrete groups and a full classification of all the related lepton mixing patterns exists now 9). Mixing angles close to the observed ones can be obtained by appealing to sufficiently large groups, but in all cases the Dirac CP phase is trivial, which is disfavored by the present data. Finally, we can combine discrete flavour symmetries with the CP symmetry and analyse the possible symmetry breaking patterns. In this case, even choosing small discrete groups one can reproduce realistic mixing angles and a non trivial Dirac phase.

One of the weak points of the approach is that there are no predictions for neutrino masses in models based on discrete symmetries. Moreover there is no hint for such symmetries from quarks. Large hierarchies and small mixing angles do not seem to require discrete groups. Extension to GUTs are pos-
sible and there are many existence proofs, but they look rather complicated. To summarize in the words of Guido 10): “In conclusion, one could have imagined that neutrinos would bring a decisive boost towards the formulation of a comprehensive understanding of fermion masses and mixings. In reality it is frustrating that no real illumination was sparked on the problem of flavor. We can reproduce in many different ways the observations, in a wide range that goes from anarchy to discrete flavor symmetries but we have not yet been able to single out a unique and convincing baseline for the understanding of fermion masses and mixings. In spite of many interesting ideas and the formulation of many elegant models the mysteries of the flavor structure of the three generations of fermions have not been much unveiled.”

Acknowledgements

I warmly thank Mario Greco and all the organizers of the Vulcano Workshop 2016 for the kind invitation and for the very nice atmosphere of the meeting. I am particularly grateful to them for the session dedicated to the memory of Guido Altarelli.

References


5. L. J. Hall, H. Murayama and N. Weiner, Phys. Rev. Lett. 84, 2572 (2000); N. Haba and H. Murayama, Phys. Rev. D 63, 053010 (2001);


LIMITS ON WARM DARK MATTER PARTICLE MASSES FROM THE ABUNDANCE OF ULTRA-FAINT GALAXIES AT HIGH REDSHIFTS IN THE HUBBLE FRONTIER FIELDS

Nicola Menci
INAF-Osservatorio Astronomico di Roma, Via di Frascati 33, Monte Porzio, Italy

Abstract
Comparing the measured abundance of the faintest galaxies with the maximum number density of dark matter halos in WDM cosmologies we set a robust limit \( m_X \geq 2.9 \text{ keV} \) for the mass of thermal relic WDM particles at a 1-\( \sigma \) confidence level, \( m_X \geq 2.4 \text{ keV} \) at 2-\( \sigma \), and \( m_X \geq 2.1 \text{ keV} \) at 3-\( \sigma \). These constitute the tightest constraints on WDM particle mass derived to date from galaxy abundance independently of the baryonic physics involved in galaxy formation. We also apply our method to WDM composed by sterile neutrinos produced via the Shi-Fuller mechanism.

1 Introduction
Warm Dark Matter (WDM hereafter) models of galaxy formation are based on DM candidates with masses in the keV scale \(^1\). In these models, the population of low-mass galaxies is characterized by lower abundances and shallower
central density profiles compared to Cold Dark Matter (CDM) due to the dissipation of small-scale density perturbations produced by the free-streaming of the lighter and faster DM particles. Thus, WDM scenarios have been proposed as a solution to some unsolved issues affecting the CDM model on small scales \( \lesssim 1 \text{ Mpc} \), like the steepness of the density profiles in the inner regions of dwarf galaxies and the over-abundance of faint dwarfs around our Galaxy and in our Local Group \(^2\), as well as in the field \(^3\), \(^4\). Indeed, while a refined treatment of baryonic effects entering galaxy formation (in particular feedback from Supernovae) can contribute to solve the problems \(^5\), feedback effects can hardly explain the over-prediction of the abundance of field dwarfs with \( V_{\text{vir}} \approx 40-60 \text{ km/s} \) \(^6\).

The effect of assuming WDM on galaxy formation strongly depends on the mass of the candidate DM particle. This determines the suppression of the density power spectrum compared to the CDM case, which drives the formation of cosmic structures. In fact, the half-mode mass \( M_{\text{hm}} \) - determining the mass scale at which the WDM spectrum is suppressed by 1/2 compared to CDM - is a strong inverse function of the WDM particle mass.

Existing astrophysical bounds on the thermal relic mass \( m_X \), have been set by different authors \(^7\), \(^8\), \(^9\) by comparing the predicted number density of satellite galaxies with the abundance of observed ultra-faint Milky Way satellites. Note however that the latter are appreciably sensitive to the assumed completeness corrections \(^10\), \(^11\). At higher redshifts \( z \approx 6 \) a limit \( m_X \gtrsim 1 \text{ keV} \) has been derived from the UV luminosity functions (LFs) of faint galaxies \( (M_{\text{UV}} \approx -16) \) \(^11\). Since these approaches are based on the comparison between observed LFs and predicted mass function of DM halos in different WDM models, the delicate issue in these methods is their dependence on the physics of baryons determining the mass-to-light ratio of faint galaxies. Uncertainties in the baryonic physics also affect \(^12\) the tighter constraints achieved so far \( m_X \gtrsim 3 \text{ keV} \), derived by comparing small scale structure in the Lyman-\(\alpha\) forest of high-resolution \( (z > 4) \) quasar spectra with hydrodynamical N-body simulations \(^13\). An effective way of bypassing the physics of baryons can be found by exploiting the maximum value \( \phi_{m_X}(z) \approx \phi(M_{\text{hm}}(m_X), z) \) of the halo mass distribution in WDM cosmology at masses close to the half-mode mass scale \( M_{\text{hm}} \) \(^14\), \(^15\), \(^16\), \(^17\). Measuring galaxy abundances larger than \( \phi_{m_X}(z) \) at a given redshift sets a lower limit on \( m_X \) which is completely independent
of the physics of the baryons, since any baryonic effect can only decrease the number of luminous galaxies compared to the number of host DM halos.

The previous applications of the above procedure at $z = 6 - 10$ \cite{18, 19} have provided lower limits $m_X \gtrsim 0.9$ keV (2$\sigma$), while $m_X \geq 1.5$ keV has been obtained comparing with galaxies at $z \approx 2$ \cite{20}. Obtaining tighter limits on $m_X$ with the above method requires reaching faint magnitudes $\approx -13$ at high redshifts $z \gtrsim 6$, which is beyond the limits of present telescopes. With the Hubble Frontier Fields (HFF) program, galaxies, with intrinsic magnitudes below the HST limits, can be detected thanks to magnification by foreground galaxy clusters. The HFF program has enabled the detection of galaxies with $M_{UV} \approx -15$ at $z \approx 6$ or $M_{UV} \leq -17$ at $z \approx 8$ \cite{21, 22}.

Recently, the observations of lensed background galaxies in Abell 2744 and MACS 0416 were used to measure the luminosity function of galaxies down to ultra-faint magnitudes $M_{UV} = -12.5$ at $z \approx 6$ (\cite{23} LFL16 hereafter). We show that such recent measurements constitute an unprecedented opportunity to derive strong constraints on the WDM particle mass $m_X$ which are independent of baryonic physics (gas physics, star formation) involved in galaxy formation.

2 Method

The computation of the halo mass function in WDM models is based on the standard procedure described and tested against N-body simulations in \cite{14, 15, 17, 16}. Here we provide a brief outline of the main steps.

The key quantity entering the mass function is the variance of the linear power spectrum $P(k)$ of DM perturbations (in terms of the wave-number $k = 2\pi/r$). Its dependence on the spatial scale $r$ of perturbations is

$$\frac{d \log \sigma^2}{d \log r} = -\frac{1}{2 \pi^2 \sigma^2(r)} \frac{P(1/r)}{r^3}.$$ \hspace{1cm} (1)

Here we have used a sharp-$k$ form (a top-hat sphere in Fourier space) for the window function $W(kr)$ relating the variance to the power spectrum $\sigma^2(M) = \int dk k^2 P(k) W(kr)/2 \pi^2$, with a normalization in the range $c = 2.5 - 2.7$ for the relation between the halo mass $M = 4\pi \rho(c r)^3/3$ and the filter scale $r$ \cite{14, 17}.

We shall consider the effect of such an uncertainty on our results.

In WDM scenarios the spectrum $P_{_{WDM}}$ is suppressed with respect to the CDM case $P_{_{CDM}}$ below a characteristic scale depending on the mass $m_X$ of
the WDM particles. If WDM is composed of relic thermalized particles, the suppression factor can be parametrized as

$$\frac{P_{WDM}(k)}{P_{CDM}(k)} = \left[1 + (\alpha k)^2\right]^{-10/\mu}. \tag{2}$$

Here the mass of the WDM mass enters through the quantity $\alpha = 0.049 \times (\Omega_X/0.25)^{0.11} (m_X/\text{keV})^{-1.11} (h/0.7)^{1.22} (h^{-1}/\text{Mpc})$, where $\Omega_X$ is the WDM density parameter, $h$ is the Hubble constant in units of 100 km/s/Mpc, and $\mu = 1.12$. The differential halo mass function (per unit log $M$) based on the extended Press & Schechter approach reads

$$\frac{d\phi}{d\log M} = \frac{1}{\bar{\rho}} \frac{\sigma^2}{M} f(\nu) \frac{d\log \sigma^2}{d\log r}. \tag{3}$$

Here $\nu \equiv \delta_c^2(t)/\sigma^2$ depends on the linearly extrapolated density for collapse in the spherical model $\delta_c = 1.686/D(t)$ and $D(t)$ is the growth factor of DM perturbations. We conservatively assume a spherical collapse model for which $f(\nu) = \sqrt{2\nu/\pi} \exp(-\nu/2)$. Assuming an ellipsoidal collapse model or including residual thermal velocities would yield even tighter constraints on the WDM particle mass.

The mass function in eq. 4 is computed after substituting eq. 1, with a power spectrum $P(k) = P_{WDM}(k)$ determined by the WDM particle mass $m_X$ after eq. 2 (for $P_{CDM}$ we adopt the form in 25)). The resulting mass functions are characterized by a maximum value at masses close to the half-mode mass. Correspondingly, the cumulative mass functions saturate to a maximum number density $\phi_{m_X}(z) \approx \phi(M_{hm}(m_X), z)$. The dependence of the scale $\alpha$ (eq. 3) on the WDM particle mass $m_X$ yields a half-mode mass ranging from $M_{hm} \approx 10^{10} M_\odot$ for $m_X = 1$ keV to $M_{hm} \approx 10^8 M_\odot$ for $m_X = 4$ keV.

3 Results

In fig. 1a we show the cumulative mass function $\phi(> M)$ computed from eq. 4 at $z = 6$ for different assumed WDM particle masses, adopting recent Planck cosmological parameters: $\Omega_m = 0.32$, $\Omega_X = 0.68$, $\Omega_b = 0.05$, $h = 0.7$, $\sigma_8 = 0.83$. All the mass functions saturate to a maximum number density $\bar{\phi}_{m_X} \approx \phi(M_{hm})$. This is compared with the observed number density $\phi_{obs}$ of galaxies with $M_{UV} \leq -12.5$ corresponding to the LFL16 LFs at $z = 6$ within 1-$\sigma$, 2-$\sigma$, and 3-$\sigma$ (shaded areas) estimated in 26).
In Figure 1b we compare $\phi_{\text{obs}}$ and $\phi_{mX}$ as a function of $m_X$. Since luminous galaxies cannot outnumber DM halos, the condition $\phi_{\text{obs}} \leq \phi_{mX}$ yields $m_X \gtrsim 2.9$ keV at 1-$\sigma$ level, $m_X \geq 2.4$ keV at 2-$\sigma$ level, and $m_X \geq 2.1$ keV at 3-$\sigma$ level. Our constraints are the tightest derived so far from galaxy counts. Although these constraints are less stringent than the 2-$\sigma$ limit $m_X \geq 3.3$ keV derived from the Lyman-$\alpha$ forest \cite{13}, our limits are entirely independent of the modelling of baryons physics which affects the constraints from the Lyman-$\alpha$ absorbers. Compared to earlier works \cite{18, 11, 20}, we derive significantly tighter constraints on $m_X$ due to the unprecedented depth reached by the LF measurements in LFL16. To provide a comparison with previous results and to show how the LFL16 measurements made it possible to significantly improve the constraint on $m_X$, we show in fig. 2a the thermal relic mass $m_X$ that can be probed by observing a given number density of galaxies $\phi_{mX}$ (in the

Figure 1: a) The cumulative mass functions computed at $z = 6$ for different values of the WDM particle mass $m_X$ shown by the labels on the right. The thickness of the lines represent the uncertainties in the theoretical predictions discussed in sect. 2. The shaded areas correspond to the observed number density of galaxies with $M_{\text{UV}} \leq -12.5$ within within 1-$\sigma$, 2-$\sigma$, and 3-$\sigma$ confidence levels.
b) For different values of the thermal relic mass $m_X$, we show the maximum value (including the theoretical uncertainties) of the predicted number density of DM halos $\phi$ at $z = 6$. The shaded areas represent the observed number density of galaxies with $M_{\text{UV}} \leq -12.5$ within 1-$\sigma$, 2-$\sigma$, and 3-$\sigma$ confidence levels.
y-axis) at different redshifts (x-axis). Such values are compared with the lower bounds set by different measurements at various redshifts. Thus, the contour corresponding to the lower tip of the arrow defines the mass $m_X$ probed by the corresponding observations (at 1-σ level). Our 1 − σ lower bound derived from LFL16 is shown by the large circle at $z = 6$ and provides the most stringent limit derived so far. We stress that the above limits are obtained for thermal relic WDM particles, whose power spectrum can be described by eq. 2. Analogous limits for the sterile neutrino mass can be derived assuming a specific production mechanism. E.g., the resonant production model depends on the lepton asymmetry $\mathcal{L}$. Thus, the corresponding spectrum depends not only on the mass of the sterile neutrino $m_{\text{sterile}}$ but also on the mixing angle $\theta$ with active neutrinos (related to $\mathcal{L}$ by the requirement that the WDM abundance

Figure 2: a) The contours show the maximum number density of DM halos (y-axis) obtained at different redshifts (x-axis) assuming different values for the WDM thermal relic mass $m_X$ (contour levels and colors). Such abundances are compared with the lower limit (at 1-σ level) set by the different UV galaxy LFs in the literature integrated down to their faintest magnitude bin at $z = 2$ [27], at $z = 3 - 4$ [28], and at $z > 6$ [25]. The thick dot corresponds to the UV LFs measured by LFL16 at $z = 6$, which provide the tightest bound on $m_X$.

b) For the case of resonantly produced sterile neutrinos, we show the constraints provided by our method on the combination of mixing angle $\theta$ and sterile neutrino mass $m_{\text{sterile}}$, and we compare them with previous upper bounds provided by XMM X-ray observations [29].
matches the observed DM cosmic density), which determines the oscillation probability \(\sin^2(2\theta)\). The spectra for a given combination of \(\sin^2(2\theta)\) and \(m_{\text{sterile}}\) can be computed \(^{31}\) and used to derive the abundance of high-redshift galaxies through eq. 4. The resulting constraints in the \(\sin^2(2\theta) - m_{\text{sterile}}\) plane are shown in fig. 2b and compared with upper bounds derived from the lack of X-ray flux resulting from different observations. We also show the point corresponding to the interpretation in terms of sterile neutrino decay of the recent unidentified X-ray line at 3.5 keV reported in observations of X-ray clusters \(^{32, 33}\).

References

GAMMA-RAYS SIGNATURES OF DARK MATTER : STATUS AND FUTURE PROSPECT

Aldo Morselli

*INFN Roma Tor Vergata*

Abstract

Indirect dark matter (DM) searches rely on detection of stable by-products of DM interactions to search for a signal of this elusive component of the Universe. Among these final products, gamma rays have recently played a major role in understanding the nature of the DM particle. We review the current status of indirect DM searches with gamma-ray observations and prospects with future instruments.

1 Introduction

High-energy phenomena in the cosmos, and in particular processes leading to the emission of gamma- rays in the energy range 300 KeV - 100 TeV, play a very special role in the understanding of our Universe. This energy range is
indeed associated with non-thermal phenomena and challenging particle acceleration processes. The Universe can be thought as a context where fundamental physics, relativistic processes, strong gravity regimes, and plasma instabilities can be explored in a way that is not possible to reproduce in our laboratories. High-energy astrophysics and atmospheric plasma physics are indeed not esoteric subjects, but are strongly linked with our daily life. Understanding cosmic high-energy processes has a large impact on our theories and laboratories applications. The technology involved in detecting gamma-rays is challenging and drives our ability to develop improved instruments for a large variety of applications.

The energy range between 300 Kev and 100 MeV is an experimentally very difficult range and remained uncovered since the time of COMPTEL. In this range a new instrument can address all astrophysics issues left open by the current generation of instruments. In particular better angular resolution in the energy range 10 MeV - 1 GeV is crucial to resolve patchy and complex features of diffuse sources in the Galaxy and in the Galactic Centre as well as increasing the point source sensitivity. This instrument can address scientific topics of great interest to the community, with particular emphasis on multifrequency correlation studies involving radio, optical, IR, X-ray, soft gamma-ray and TeV emission.

Above 100 MeV, thanks to the launch of the Fermi-LAT satellite and to the advent of a new generation of imaging air Cherenkov telescopes (H.E.S.S., MAGIC, VERITAS), several thousand gamma-ray sources are known today revealing an unexpected ubiquity of particle acceleration processes in the Universe.

Major scientific challenges are still ahead, such as the identification of the nature of Dark Matter, the discovery and understanding of the sources of cosmic rays, or the comprehension of the particle acceleration processes that are at work in the various objects.

The identification of the nature of Dark Matter can be done with the detection of gamma rays and cosmic rays from the annihilation or decay of dark matter particles. This is a promising method for identifying dark matter, understanding its intrinsic properties, and mapping its distribution in the universe.
Astrophysical searches for dark matter (DM) are a fundamental part of the experimental efforts to explore the dark sector. The strategy is to search for DM annihilation products in preferred regions of the sky, i.e., those with the highest expected DM concentrations and still close enough to yield high DM-induced fluxes at the Earth. For that reason, the Galactic Center (GC), nearby dwarf spheroidal galaxy (dSphs) satellites of the Milky Way, as well as local galaxy clusters are thought to be among the most promising objects for DM searches. In particular, dSphs represent very attractive targets because they are highly DM-dominated systems and are expected to be free from any other astrophysical gamma-ray emitters that might contaminate any potential DM signal. Although the expected signal cannot be as large as that from the GC, dSphs may produce a larger signal-to-noise ($S/N$) ratio. This fact allows us to place very competitive upper limits on the gamma-ray signal from DM annihilation $^1,^2,^3$), using data collected by the Large Area Telescope (LAT) onboard the Fermi gamma-ray observatory $^4$). These are often referred to as the most stringent limits on DM annihilation cross-section obtained so far.

Despite these interesting limits derived from dSphs, the GC is still expected to be the brightest source of DM annihilations in the gamma-ray sky by several orders of magnitude. Although several astrophysical processes at work in the crowded GC region make it extremely difficult to disentangle the DM signal from conventional emissions, the DM-induced gamma-ray emission is expected to be so large there that the search is still worthwhile. Furthermore, the DM density in the GC may be larger than what is typically obtained in $N$-body cosmological simulations. Ordinary matter (baryons) dominates the central region of our Galaxy $^5$). Thus, baryons may significantly affect the DM distribution. As baryons collapse and move to the center they increase the gravitational potential, which in turn forces the DM to contract and increase its density. This is a known and qualitatively well understood physical process $^6$). It is also observed in many cosmological simulations that include hydrodynamics and star formation $^8$). If this is the only effect of baryons, then the expected annihilation signal will substantially increase $^5,^7$).

A preliminary analysis of Fermi LAT observations of the GC region was presented in $^{13, 14}$), with an observation of an excess of gamma rays in the
Figure 1: Comparison of constraints on the DM annihilation cross section for the $\bar{b}b$ (left) and $\tau\tau$ (right) channels 29) with previously published constraints from LAT analysis of the Milky Way halo (3$\sigma$ limit) 24), 112 hours of observations of the Galactic Center with H.E.S.S. 9), and 157.9 hours of observations of Segue 1 with MAGIC 10). Closed contours and the marker with error bars show the best-fit cross section and mass from several interpretations of the Galactic center excess 15).

3 – 5 GeV energy range from the GC region. These results produced a lot of activity outside the Fermi collaboration with claims of evidence for dark matter in the Galactic Center (i.e. 15, 16) and references therein).

This possibility was already considered in the analysis of the EGRET galactic center excess 27) but there are other possible explanations, e.g. a population of millisecond pulsars around the Galactic Center below the Fermi threshold 19), 20).

A third possibility is related to past activity of the Galactic Center 17), 18). In this case the excess can be connected to the Fermi bubble and it will be very important to see how this bubble is structured in the GC region.

The analysis of the Fermi Collaboration 21) using 5 years of data and the Pass 7 event selections 22) and the ongoing analysis with 6.5 year of data and the Pass 8 event selections 28) confirm the excess but confirm also that when all the uncertainties on the excess morphology and spectrum related to the modeling of the various components of gamma-ray emission in that region, in particular in the distribution of interstellar gas along the line of sight, in the low latitude emission from the Fermi bubbles, and in the abundance of cosmic ray sources in the innermost Galaxy are considered the spectrum varies significantly and it is not possible to discriminate between the different hypotheses. The new
Figure 2: Comparison of constraints on the DM annihilation cross section for the $\tau\tau$ channel \textsuperscript{29} with Antares \textsuperscript{11}, IceCube-DeepCore \textsuperscript{12} and MAGIC \textsuperscript{10} analysis of the dSphs with the use of Pass 8 begin to constrain some of the preferred parameter space for a DM interpretation of a gamma-ray excess in the Galactic center region. As shown in Figure 1, for interpretations assuming a $b\bar{b}$ final state, the best-fit models lie in a region of parameter space slightly above the 95% CL upper limit from this analysis, with an annihilation cross section in the range of $(1-3) \times 10^{-26} \text{ cm}^3 \text{s}^{-1}$ and $m_{DM}$ between 25 and 50 GeV. However, uncertainties in the structure of the Galactic DM distribution can significantly enlarge the best-fit regions of $\langle \sigma v \rangle$ channel, and $m_{DM}$. Figure 2 shows a comparison of constraints on the DM annihilation cross section for the $\tau\tau$ channel \textsuperscript{29} with Antares \textsuperscript{11}, IceCube-DeepCore \textsuperscript{12} and MAGIC \textsuperscript{10}. One can see that the Fermi limits are the best limits below 2 TeV.

At lower energies a new instrument like Gamma-Light \textsuperscript{30}, or ASTROGAM \textsuperscript{31} can really improve these results both in the Galactic center and in the dSphs limits.

The project for an improved version of ASTROGAM, e-ASTROGAM, is being prepared in reply of the fifth ESA call for medium mission (M5) that
will be released in the fall of 2016.

Interactions of photons with matter in the e-ASTROGAM energy range is dominated by Compton scattering from 0.1 MeV up to about 15 MeV in silicon, and by electron-positron pair production in the field of a target nucleus at higher energies. e-ASTROGAM maximizes its efficiency for imaging and spectroscopy of energetic gamma rays by using both processes. The e-ASTROGAM instrument is based on double-sided Silicon detectors coupled to front-end-electronics capable of acquiring analog information on energy deposition in the range 20-1000 keV with high efficiency and high signal-to-noise. Both Compton events induced by photons in the range 0.3-30 MeV and pair production events in the 30 MeV - 30 GeV range can be detected by the e-ASTROGAM Tracker equipped with a Calorimeter and an Anticoincidence system. Figure 3 shows representative topologies for Compton and pair events.

For Compton events, point interactions of the gamma ray in tracker and calorimeter produce spatially resolved energy deposits, which have to be reconstructed in sequence using the redundant kinematic information from multiple interactions. Once the sequence is established, two sets of information are used
Figure 4: Point Spread Function (PSF, 68% containment radius) of the e-ASTROGAM gamma-ray detector. For comparison, we show the Fermi-LAT Pass7 PSF and the COMPTEL instrument. In the Compton domain, the performance of e-ASTROGAM and COMPTEL is the FWHM of the angular resolution measure (ARM).

for imaging: the total energy and the energy deposit in the first interaction measure the first Compton scatter angle. The combination with the direction of the scattered photon from the vertices of the first and second interactions generates a ring on the sky containing the source direction. Multiple photons from the same source enable a full deconvolution of the image, using probabilistic techniques. For energetic Compton scatters (above 1 MeV), measurement of the track of the scattered electron becomes possible, resulting in a reduction of the event ring to an arc, hence further improving event reconstruction. Compton scattering depends on polarization of the incoming photon, hence careful statistical analysis of the photons for a strong (e.g., transient) source yields a measurement of the degree of polarization of its high-energy emission. Pair events produce two main tracks from the electron and positron at small opening angle. Tracking of the initial opening angle and the plane spanned by
electron and positron enables direct back-projection of the source. Multiple scattering in the tracker material (or any intervening passive materials) leads to broadening of the tracks and limits the angular resolution at low energies.

The nuclear recoil taking up an unmeasured momentum results in a small uncertainty, usually negligible compared to instrumental effects. The energy of the gamma ray is measured using the calorimeter. Polarization information in the pair domain is given by the azimuthal orientation of the electron-positron plane. The Point Spread Function of e-ASTROGAM is shown in Figure 4, and the sensitivity is shown in Figure 5 is for an effective exposure of 1 year of a high Galactic latitude source. Sensitivities above 30 MeV are given at the 5-sigma confidence level, whereas those below 10 MeV (30 MeV for COMPTEL) are at 3-sigma.

3 Dark Matter Studies in the MeV - GeV domain

One of the major scientific objectives of e-ASTROGAM is the search for dark matter (DM) by means of the production of secondary gamma-rays after the annihilation or decay of the DM particle candidates. The importance of e-ASTROGAM for DM searches can be seen in Figure 6 where the differential γ-ray energy spectra per annihilation of WIMP are plotted. As one can...
Figure 6: Left: differential energy spectra per annihilation for a few sample annihilation channels and a fixed WIMP mass (200 GeV) and differential γ-ray energy spectra per annihilation for a fixed annihilation channel (bb) and for different values of WIMP masses \(^{27}\). For comparison we also show the emissivity, with an arbitrarily rescaled normalization, from the interaction of primaries with the interstellar medium. Right: The solid lines are the total yields for different annihilation channels, while the dashed lines are components not due to \(π^0\) decays.

see the bulk of the emission even for high WIMP masses is in the energy range 5 MeV - 100 MeV. Decaying DM can also produce a detectable line in the e-ASTROGAM energy range that might be detectable out of the continuum. Together with Fermi and CTA, e-ASTROGAM will probe most of the space of WIMP models with thermal relic annihilation cross section. Resolving the inner region of our Galaxy at high-energies remains one of the outstanding problems of modern astrophysics. Despite several attempts, the origin of positrons currently annihilating at the rate of \(2 \cdot 10^{43} \text{ s}^{-1}\) from the inner Galaxy is not accounted for by current models of star formation and compact object activities in the region. Recent data show that in addition to the central bulge also the inner disk is producing 511 keV emission. Candidate positron sources include: the central black hole activity, massive stars, Supernovae, compact binaries, pulsars, and possibly DM annihilation/de-excitation. The much improved e-ASTROGAM sensitivity at the electron-positron annihilation energy will be
used for a high-resolution mapping of the mysterious 511 KeV radiation. In the Fermi-LAT analysis of the Galactic Center the diffuse gamma-ray backgrounds and discrete sources, as we model them today, can account for the large majority of the detected gamma-ray emission from the Galactic Center.

Nevertheless a residual emission is left, not accounted for by the above models of standard astrophysical phenomena. In the crowded Galactic Center region the analysis to disentangle a possible DM signal from conventional emissions has still large uncertainties due to the extremely difficult subtraction of the Galactic diffuse emission and the contribution of unresolved sources. The very good angular resolution of e-ASTROGAM at low energies will help to resolve sources in the galactic center region and to disentangle the possible DM contribution, see Figure 7. e-ASTROGAM will also perform indirect DM detection searches in dwarf spheroidal galaxies and put constraints on DM contribution to the largely unknown diffuse extragalactic gamma-ray background in the spectral range 0.3 - 100 MeV. Models will be tested in a spectral range not yet currently studied.
References


17. J. Petrovic et al., 2014, JCAP, 10, 052
A realistic model for cosmic-rays production and the galactic center GeV excess.

Marco Taoso

Instituto de Física Teórica (IFT) UAM/CSIC, calle Nicolás Cabrera 13-15, 28049 Cantoblanco, Madrid, Spain

Abstract

Analysis of the Fermi-LAT gamma-ray data have uncovered an excess emission in the inner Galaxy, peaking at energies around few GeV and extended up to 10 degrees. Driven by the evidence of a large gas density in the inner kpc of the Galaxy, correlated with an impressive cosmic-rays acceleration, we investigate the possibility of addressing this excess in terms of ordinary cosmic-ray sources and standard steady-state diffusion. Remarkably, we find that this astrophysical scenario reproduces the morphological features of the excess, potentially explaining most of this emission.

1 Introduction

There is a remarkable agreement between models for the diffuse γ-ray emission in the Galaxy and data from the all-sky survey from the Fermi Large Area Telescope. A fairly good match is obtained in most region of the sky implementing only minor readjustments to a standard recipe based on: i) supernova remnants
Figure 1: Spectrum of the various contributions to the total $\gamma$-ray flux, pre- and post-template-fitting, compared to the Fermi-LAT data. Left panel: Model A+DM. Right panel: Model A+spike+DM.

(SNRs) as cosmic-ray (CR) sources; ii) the steady-state propagation of CRs in the Galaxy as tuned on local CR measurements; and iii) $\gamma$-ray emitting targets, namely the gas and the interstellar radiation field (ISRF), as indirectly derived from observations at other wavelengths.

One notable exception is the Galactic Center (GC) region. Several analyses have reported since 2009 an excess of gamma-rays from the inner few degrees around the GC (e.g. 1). This extended emission appears spherically symmetric and peaks at energies around few GeVs. Remarkably, all these features are compatible with those expected for a $\gamma$-ray signal produced by annihilations of dark matter (DM) particles 2, 3). Interestingly, the DM annihilation cross-section required to explain the emission is of the order of the typical one for thermal Weakly Interacting Massive Particles (WIMPs). This scenario therefore seems quite appealing, being also compatible with well motivated theories Beyond the Standard Model. The DM hypothesis is also in agreement with the constraints from other observations, in particular bounds from Fermi-LAT observations of dwarf galaxies and measurements of the anti-protons CRs 4).

Other competing scenarios accounting for the extra-emission have been suggested, notably the presence a population of unresolved millisecond pulsars 5), a bursting star-forming activity in the past near the GC 6), or non-thermal bremsstrahlung produced by a population of electrons interacting with neutral gas in molecular clouds 7). At the moment is still unclear whether there is a preferred scenario and further investigations are ongoing.
Figure 2: Residual counts obtained for the Model A (left panel, without the inclusion of DM template), for the Model A + DM (central panel), and for the spike model (right panel, without the inclusion of DM template).

The modeling of the diffuse gamma-ray emission in the GC region is a challenging task. All the three ingredients that we have listed above for the standard recipe show problematic aspects. From the point of view of CR propagation, there are several features making it likely that transport properties in the GC region differ significantly from average properties in the Galaxy, in particular the evidence for strong magnetic fields and large convective winds. Furthermore, the catalogues used for CR source distributions, mainly the observed distributions of SNRs and pulsars, are not optimal for the inner Galaxy. In fact, observations show that the GC harbour significant star formation and a large rate of Supernova explosions compared to the average value in the Galaxy (according to ref. 8 it is roughly a factor of 250 higher than the mean rate in the Galaxy). This is the consequence of a large reservoir of molecular gas filling the inner part of the Galactic bulge (the Central Molecular Zone 9) (CMZ) and the very peculiar physical properties of this environment.

In 10, we have shown that the enhancement of cosmic-ray sources in the CMZ drives major consequences on the $\gamma$-ray emissivities in the GC region. The main impact of this extra source in the Galactic center analysis is to enhance the IC emissivity introducing a morphological imprint analogous to DM pair annihilations. In the following we are going to details the main results of this analysis.
Figure 3: **Top left panel.** We compare the test-statistic of the models under consideration; a positive difference between two models means that the second model performs better. **Top right panel.** We compare the $\chi^2$ of the longitude profiles for the same models. **Bottom panels.** The same as the top right panel, for latitude and radial profiles.

2 The template fitting analysis

In order to compare the Fermi data with a theoretical model, we have performed a template fitting analysis. We have considered different morphological templates\(^1\), assigned to components connected to different physical emission mechanisms. They are: i) a template for the sum of the $\pi^0$ and bremsstrahlung

\[^1\text{In our analysis, all the diffusion emission models have been obtained using DRAGON and GammaSky}^{11}\)
emission (correlated with the gas), ii) Inverse Compton (IC) term (correlated with the ISRF), iii) an isotropic emission for the extra-Galactic background component, iv) a template for the so-called Fermi bubbles emission and v) templates for the point-sources based on the 3FGL catalogue provided by the Fermi collaboration. Each template is free to fluctuate in each energy bin and the normalization is found maximizing a Poisson likelihood. In this way one accounts for possible spectral distortions and normalization uncertainties of the different components within the theoretical model. Indications for an extra component can be claimed if a significant improvement of the overall fit is found when repeating the same exercise with an additional physically motivated template. We focus on a region of interest (ROI) centered in the GC with $|l| < 20^\circ$ $2^\circ < |b| < 20^\circ$ in Galactic coordinates.

First we adopt a reference diffuse emission model, labelled as ModelA in ref. 3). As shown in fig. 2, this model fails to correctly reproduce the gamma-ray emission in the GC region, and a roughly spherically symmetric excess emerges from the data. If we include in the analysis a template for DM annihilations (here modeled according to a generalized Navarro-Frenk-White distribution with $\gamma = 1.26$), the excess disappears (center plot of Fig. 2) and the presence of this template is preferred by the fit by a large statistical significance. Fig. 1 shows the best-fit spectra of the different templates: the DM flux peaks at energies of few GeVs. As proved in ref. 3), the specific choice of propagation parameters is not crucial for the template-fitting analysis and the properties of the excess are stable against a large set of models adopted for the Galactic diffuse emission.

However, standard diffusion models (including "ModelA") adopt as CR source function a smooth interpolation from SNR or pulsar catalogues. As already mentioned in the Introduction, these parametrization clearly do not include a satisfactory description of the Galactic bulge region. Therefore, we consider an extra steady-state CR source term, which we model as a Gaussian (hereafter "spike") with normalization $N$ and spatial extension $\sigma$ ($r$ is the galactocentric distance):

$$Q_{\text{spike}} = Q_0 \exp \left( -\frac{r^2}{\sigma^2} \right).$$  \hspace{1cm} (1)
3 A novel reference model

We choose $\sigma = 300$ pc, as sample value for the spike width, in agreement with astro-physical inputs. The normalization cannot be too large since the SFR in the GC cannot exceed few percent of the total rate in the Galaxy. Based on this requirement, we fix its value verifying \textit{a posteriori} that the spike emission absorbs the majority of the GC excess.

In fig. 1 and 2 we show the results of the template fitting analysis obtained with $N = 2.2\%$. The residual map, obtained adopting replacing the DM template with the spike is analogous to the DM case. If we consider the possibility of having at the same time the spike and the DM component, the contribution of the DM template is significantly reduced. It is consistent with zero (within the error bars obtained from the template fitting) in a large energy range, and –most importantly – gives rise to a featureless spectrum. In other words, most of the GC excess is absorbed by the presence of the spike.

This is due to the morphological properties of the IC template in this novel diffusion emission model. The spike, in our ROI, mainly affects the IC template, since the galactic plane, where most of the $\pi^0$ and bremsstrahlung is produced, is masked out. This enhanced source term induces an IC emission peaked at the GC and with morphological properties similar to those of the DM template. However, it is important to notice that, differently from the DM scenario, we are not treating the extra ingredient as an independent term, rather we are correlating its spectrum to the overall IC emission.

In order to scrutiny the performance of our model, in fig. 3 we compare, for each energy bin, the Test Statistic (TS) of the models under consideration ($TS = -2\Delta \log \mathcal{L}$). The presence of the DM template or the spike greatly improves the fit. More importantly, these two scenarios perform similarly at the level of statistical preference.

We also compare models and data analyzing the $\gamma$-ray profiles in longitude, latitude and radial directions (see ref. \textsuperscript{10} for details). We find that the spike provides a fit comparable in quality with the one obtained using the DM template in the mid-energy region, where the evidence of the GC excess was considered stronger. From the longitudinal profiles, we notice instead a preference for the DM template in the low energy bins. The presence of the spike seems to be disfavored since it produces an overshoot at low latitudes.
\(|l| \lesssim 4^\circ\) for \(E_\gamma \lesssim 0.5\) GeV. However we remark that: i) systematic errors, not included here, dominate at low energies. Their estimate is crucial to assess the significance of this discrepancy. ii) The CR transport at the GC is far from being understood and non-standard properties of diffusion, very much likely in such a complex environment, may completely alter the simplified picture that we have adopted in the description of the GC excess, especially at low energies.

4 Conclusions

We have shown that the presence of the spike depicts a viable astrophysical scenario potentially able to fully explain the GC excess\(^2\). Similar results have been found recently in ref. 12\(^2\), based on a detailed gas model of the CMZ as a tracer for the cosmic-ray injection. Further improvements of our models are necessary to describe the complex GC region, and possibly isolate and characterize new faint emissions, such that produced by DM annihilations.

References


\(^2\)Although we chose Model A as a reference case, we have verified that our results are not strongly dependent on the CR transport model. Moreover, they can be obtained also for different values of \(\sigma\) in the range 100 – 400 pc.


Abstract

About a century of experimental observations and theoretical arguments allows one to conclude that a large fraction of the Universe is composed by Dark Matter (DM) particles. Many possibilities are open on their nature(s) and interaction types. Moreover, the poor knowledge of many fundamental astrophysical, nuclear and particle Physics aspects as well as of some experimental and theoretical parameters, the different used approaches and/or target materials, etc. leave open space in serious comparisons. A model independent approach based on the investigation of the DM annual modulation signature with widely sensitive target materials and full control of all the detectors and running features at the needed level has allowed to unambiguously test their presence at galactic scale. Some arguments are shortly addressed here.
1 Introduction

In theories extending the Standard Model of particle physics, many candidates having different nature and interaction types have been proposed as DM particles, as e.g.: SUSY particles (as e.g. neutralino or sneutrino in various scenarios), inelastic DM in various scenarios, electron interacting DM, a heavy neutrino of the 4-th family, sterile neutrino, Kaluza-Klein particles, self-interacting DM, axion-like (light pseudoscalar and scalar candidate), mirror DM in various scenarios, Resonant DM, DM from exotic 4th generation quarks, Elementary Black holes, Planckian objects, Daemons, Composite DM, Light scalar WIMP through Higgs portal, Complex Scalar DM, specific two Higgs doublet models, exothermic DM, Secluded WIMPs, Asymmetric DM, Isospin-Violating DM, Singlet DM, Specific GU, SuperWIMPs, WIMPzilla, Dark Atoms (as O-Helium), etc.; a wide literature is available. Moreover, even a suitable particle not yet foreseen by theories could be the solution or one of the solutions.

In fact, considering the richness in particles of the visible matter which is less than 1% of the Universe density, one could also expect that the DM particles in the Universe can also be multicomponent. It is worth noting that often the definition “WIMP” is used as synonymous of DM particle; on the contrary it refers not to a specific particle, but to a class of different particles which can also have well different phenomenologies; moreover, many other DM candidates with well different nature and interaction types are available.

Often, the elastic scattering on target nuclei is the considered interaction process, but other processes are possible and considered in literature, as e.g. those where also electromagnetic radiation is produced. Hence, considering the richness of particle possibilities and the existing uncertainties on related astrophysical (e.g. halo model and related parameters, etc.), nuclear (e.g. form factors, spin factors, scaling laws, etc.) and particle physics (e.g. particle nature, interaction types, etc.), a widely-sensitive model independent approach is mandatory as well as full control of the running conditions. Most activities in the field are instead based on a particular a priori assumption on the nature of the DM particle and of its interaction, in order to try to overcome — by various kind of events subtraction/rejection — the limitation arising from their originally measured counting rate. On the other hand, it is worth noting that experiments at accelerators may prove — when they can state a solid model independent result — the existence of some possible DM candidate particles,
but they cannot credit that a certain particle is a/the only solution for DM particle(s). Moreover, DM candidate particles and scenarios (even e.g. in the case of the neutralino candidate) exist which cannot be investigated at accelerators.

The expected energy distribution for the interactions of DM particles in a terrestrial detector depends — among others — on their density and velocity distribution at Earth’s position. However, the experimental observations regarding the dark halo of our Galaxy do not allow one to get information on this crucial aspect without introducing a model for the Galaxy matter density. Because of its simplicity, the isothermal sphere model (which consists in a spherical infinite system with a flat rotational curve) is a widely used assumption for the DM density distribution, and thus in the evaluation of DM expected rates. However, many of its underlying assumptions (sphericity of the halo, absence of rotation, isotropy of the dispersion tensor, flatness of the rotational curve) are not strongly constrained by astrophysical observations. Moreover, the isothermal sphere is strictly unphysical and may only represent the behavior of the inner part of physical systems, since it has a total infinite mass and needs some cutoff at large radii. Thus, the use of more realistic halo models is mandatory in the interpretation and comparison procedures of different experiments, since the model dependent results can significantly vary \(^1, 2\).

In conclusion, the uncertainties still present on the shape of the DM halo and on the density and velocity distribution prevent the definition of a “standard” halo and illustrate how the comparisons among the experiments of direct detection of DM particles can be consistent even just considering this particular aspect (also see Ref. \(^3\)). Moreover, many other experimental and theoretical uncertainties exist and must be considered in whatever suitable model dependent analysis and comparison among the experiments of direct detection of DM particles.

2 The Dark Matter particles direct detection

Considering the many available DM candidate particles and scenarios, and the existing uncertainties on the astrophysical, nuclear and particle physics, a model independent approach, a ultra-low-background suitable target material, a large exposure and a full control of running conditions are mandatory to pursue a widely sensitive direct detection of DM particles in the galactic halo.
Actually, most activities in the field release marginal exposures even after many years underground, and often they do not offer suitable information e.g. about operational stability and procedures during the running periods, and generally base their analysis on a particular *a priori* assumption on the nature of the DM particle and its interaction, and on all the involved aspects of the overall scenario and related parameters. They assume the elastic scattering on target nuclei as the DM particles interaction with matter and pursue through data selection and several/many subtraction procedures the selection of a recoil-like sample in the data. It is worth noting that both the specific nature of the candidate and the kind of interaction are not identified since several candidates can give rise to nuclear recoils and with different kind of interaction types, and known undistinguishable recoil-like events from background exist. Moreover, e.g. the applied subtraction procedures are — by the fact — statistical and cannot offer an unambiguous identification of a similar signal because of known existing recoil-like indistinguishable background; tails of the subtracted populations can play a role as well. Finally, the electromagnetic component of the counting rate, statistically “rejected” by several procedures in this approach, can contain either the signal or part of it, and it will be lost.

A regards experimental activities with liquid noble gases - more recently considered in the field - both single and dual phase liquid/gas detectors (as XENON, LUX, DARKSIDE) (see e.g. Ref. 4, 5) and refs therein), the released results suffer e.g. because of their largely disuniform and non-linear response, of physical energy thresholds not suitably proved, of absence of routine calibration in the same running conditions, of the fact that - despite of the small light response (2.28 photoelectron/keVee) - an energy threshold at 1 keVee is claimed, the energy resolution is poor and its naive convolution give rise to illusory sensitivity to low mass candidates in the single — largely arbitrary — fixed scenario they adopt, the behaviour of the light yield for recoils at low energy is uncertain, in the scale-up of the detectors the performances deteriorate, etc. For detailed discussion the reader can refer to the dedicated paper 5) and in other literature.

A positive hint for a signal of light DM candidates has been reported by the CoGeNT experiment 6, 7).

In the double read-out bolometric technique, the heat signal and the ionization signal are used to try to discriminate between electromagnetic and
recoil-like events (as CDMS and EDELWEISS). Generally the published exposures are absolutely marginal and hugely selected. Some comments can be found e.g. in 8). In these very small exposure experiments few recoil-like events survive the many selections/subtractions cuts applied in the data analysis; these events are generally interpreted in terms of background. In particular, the results of CDMS-II with the Si detectors were published in two close-in-time data releases 9, 10); while no events in six detectors (corresponding exposure of only 55.9 kg×day before analysis cuts) were reported in the former 9), three events in eight detectors (corresponding raw exposure of 140.2 kg×day) were reported over the residual background, estimated after subtraction: ≃ 0.4 in the second one 10).

Finally I remind the case of the CRESST-II experiment, which exploits the double read-out bolometric technique, using the heat signal due to an interacting particle in the CaWO$_4$ crystals and the heating of another device by scintillation light produced in the crystal at same time. The light signal is very poor and the possibility to efficiently collect all is — in my opinion — questionable. However, a statistical discrimination of nuclear recoil-like events from electromagnetic radiation is performed, and many cuts and selection procedures are applied. A previous run (8 detectors of 300 g each one, for an exposure of about 730 kg × day) showed that, after selections, 67 nuclear recoil-like events were observed in the Oxygen band 11) and a 4σ effect for possible signal was claimed. However, this result has been not confirmed in last run 12), where however a more marginal exposure has been used (52 kg × day and energy threshold of 0.6 keV). This discrepancy confirms the difficulties in managing the systematics in such kind of experiment.

In conclusion, suitable experiments offering a model independent signature for the presence of DM particles in the galactic halo are mandatory, as those realized by DAMA (see next section).

3 The DM model independent results of DAMA

To obtain a reliable signature for the presence of DM particles in the galactic halo, it is necessary to exploit a suitable model independent signature. With the present technology, one feasible and able to test a large range of cross sections and of DM particle halo densities, is the so-called DM annual modulation signature 13). The annual modulation of the signal rate originates from the
Earth revolution around the Sun. In fact, as a consequence of its annual revolution around the Sun, which is moving in the Galaxy traveling with respect to the Local Standard of Rest towards the star Vega near the constellation of Hercules, the Earth should be crossed by a larger flux of DM particles around \( \sim 2 \) June (when the Earth orbital velocity is summed to the one of the solar system with respect to the Galaxy) and by a smaller one around \( \sim 2 \) December (when the two velocities are subtracted). Thus, this signature has a different origin and peculiarities than the seasons on the Earth and than effects correlated with seasons (consider the expected value of the phase as well as the other requirements listed below). This DM annual modulation signature is very distinctive since the effect induced by DM particles must simultaneously satisfy all the following requirements: (1) the rate must contain a component modulated according to a cosine function; (2) with one year period; (3) with a phase that peaks roughly around \( \sim 2 \)nd June; (4) this modulation must be present only in a well-defined low energy range, where DM particles can induce signals; (5) it must be present only in those events where just a single detector, among all the available ones in the used set-up, actually “fires” (single-hit events), since the probability that DM particles experience multiple interactions is negligible; (6) the modulation amplitude in the region of maximal sensitivity has to be \( \lesssim 7\% \) in case of usually adopted halo distributions, but it may be significantly larger in case of some particular scenarios such as e.g. those in Ref. 14, 15).

This signature has been exploited with large exposure — using highly radiopure NaI(Tl) as target material — by the former DAMA/NaI \((\sim 100 \text{ kg sensitive mass})\) experiment and by the currently running DAMA/LIBRA \((\sim 250 \text{ kg sensitive mass})\), within the DAMA project 16, 17, 2, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39).

The DAMA/NaI and DAMA/LIBRA-phase1 results give evidence for the presence of DM particles in the galactic halo, on the basis of the exploited model independent DM annual modulation signature, at 9.3 \( \sigma \) C.L. The modulation amplitude of the single-hit events in the \( (2-6) \text{ keV} \) energy interval in NaI(Tl) target is: \((0.0112 \pm 0.0012) \text{ cpd/kg/keV}\); the measured phase is \((144 \pm 7) \text{ days}\) and the measured period is \((0.998 \pm 0.002) \text{ yr}\), values well in agreement with those expected for DM particles. No systematic or side reaction able to mimic the exploited DM signature has been found or suggested by anyone over more than a decade.
Recently an investigation of possible diurnal effects in the single-hit low energy scintillation events collected by DAMA/LIBRA-phase1 has been carried out \(^{36}\). A model-independent diurnal effect with the sidereal time is expected for DM because of Earth rotation. At the present level of sensitivity the presence of any significant diurnal variation and of diurnal time structures in the data can be excluded for both the cases of solar and sidereal time; in particular, the DM diurnal modulation amplitude as a function of the sidereal time expected – because of the Earth diurnal motion – on the basis of the DAMA DM annual modulation results is below the present sensitivity \(^{36}\). It will be possible to investigate such a diurnal effect with adequate sensitivity only when a much larger exposure will be available and exploiting the lower energy threshold as in the presently running DAMA/LIBRA-phase2. For completeness we recall that a recent analysis has been performed considering the so called “Earth Shadow Effect” \(^{38}\).

After a first upgrade in 2008, a further upgrade of DAMA/LIBRA has been performed at the end of 2010 when all the PMTs have been replaced with new ones having higher quantum efficiency \(^{30}\). Since then, after tests and optimization periods, the DAMA/LIBRA-phase2 is continuously running in order: (1) to increase the experimental sensitivity lowering the software energy threshold of the experiment; (2) to improve the corollary investigation on the nature of the DM particle and related astrophysical, nuclear and particle physics arguments; (3) to investigate other signal features and second order effects. DAMA/LIBRA also continue its study on several other rare processes \(^{33, 34, 35, 36, 37, 38, 39}\).

The DM model-independent DAMA result is compatible with a wide set of scenarios regarding the nature of the DM candidate and related astrophysical, nuclear and particle Physics. For example some given scenarios and parameters are discussed e.g. in Ref. \(^{2, 18, 19, 20, 21, 22, 23, 24, 26, 32, 39}\). Further large literature is available on the topics \(^{40}\); other possibilities are open. Let us remark that no other experiment exists, whose result can be directly compared – at least in principle – in a model-independent way with those by DAMA/NaI and DAMA/LIBRA.

It should also be stressed that the so-called “Snowmass plot” (and the analogous reported in the Ref. \(^{41}\)), where in the plane cross section on nucleon vs particle mass there are depicted all together some kinds of exclu-
sion plots, allowed regions and sensitivity curves, has not universal validity for the reasons given above and other ones (see for example discussions in Ref. 2, 26, 18, 4, 5, 42)).

4 Future perspectives for the DM directionality approach

The directionality approach — based on the study of the correlation between the recoil direction of the target nuclei and the Earth motion in the galactic rest frame — can offer a good approach to study those DM candidate particles able to induce just nuclear recoils. In particular, in the case of DM candidate particles interacting with nuclei the induced nuclear recoils are expected to be strongly correlated with the impinging direction of DM, while the background events are not; therefore, the study of the nuclear recoils direction can offer a way for pointing out the presence of these DM candidate particles.

This approach has some technical difficulties because it is arduous to detect the short recoil track. Different techniques are under consideration but, up to now, they are at R&D stage and have not produced yet competitive results in the field (see e.g. DRIFT, DMTPC, DAMIC, NEWS). In fact, they are generally limited by the difficulty of detecting very short tracks and of achieving high stability, large sensitive volume and very good spatial resolution. To overcome such a difficulty, it has been suggested the use of anisotropic scintillator detectors 43, 44, 45); their use was proposed for the first time in Ref. 43) and revisited in Ref. 44).

In particular, low background ZnWO$_4$ crystal scintillators have been recently proposed since their features and performances are very promising 46). In fact, both the light output and the scintillation pulse shape depend on the impinging direction of heavy particles (p, alpha, nuclear recoils, etc.) with respect to the crystal axes and can supply two independent ways to study the directionality and to discriminate the electromagnetic events (that does not give rise to any anisotropic effects).

Other advantages offered by ZnWO$_4$ detectors are very good radio-purity starting levels (about 0.1 cpd/kg/keV at low energy) and the potentiality to reach energy thresholds at keV level. Both these features can also be improved (e.g., the light yield shows a significant enhancement when working at low temperatures — about 100 K — and better radiopurity levels can be reached with dedicated R&D). Discussions can be found in Ref. 46).
5 Conclusions

The DM model independent annual modulation signature with widely sensitive target materials still remains a major approach, offering an unique possibility for detection; it requires well known techniques, full proved detector stability, well known and proved detector response in all the aspects, etc.

The DAMA positive model independent evidence for the presence of DM particles in the galactic halo is supported at very high confidence level. It has been shown in literature that this is compatible with many DM scenarios.

At present DAMA/LIBRA-phase2 is running with a lower software energy threshold.

References


Euclid: the next ESA mission for Cosmology

Roberto Scaramella
Osservatorio Astronomico di Roma
on behalf of the Euclid Collaboration

Abstract

Euclid is an approved ESA mission \(^1\), whose main aim is to study via gravitational lensing and the clustering of galaxies the cosmological Dark Sector (Energy, Matter). It will measure billions of galaxies over \(\sim15,000\) square degrees of the extragalactic sky in optical imaging, NIR photometry and NIR slitless spectra. The huge database will be also invaluable for most areas of astronomy. We give a brief overview of the mission goals and methods, more details can be found in \(^1\) and \(^2\).

1 Introduction

1.1 The Why

Since the later thirties of the past century Astronomy and then Physics have tried to solve the puzzle of the presence of the "Dark Matter" [DM], i.e. its
abundance, mass, origin, interaction properties and uniqueness (more than a single kind?). Now, since the beginning of this century another riddle has added, i.e. the cause of the observed acceleration of the expansion of the Universe. The observations of distant SuperNovae and of the Cosmic Microwave Background [CMB] show that the spatial curvature is zero but this cannot be ascribed to an universe of critical density and therefore suggest the presence of a cosmological constant $\Lambda$. However, it is well known that such a constant has not much appeal and therefore people have been investigating two alternative lines of approach: either the presence of an evolving quantum field (the "Dark Energy" [DE]) or some modification to the General Relativity on large scales (for a review of theoretical models related to Euclid see ⁵); from now on we will refer only to the DE but still implying the possibility of modified GR). It must be recalled that most of the effects of the DE appears at redshift $z \lesssim 3$ (see Fig. 1 of ⁶) and this is the epoch of major interest for Euclid, fully complementary to the informations we get from the CMB.

1.2 The How

In order to study these issues one needs to have not only geometrical test which determine the expansion history ($H(z)$) but also the characteristics (i.e. the speed) of the growth of the gravitational instability, that is the derivative with time (or expansion) of the matter density contrast, $\delta \equiv \delta \rho / \rho$. Usually this is parametrised like: $d[ln(\delta)]/d[ln(a)] = f = \Omega', \gamma_{\Lambda CDM} \simeq 0.55$ and $a$ denotes the expansion factor, $a \equiv 1/(1+z)$. Different models of DE reflect into different $\gamma$. It is important to notice that the latter collapse not only comprises but is also influenced by the dark matter characteristics.

Among several probes two are the most promising and actively pursued with ongoing and planned astronomical surveys. One is the Galaxy Clustering [GC], the other is the Weak Lensing [WL]. Both methods use galaxies as tracers (as direct particles and as background shapes, respectively) and, being statistically based, need huge samples to reach the wanted precision. The possible presence of a quantum field, responsible for the expansion is tested via its equation of state, $w = p/\rho$. The case of cosmological constant corresponds to $w = -1$, the more general case has a different values and is a function of the expansion factor, such as $w = w(a)$. A Taylor expansion can be done as
\[ w = w_0 + w_a (1 - a) \]  and limits are sought on the possible values in the plane \( w_0 - w_a \). Different experiments can contain different areas in such a plane and the inverse of the area if called Figure of Merit [FoM] \(^7\), so that the more informative (the better) is the experiment, the larger is the area. Euclid aims to reach from the combination of its probes alone at least \( FoM = 400 \) \(^1\).

1.3 Why space

Nowadays, tens of astronomical surveys are ongoing or planned in order to study these questions, with months, even years of observing time on large telescopes or even new, dedicated facilities. So why a space mission is needed? The answer on the interpretation side is that we need not only a large precision (a small uncertainty, so need large areas for many independent volumes and billion of galaxies for Poisson) but also and especially a large accuracy, i.e. to keep under control systematics as much as possible \(^8\) \(^9\). Now, to go in space means to face some new problems but also to get rid to all those given by the atmosphere and its variability.

In fact for lensing one needs exquisite imaging and extreme stability of the whole apparatus so to ease the problem of getting the maximum fidelity in the model of the actual point spread function [PSF], needed to extract the intrinsic distortion of the galaxy shapes. \(^10\) This knowledge and its stability are the major problem in ground-based surveys.

For the NIR data, differently, apart from the stability the space is useful because of a drastically smaller background with respect the one given by our atmosphere.

The large decrease in background dramatically increases the signal to noise ratio in the NIR bands, making even a small telescope competitive with the large ground based ones such as that in order to cover same areas at same depths a ground based NIR survey would need to spend several tens of years.\(^1\).

\(^1\)James Webb Space Telescope will be in orbit as well and with its diameter of 6 m will go much deeper and faster than Euclid but only on small areas: hence the two are complementary and JWST will possibly benefit of targets preselected from the Euclid surveys.
Figure 1: An estimate of the spectrum of the Zodiacal background with evidenced the range covered by Euclid instruments (adapted from 11).

2 The mission

The Euclid mission is expected to last six years and to be launched in year 2020. All data (raw and reduced) will be public and delivered in releases during the mission and after its completion.

3 Spacecraft and orbit

For the primary science goals of investigating dark energy with weak lensing and galaxy clustering, Euclid will need to perform a wide survey over a large fraction of the extragalactic sky. It is referred to as the Euclid Wide Survey [EWS]. The EWS will allow us to make the measurements necessary to perform the primary science. Its main characteristics are set by the Weak Lensing [WL] and Galaxy Clustering [GC] requirements and are mainly reflected in the area to be covered and the depth to be reached ensuring that the required average numbers of galaxies useful for these probes is met.

ESA provides the platform, the service module and the telescope. The
main mirror has a diameter of 1.2 m and a net collecting area of 1 m². Light then is split from a dichroic into a visible and a NIR channel which feed the instruments.

Euclid will be placed in orbit at the L2 position (second Lagrangian point) so to have maximum thermal stability and small occultation from Earth.

3.1 The Instruments

The instruments are provided by some of the European National Space Agencies, and Scientific Institutes which formed an ad hoc consortium, The Euclid Consortium [EC]. NASA contributes the NIR detectors.

3.1.1 VIS

The instrument VIS is devoted to the the weak lensing measurements 1) 12). It will take an image in a single r+i+z band covering the wide interval 550-900 nm over a field of view of ~0.5 deg². By stacking 4 dithered exposures with a total of 2260 sec, VIS will reach the depth of $m_{AB} = 24.5$ (at 10$\sigma$) for sources with diameter (FWHM) ~0.3arcsec. The image sampling of the 36 CCD’s is 0.1arcsec.
3.1.2 NISP

The NISP instrument \(^1\) \(^{13}\) works either as an imager or as a slitless spectrometer in the Near Infrared. Its focal plane, matched in area to that of VIS, contains an array of 16 NIR detectors with 0.3arcsec per pixel and 2.3\(\mu m\) cut-off wavelength. When operated in photometer mode [NISP-P] it can take images in y, J and H band. A grism wheel separated from the filter one holds three similar NIR-red grisms which differ in dispersion direction plus one NIR-blue grism. In the slitless spectrometry mode [NISP-S], the light is dispersed by a grism in the wavelength range 0.92 to 1.85 \(\mu m\) with a constant \(\Delta \lambda\) at a spectral resolution \(\lambda/\Delta \lambda > 380\) for an object of 0.5 arcsec diameter.

3.2 The Observations

The Euclid mission then is primarily defined as a unique survey intended to repeat a pre-defined sequence of observations over the whole useful sky defined by the Wide and Deep Survey requirements and their associated calibration sequence. The Survey concept aims at tiling the selected areas of the sky with the reference elementary observation pattern in the most efficient way. \(^{14}\)

The EWS will be covered with \(\sim\) 30,000 fields, each covering 0.5 sq deg with four dithered exposures, with a minimal overlap among them. The exposure times and sequences are fixed to ensure the best repeatability and calibration. The price to pay, however, is to have a survey of varying depth, because the back/foregrounds, i.e. zodiacal, scattered light, extinction, vary with direction (and time of observation for the first). Therefore the goal values are set and met on the average over the whole.

An important aspect and a strong limit to the possible pointing is given by the tight constraints on the thermal stability which translate into very small ranges for the possible pointing angles. In practice Euclid will observe almost always orthogonally to the sun. This then limits drastically the visibility of a region according to its ecliptic latitude: at the two poles the visibility is perennial but only over a radius of 3 degs for each cap. For a cap of radius up to 15 degs the visibility is a continuous stretch of \(\sim\) six months. For other areas each direction can be seen only twice at year (this is obtained by flipping the telescope six months apart), with a visibility which becomes shorter and shorter the closer to the ecliptic plane.

The major sources of problems are the background of zodiacal light, which
is maximum at the ecliptic plane and minimal at the poles, the extinction by dust and the scattered light from stars. Both the last issues suggest to stay away from the galaxy plane but the first pushes to stay far from the ecliptic plane as well. Therefore the less contaminated areas are in four connected regions, given by avoiding those two strips. In addition, single spots will be contaminated by scattered light from bright stars as well. These effects are readily seen in Fig.3

![Figure 3: An estimate of the SNR for the VIS and H photometry.](image)

### 3.2.1 The wide Survey

Each wide field will be observed only once. The basic cycle of a sequence of observations is first an exposure of VIS and NISP-S done in parallel with no motions of wheels. Then, VIS is closed and it starts the series of y, J, H band photometry, which need rotations of the grism wheel and of the filter wheel. The same sequence is repeated three more times after a small dither in between for a total of four exposures on each field. There is a variation, however, in the four otherwise identical sequences: in each of these the main axis of the dispersion of the red grism is rotated so to have three independent orthogonal spectra for each object, which helps in diminishing confusion given by contaminating neighbours\(^2\). The dither pattern is chosen such that every pixel of the joint focal plane is seen at least three times.

\(^2\)At the end one has a stack of four red spectra for the wide. This is an update with respect earlier solutions, which also had blueward NIR spectra for the wide \(^1\)
Figure 4: two views of the Euclid survey. Colours of the areas differ in time of observation.

3.2.2 Calibrations

We have two kind of calibrations: (i) the instrument calibrations done against standards (e.g. lamps or specific objects such as Planetary Nebulae for spectroscopy) or by repeated observations of the same field (for flats and stability). These are several and complex and will take in total ~ 6 months of time. Then we have (ii) sample characterisations: we need to know very well the intrinsic and observable characteristics of the galaxy sample under study. To do so one needs deeper observations with nearly 100% knowledge of the galaxies which are observed in the (shallower) wide survey. This will be achieved by devoting six months of observing times to three deep fields.

3.2.3 The Deep Fields

Three deep fields will be observed several times so to reach the two goals of having: (a) ~ 40 sq deg of spectroscopy with nearly zero confusion, i.e.
observed at least ten times with 30 independent spectral directions, separated by at least 5 degs each with no gap larger then 20 degs. And (b) $\sim 40$ sq deg of imaging at least two magnitudes deeper than the average wide $^{15}$, which translates in a total of 40 repeated passes. Constraint (a) is quite difficult since it requires strict timings for the ten separated visits. A solution fulfilling is found which also maximises science output with three Euclid Deep Fields [EDFs]. At the Northern Ecliptic Pole we have EDF-N, a two tier field (20 sq degs fulfilling (a) with ten visits and its inner ten sq degs visited 30 times more for (b)). Then we have EDF-S1 which is a 20 sq deg field close to the South Ecliptic Pole$^3$ with both (a) and (b). Finally is EDF-S2, a ten sq deg field centered on the Chandra Deep Field South and observed for (b) only. For general science purposes in the EDFs will be used also the on board NIR blue grism besides the mandatory NIR red exposures.

3.3 The Ground Segment

The dataset produced by Euclid will be enormous, not only in storage but also in complexity, given three different modes of operations (VIS, NISP-P, NISP-S), all the related calibrations. To this adds the need of dealing with ground based photometry (from various surveys) used to complement the NIR photometry from Euclid so to get reliable photometric redshifts needed for the WL tomography.

Therefore the Ground segment in terms of complexity, human and hardware resources has an impact comparable to the one of the instruments and is entirely provided by the EC.

The actual data reduction will be done in national centres (all can equally deal with part of the data) but the previous step, i.e how to reduce data and produce scientific useful material is done in different Organisational Unites, who are responsible of the algorithms and the validation of the results. $^{16}$). Each of these unit focuses none of the major steps, e.g. there are three for the Euclid raw data, one for the ground based data, one for the matching of different data sets, one for photometric redshifts, one for the measurement of emission lines, one for the shear, one for simulations, one for the production

$^3$close because one needs to stay clear of the outskirts of the Large Magellanic Cloud
of science ready data (e.g. correlation functions, catalog of clusters, power spectra etc).

4 Acknowledgements

Euclid is an enterprise at which are working hundreds of people at ESA and Euclid Consortium. I acknowledge partial support by ASI contract 1/023/12/0.

References


232
THE RELATIVISTIC ALL-SKY ANALYSIS WITH GAIA

Mariateresa Crosta
Istituto Nazionale di Astrofisica - Osservatorio Astrofisico di Torino
Via Osservatorio, 20 - Pino Torinese, Torino

Abstract

By providing an homogenous all-sky survey of high precision parallaxes, space motion (proper motions and radial velocities) and astrophysical characterization for more than one billion stars throughout the Galaxy and thanks to the depth of the volume achievable, Gaia will deliver a huge amount of astrometric, spectroscopic, and photometric data. Gaia will contribute also to the determination of an optical reference frame by observing many thousands of quasars. In doing so Gaia will have a huge impact across many fields, including many branches of stellar astrophysics (details of the structure and stellar evolutionary phases), exoplanets, solar system objects, the cosmic distance ladder (through a model independent of the primary calibrators) and fundamental physics. New “accurate” distances and motions of the stars within our Galaxy will provide access to the cosmological signatures left in the disk and halo offering independent, direct and detailed comparisons the predictions of the most advanced cosmological simulations. But all the above goals will not be achieved without the correct characterization and exploitation of the “relativistic”, i.e. very high
accuracy, astrometric data. Since a Gaia-like observer is positioned inside the Solar System, the measurements are performed in a weak gravitational regime which can be regarded as “strong” when one has to compare these slow varying fields with the accuracy achievable by Gaia.

1 Gaia Mission

Gaia (European Space Agency, ESA) is the first astrometric mission of the twenty-first century dedicated to the study of the Milky Way and was successfully launched on 19th December 2013 from the European base of Kourou in French Guyana.

At L2 Sun-Earth system Gaia is performing absolute parallaxes, combining at the same time two different stellar directions in one focal plane, observing all objects that pass away in its two fields of view, and scanning repeatedly the sky for at least 5 years. Precession at fixed angle to the Sun (45 degrees) ensures sky coverage. Nearly one-two billion astronomical objects will be observed on about 80 times, leading to around 630 CCD transits, so a total of more than 150 billion measurements at the end of the mission. Routine science operations started end-of-August 2014, after an extended commissioning phase which formally ended on July 18, 2014, followed by approximately 1-month of science calibrations with Gaia in EPSL (spin axis not precessing at 45 deg on ecliptic) using Ecliptic Pole special catalog.

Gaia’s survey provides the detailed 3D distributions and space motion of some 1 billion individual stars in our Galaxy and beyond, extended to G=20.7 (i.e., V=21), but not complete at this magnitude limit.

The schedule for the first general all-Gaia data delivery (DR1) has been confirmed on 14th September 2016 (corresponding to 1000 days into science operations in Nominal Scanning Law). DR1 contains the five-parameter astrometric solution - positions, parallaxes, and proper motions - for stars in common between the Tycho-2 Catalogue and Gaia (TGAS), namely for 2 million stars complete to V=11.5 (solar neighborhood, open clusters and associations, moving groups, ..) with sub-milliarcsec accuracy (10 % at 300 pc), while at the end-of-mission the astrometric accuracies are expected better than 5-10µas (microarcsecond) for the brighter stars and 130-600µas for faint targets.

The location of an object in astrometry is considered reliable if the rela-
tive error in parallax is less than 10%. This implies that with the microarc-second level of accuracy we get the galactic scale. Such a depth allows Gaia to contribute to our knowledge of Galaxy origin and formation. Galactic structure and dynamics; it will provide detailed information to better understand the physics of stars and their evolution; tens of thousands of brown dwarfs and white dwarfs will be identified; ten million binaries within 250 pc will be resolved; many thousands extra-solar planets and thousands of extragalactic supernovae will be discovered; 500,000 quasars will be pinpointed for celestial reference frames; the solar-system observations will include hundreds of thousands of minor planets, near-Earth objects, inner Trojans and new trans-Neptunian objects; finally, even fundamental physics will be tested (section 3). For details, refer to http://www.cosmos.esa.int/web/gaia.

2 Relativistic astrometric sky modeling

Having a control on the error budget at the level of $\mu$as for Gaia is even more critical if one considers that the solar system generates perturbations of the order of accuracy of the measurements. This turns out to trace back the direction of light to the position of the star from within the ever-present and ever-changing gravitational fields of our solar system. Consequently, also the retarded time terms due to the varying gravitational fields of the bodies need to be taken into account, namely the time when the gravitational field of the source actually began to propagate along the light cone. The major effects are the deflections of light due to the planets: already at the first post newtonian approximation they produce overlapping contributions up to the order of several $\mu$as; the contribution amounts just 1 $\mu$as, for example, at 180 deg from the limb of the Sun and at 90 deg from that one of Jupiter.

Therefore, achieving high astrometric accuracy translates into a fully self consistent relativistic model suitable to describe correctly the observables.

Thanks to the need of using General Relativity (GR) for Gaia, nowadays there exist different ways to model an astrometric observable. Their availability is required in order to consolidate the results. From the experimental point of view, in fact, relativistic astrometry opens a largely uncharted territory and it is of capital importance to allow the existences of different and cross-checked models which exploit different solutions to interpret the same experimental data. In this regard, inside the Data and Processing Analysis Consortium (DPAC)
constituted for the Gaia data reduction, two models are considered: i) GREM (Gaia RElativistic Model, baselined for the Astrometric Global Iterative Solution for Gaia (AGIS), and ii) RAMOD (Relativistic Astrometric MODels) implemented in the Global Sphere Reconstruction (GSR) of the Astrometric Verification Unit (AVU) at the Italian data center (DPCT), the only center, together with the DPC of Madrid, able to perform the calibration of positions, parallaxes and proper motions of the Gaia data.

RAMOD stands originally for Relativistic Astrometric MODel, conceived to solve the inverse ray-tracing problem in a general relativistic framework not constrained by a priori approximation \(^1, 2\). RAMOD is, actually, a family of models of increasing intrinsic accuracy all based on the measurement protocol in GR \(^3\), where light propagation is expressed in a general relativistic context, not necessarily applied only to astrometry. RAMOD can be adapted to many different observers settings. The solutions interface numerical \(^4\) and analytical relativity \(^5\).

Since both models are used for the Gaia data reduction, any inconsistency in the relativistic model(s) would invalidate the quality and reliability of the scientific outputs. Indeed, the main Solar System curvature perturbation amounts approximately to 100 microarcsecond, which will cause the individual parallaxes to fast degrade beyond 1 kpc, while completely invalidating the most accurate calibration of, e.g., the primary distance calibrators. This alone is sufficient reason for allowing the existence and making a theoretical comparison of different approaches a necessity.

3 All-sky GR testing with Gaia

The relativistic observable takes into account the measured abscissa along the scanning direction. In principle, once determined the local-line-of sight according to the RAMOD solutions and defined an appropriate relativistic attitude, each observation is a function of the Astrometric, Attitude, Instrument, and Global parameters which are accumulated in a large system of linearized observation equations in the case of GSR. Direct solution, no block-adjustment, of such a system via an iterative method provides estimates of variances. The dependence on the PPN parameter \(\gamma\) - which measures the amount of curvature produced by unit rest mass, equal to one in GR - gives the estimation of such a parameter as a by-product of the sphere reconstruction. Then, given the suit-
able relativistic models for analyzing the data, a mission like Gaia, repeatedly observing over 5 years millions of bright and stable stars uniformly scattered across the sky to a precision of 10-20 micro-arcsecond, will constitute by far the largest and most thorough astronomical experiment in testing GR ever attempted since its formulation (one century ago), possibly with the sensitivity for testing the dilaton-runaway scenario 6). Gravity theories alternative to GR require the existence of this scalar field and predict it fades with time, so that this residue would manifest itself through very small deviations from Einstein's GR in the weak field regime. Very accurate global astrometry is a very powerful and independent tool to unveil the presence of this scalar field, providing even available scenarios without dark components. Current simulations with the present configuration of Gaia suggest a final estimate of $|\gamma - 1|$ at the $10^{-6}$ level of accuracy, mainly due to a trade-off in the measurement performances between the faint and the bright end of the stellar sample.

While global tests will be done toward the mission’s end, when most of the observations will be collected, differential experiments, exploiting the precision of the elementary measurements, can be implemented also in the form of repeated Eddington-like experiments by comparing the evolution of angular distances in bright stellar asterisms consecutively observed by the satellite within a few planet’s radii from the limb of a giant planet like Jupiter. Results based on simulated observations of actual compound observed fields near Jupiter’s orbit - against a selected reference frame of fiducial stars for different scanning directions - prove Gaia’s ability to test the light deflection due to Jupiter’s quadrupole, predicted by GR and yet to be detected, with opportunities quite early (February 2017).

Gaia accurate space-phase structure can fully probe the Milky Way outer halo (i.e. mass content and distribution) and compare the prediction of ΛCDM models in situ. The aim is to search for new kinematic streams in the local halo and redefine membership of known streams. Cold Dark Matter (CDM) models predict that structures grow by hierarchical merging, mainly driven by dynamical friction and tidal disruption, leaving streams and substructures as relics of this process considered as tracers of the distribution of dark matter. Simulations predict the presence of hundreds of streams in the solar vicinity. However, although several groups of halo stars originating from common progenitor satellites have already been identified within a few kpc of the Sun, their
small velocity dispersion inside the streams requires a very high precision on 3D-velocity (about 5 km/s) to unambiguously separate them from the field. Recent simulations taking into account dynamical friction with the addition of the Gaia errors, show the possibility to detail the localization of the different streams in the halo structure 7).

Finally, the proper motions measured by Gaia have the potential to further confirm the Galactic warp 8).

4 Conclusion

All the goals of Gaia will not be achieved without a correct implementation of General Relativity in the data processing and analysis.

Gaia will not only greatly enhance our knowledge of the Galactic structure, but it will also provide precise information allowing astronomers to frame a much more detailed kinematical picture of our Galaxy than what presently available. A 6-dimensional accurate reconstruction of the individual stars across a large portion of the Milky Way necessarily needs extremely accurate astrometric observations modeled within a fully, comparably accurate, relativistic framework. Once a relativistic model for the data reduction has been implemented, any subsequent scientific exploitation should be consistent with the precepts of the theory underlying such a model. Any discrepancy between the relativistic models, if it can not be attributed to errors of different nature, will mean either a limit in the modeling/interpretation - that a correct application of GR should fix - and therefore a validation of GR, or, maybe, a clue that we need to refine our approach to GR. Moreover, given the number of celestial objects (a real Galilean method applied on the sky!) and directions involved (the whole celestial sphere!), the realization of the relativistic celestial sphere is not only a scientific validation of the absolute parallax and proper motions obtained with Gaia. Reaching 10-20 $\mu$as accuracy on individual parallax and annual proper motions for bright stars ($V < 16$) is also the key possibly to perform the largest GR experiment ever attempted from space with astrometric methods (since 1919).

And beyond the micro-arcsecond? Gaia represents only a ground step, increasing the level of accuracy requires to refine consistently the metric of the solar system, the solutions for the null geodesic, the observables, the attitude, and so on..
Therefore, the astronomers need to be ready to exploit all of the scientific potential of the local measurements entangled to the varying gravitational fields from within the Solar System and to maximize its impact. After Gaia, Astrometry becomes part of fundamental physics and, in particular, in that of gravitation.

Acknowledgments

This contribution was supported by ASI under contract to INAF 2014-025-R.1.2015 (Gaia Mission - The Italian Participation to DPAC - Operational Phase).

References

THE LARGE SCALE STRUCTURE PATH TO THE HIGH ENERGY UNIVERSE

Enzo Branchini
Department of Mathematics and Physics. University Roma Tre. Rome, Italy.
INFN. Sezione di Roma Tre. Rome. Italy.

Abstract

We illustrate the importance of the Large Scale Structure of the Universe beyond the usual perimeter of observational cosmology by means of two case studies, traditionally investigated within the framework of high energy astrophysics: the nature of the unresolved extragalactic $\gamma$-ray background and the search for the missing baryons in the intergalactic medium.

1 Introduction

The large scale structure [LSS] of the Universe is emerging as one of the most effective probes to investigate the nature of Dark Matter [DM] and that of Dark Energy. As these are probably the most important open problems in Cosmology and Fundamental Physics a number of large galaxy surveys are being designed to trace the LSS over increasingly large volumes of the Universe across a wide

240
range of epochs. These datasets will contain a tremendous amount of information that will impact many areas of astrophysics beyond observational cosmology, including high energy phenomena.

In this work we briefly illustrate how the LSS can be exploited to address two outstanding problems: (i) the nature of the diffuse γ-ray background and (ii) the missing baryons problem. We shall show and discuss the most recent results based on the available data and highlight the impact that next generation surveys will have on both issues.

The layout of the paper is as follows. In Section 2 we use the cross-correlation between the LSS and the unresolved γ-ray background to investigate the origin of the latter. In Section 3 we show that the filamentary structures in the LSS is an effective signposts for the elusive Warm Hot Intergalactic Medium [WHIM], i.e. the likely reservoir of the atoms that are missing from the baryon budget. We summarise the main results and consideration in Section 4.

2 The nature of the γ-ray background and Dark Matter.

Our understanding of the extragalactic γ-ray sky has improved dramatically thanks to the Large Area Telescope (LAT) on board of the Fermi satellite. We now know that most of the resolved extragalactic sources are blazars. What we still don’t know is whether these sources also account for the unresolved Diffuse Gamma-Ray Background (DGRB) \(^2\). The analysis of its energy spectrum has shown that blazars, misaligned AGNs and star forming galaxies [SFGs] can indeed explain a large fraction of the DGRB. However, current data neither clarify the relative contributions of these sources nor rule out the possibility that some other type of γ-ray source, like annihilating (or decaying) DM particles contribute to the DGRB. \(^1\), \(^4\)

Additional, independent constraints can be obtained by studying the angular correlation properties of the DGRB. Unfortunately, the standard autocorrelation analysis so successfully applied to the microwave background, is not very effective here. The reason is that uncertainties in the Galactic γ-ray emission models, do not guarantee an accurate subtraction of the Galactic foreground so that residuals can generate a spurious correlation signal. An effective way to overcome this problem is to cross-correlate the DGRB with any catalogue of extragalactic objects that trace the same LSS in which the extragalactic γ-ray sources reside and do not correlate with the local Galactic
This is the idea behind the suite of analyses that we have performed, in which we cross-correlated the Fermi-LAT DGRB maps with several different catalogs of extragalactic sources spanning different redshift ranges \(^6\).

More specifically, we have considered the 5-year Fermi DGRB maps obtained after subtracting the contribution of resolved \(\gamma\)-ray sources, Galactic diffuse emission generated in the disk and off the Galactic plane in the Fermi bubble and Loop I structures. Events have been divided in three energy bins with \(E > 0.5\), \(> 1\) and \(> 10\) GeV and maps have been further cleaned by removing multipoles with \(\ell > 10\). Then we considered maps of discrete extragalactic sources obtained from five different datasets: (i) the SDSS DR6 quasar catalog, (ii) the 2MASS galaxy catalogue, (iii) the radio sources of the NVSS catalogue, (iv) the luminous red galaxies in the SDSS DR8 catalog and (v) the main galaxy sample from the SDSS DR8 catalogue. Finally, we computed the 2-point cross-correlation function between the objects in each of the catalogues and the flux in the DGRB map. The typical result, plotted in Fig. 1, is that of a positive cross-correlation signal on angular scales below \(1^\circ\).

To understand the nature of this cross-correlation we compared these measurements with the following model cross-angular power spectrum

\[
C_{\ell}^{(\gamma g)}(\chi) = \int \frac{d\chi}{\chi^2} W_{\gamma}(\chi) W_g(\chi) P_{\gamma g}(k = \ell/\chi, \chi),
\]

where \(C_{\ell}\) is the amplitude of the power spectrum at the multipole \(\ell\), \(\chi\) is the radial co-moving distance, \(P_{\gamma g}(k)\) is the 3D cross-power spectrum between \(\gamma\)-ray emitters and the discrete source (e.g. 2MASS galaxies) and \(W(\chi)\) is the so-called window function that characterizes the distribution of objects and \(\gamma\)-ray emitters along the line of sight. To model the 3D spectrum we used the so-called halo model. \(P_{\gamma g}(k)\) encodes the information on the relative clustering between LSS tracers and \(\gamma\)-ray emitters. \(W_{\gamma}\) represents the \(\gamma\)-ray intensity along the
line of sight and depends on the intrinsic properties and redshift distribution of different types of γ-ray emitters: blazars, misaligned AGNs, SFGs and more exotic sources like DM particles that can annihilate or decay into γ–photons. $W_{\gamma}(\chi)$ weights the contribution of the discrete sources at various redshifts and, therefore, depends on the observed redshift distribution of the objects in each of the catalogues listed above.

Figure 2: 95% C.L. upper limits on the DM annihilation rate as a function of its mass for different final states ($b\bar{b}$, $\mu^+\mu^-$, $\tau^+\tau^-$, $W^+W^-$). The more conservative estimate (black curve) is obtained by assuming that only the DM contributes to the DGRB. The more realistic case of a DGRB contributed by all plausible astrophysical sources is shown by a red curve.

Eq. 1 reveals the important tomographic aspect of the cross-correlation analysis. Different γ-ray emitters peaks their emission at different redshifts so that for each type of γ-ray source $W_{\gamma}$ is significantly different from zero in a different redshift-range. As a consequence, one check whether a particular emitter does contribute to the DGRB by choosing an LSS-tracer with a window function $W_g$ that overlap with that of the potential emitter. The possibility of measuring the cross-correlation in different energy bins, further sharpens the possibility to disentangle the various contributions.

Comparing the model with the measured cross-correlation signal allows one to constrain the contribution of the known astrophysical γ-ray sources to
the EGB as well as the DM properties. For example one can constrain the mass and the annihilation cross-section for annihilation, as shown in Fig. 3. These constraints are consistent and competitive with those obtained by other indirect DM detection strategies targeting local dwarf galaxies of the Galactic center 7).

3 Searching for the missing baryons.

In the current framework of LSS formation, coherent structures build up over time forming the filaments of the cosmic web. At $z < 1$ their density contrast is sufficiently large to shock-heat the baryon component to temperatures $10^5$–$10^7$ K, forming the so called WHIM. The experimental evidence for such component is scarce and its detection and characterisation is a major target of several, ongoing observational campaigns. The expected association between the WHIM and the LSS filaments has now some experimental support from X-ray observations of the outskirts of galaxy clusters, i.e. the nodes of the cosmic web connecting the filaments 3). It is therefore clear that filamentary structures in the LSS can be used as effective signpost for the WHIM detection.

The volumes probed by current galaxy redshift survey is already large enough to trace the filamentary structures in the galaxy distribution. A typical example is illustrated in Fig. in which we show the filaments detected in the distribution of galaxies in VIPERS, a recently completed galaxy redshift survey of $\sim 90,000$ objects at $z \sim 0.8$.

We have proposed a new approach that uses galaxy luminosity density as a tracer of the WHIM. Using hydrodynamical simulations we have shown that the large scatter in the already known correlation between WHIM gas overdensity, $\delta_{\text{whim}}$, and galaxy luminosity over-density, $\delta_{\text{LD}}$, becomes significantly tighter and consistent with linear ($\delta_{\text{whim}} = 0.7 \pm 0.1 \times \delta_{\text{LD}}^{0.9\pm0.2}$) when is restricted to filaments 5).

This confirms that filaments in the galaxy distribution are preferential location for the WHIM. To test this hypothesis we considered the filaments along the line of sight to the blazar H2356-309 and used the measured galaxy luminosity density to predict position and density of the WHIM. We found evidence for the WHIM in correspondence of the Sculptor Wall and Pisces-Cetus superclusters in agreement with the redshifts and column densities of the tentative WHIM detections already obtained.
4 Conclusions

In this paper we have worked out two examples that illustrate the importance of the LSS as tool not only for cosmology but also for high energy astrophysics. In the first example we have shown that LSS can be used to investigate the nature of the extragalactic DGRB. Using wide, redshift surveys of extragalactic objects to trace the LSS at different redshifts it has been possible to investigate the nature, abundance and redshift distribution of the unresolved sources that contribute to the DGRB, including DM. In fact, the constraints on the mass and annihilation cross section (or decay time) of the DM particle that one derives by cross correlating the LSS with the DGRB maps are competitive with the best indirect DM detection techniques proposed so far. Yet, current catalogues allows one to span a wide redshift range but with a limited resolu-
tion. Next generation surveys like Euclid, DESI or SKA will be large enough to significantly increase the number of redshift bins and, consequently, improve the quality of the constraints.

The second example uses the LSS to locate the elusive WHIM. We have shown that once the filamentary structures in the spatial distribution of galaxies are traced, then one can use the local galaxy luminosity to predict the location, the density of the WHIM gas and, consequently, the probability of its detection with X-ray observations. Also in this case, the advent of the next generation redshift surveys will greatly enhance our ability to detect and, eventually, characterize the WHIM in the X-ray band, both in emission and in absorption, with the planned Athena satellite mission.

References

THE FIRST STARS IN THE UNIVERSE

Umberto Maio
INAF – Osservatorio Astronomico di Trieste, 34143 Trieste (Italy)
Leibniz Institute for Astrophysics, 14482 Potsdam (Germany)

Abstract
The basic processes of the formation of the first stars in the primordial Universe are outlined and the implications for cosmological structure formation discussed. By employing theoretical and numerical models of cosmic structure evolution embedded within N-body hydrodynamical chemistry simulations, predictions for the production of the first heavy elements in the Universe are given. These results are then compared against measured data of UV luminosities and metal abundances in different kinds of observations in order to draw conclusions on the chemical and thermal state of the cosmic medium at different cosmological epochs.

1 Introduction
Cosmic structures originate from the growth of matter perturbations at early times in an expanding Universe. Baryonic objects form from the in-fall and
cooling of gas into the dark-matter potential wells since very high redshift ($z$). A star forming gas ‘cloud’ can form if radiative losses are sufficient to make the gas condense and fragment. At high $z$, gas cooling is dominated by H, He and H-based molecules, like H$_2$ and HD. After pollution from freshly formed stars, metals also contribute to gas cooling.

Primordial epochs are important for our understanding of the formation of the very first stars and (proto-)galaxies. These are the sites which witness the occurrence of the first heavy elements in the Universe and their spreading in the surrounding regions. Unfortunately, there is still a lack of definitive knowledge about the features of the first stars, such as their masses, metal yields and luminosities, so that it is difficult to give exact predictions for cosmic pollution in different environments. Furthermore, the role of molecules and metals is crucial in dictating the transition from the formation of stars in pristine environments (population III, Pop III, regime) to a regime which is dominated by cooling due to heavy elements (population II-I, Pop II-I, regime).

To assess these problems in a quantitative way, deep studies of the hydrodynamical and chemical properties of cosmic medium are required. Here we will present and summarise results from numerical simulations taking into account the most relevant physical processes to draw conclusions on primordial star formation and on the impacts of the first stars on the following generations of cosmological structures.

2 Numerical simulations

For a consistent picture of primordial structure formation it is important to consider a number of processes to be included into N-body hydrodynamical numerical calculations. We use the cosmological parallel code Gadget$^2$ and extend it to include non-equilibrium molecular chemistry, cooling, stellar evolution and metal pollution$^1$, $^4$, $^6$, $^13$.

Molecule formation and evolution determine the first collapsing events through H$_2$ and HD cooling and lead Pop III star formation.

The occurrence of heavy elements from Pop III stars increase the efficiency of gas cooling capabilities, hence metal chemistry become important for the formation of second generation stars.

Finally, stellar evolution processes dictate the amounts of photons and heavy elements ejected by stars with different masses and lifetimes via super-
nova (SN) explosions or asymptotic giant branch (AGB) winds.

The transition from the Pop III to the Pop II-I regime is accounted for according to the underlying metallicity ($Z$) of the collapsing gas. If star forming gas is pristine or its metallicity is below a threshold limit of $10^{-4}$ the solar value ($Z_{\odot}$) the actual stellar population is assumed to be Pop III with a stellar mass range [100, 500] M$_{\odot}$ and an initial mass function (IMF) having a slope of $-2.35$. If local metallicities are $Z > 10^{-4} Z_{\odot}$ then a Pop II-I regime is assumed with a Salpeter IMF over [0.1, 100] M$_{\odot}$. We stress that the different mass ranges between Pop III and Pop II-I generations imply also different metal yields and lifetimes with obvious consequences on the patterns and timescales of cosmological metal enrichment.

From simulated outputs it is possible to extract the main properties of dark-matter haloes, cosmic gas and evolving galaxies and to explore their correlation properties.

The background cosmology adopted is a model with cold dark matter and cosmological constant $\Lambda$ ($\Lambda$CDM), with expansion parameter $H_0 = 70 \text{ km/s/Mpc}$, present total matter density parameter $\Omega_0 = 0.3$, baryon density parameter $\Omega_{0,b} = 0.04$, $\Lambda$ density parameter $\Omega_{0,\Lambda} = 0.7$, mass variance within 8-Mpc/$h$ radius sphere $\sigma_8 = 0.9$ and spectral index $n = 1$.

3 Results

3.1 The first Gyr

The contribution of the first Pop III stars to the cosmic star formation rate (SFR) density is shown in Fig. 5 by Maio et al. 2010. Given the uncertainties on the critical metallicity for the transition, values of $10^{-6}$, $10^{-5}$, $10^{-4}$, $10^{-3}$ $Z_{\odot}$ have been tested, as well (see also Fig. 6 and 7 therein for further parameter dependences). The contribution from Pop III regime drops dramatically from about unity at $z > 16$ down to $\sim 10^{-3}$ at $z < 11$ quite independently from the exact critical metallicity adopted. This means that at such early epochs ($< 0.5 \text{ Gyr}$) the role of pristine star formation is already marginal and the Universe is already enriched in a sensible way. The total cosmic star formation results dominated by polluted Pop II-I haloes, despite their number fraction is not extremely large (below 10 per cent), as shown by Biffi & Maio (2013) in the top and bottom panel of their Fig. 3.
The picture remains qualitatively similar even after a broader parameter space exploration. It emerges a primordial Universe which is rapidly enriched and dominated by Pop II-I stellar population after a relatively short time from the onset of cosmic star formation.

3.2 Theory and data

Theoretical results can be compared against available data for luminosity functions (LFs) and specific star formation rates (sSFR).

The LF provides the fraction of objects with a given luminosity (or magnitude) observed at any given redshift. In the recent years many data for the high-z Universe have become available and they are precious to probe the early epochs at \( z > 5 \). Results on LFs at such primordial epochs date back to Salvaterra et al. (2013) where it was found a reasonable agreement between theory and observations for \( z > 6 \) (Fig. 1). At \( z \sim 6 \) or lower, the observed LF usually results obscured by effects due to the growth of dust grains in the interstellar medium which need to be taken into account.

Data for the sSFR are more uncertain and their scatter is up to one dex, therefore it is quite difficult to disentangle different theoretical models. Nevertheless, the typical trend of increasing sSFR for increasing \( z \) displayed by data is broadly in line with expectations (Fig. 12).

3.3 Implications for high-z GRBs

Primordial gamma-ray bursts (GRBs) are thought to be originated by the collapse of the first massive stars. These events can form a black hole and can be accompanied by jets. Since massive stars are short-lived, early GRBs basically trace star formation episodes and provide information about the local environment.

A number of theoretical studies in this field have shown the potential of GRBs to infer the properties of their host galaxies and place constraints on the primordial Pop III regime. It turns out that (Salvaterra et al.; 2013) typical stellar masses of primordial GRB hosts peak at \( \sim 10^7 \, M_\odot \) and corresponding SFR values range between \( \sim 0.01 \) and \( 0.1 \, M_\odot/yr \), giving \( sSFR \sim 5 - 10 \, \text{Gyr}^{-1} \). The most popular magnitudes for UV luminosities are in the range \([-20, -12]\), while expected metallicities peak around \( Z \sim 10^{-1.5} \, Z_\odot \).

These findings are in agreement with available data at \( z > 5 \), as shown in...
Fig. 5, and support the idea that small primordial proto-galaxies produce most of the ionising photons at early times (Fig. 3-4 therein).

Further analyses have stressed the most suited targets for Pop III searches at higher $z$. In particular, Ma et al. (2015) have explored the indirect Pop III signatures imprinted in enriched Pop II-I hosts in order to disentangle the two stellar populations. Thus, Pop II-I star forming galaxies pre-enriched by very massive Pop III stars appear to have typical metallicities $Z \lesssim 10^{-2.8} Z_\odot$ and peculiar abundance ratios, such as $[\text{Si}/\text{O}] < -0.6$, $[\text{S}/\text{O}] < -0.6$, $[\text{C}/\text{O}] > -0.4$. Obviously, such criteria depend on the assumed mass of primordial stars, hence, spectral data can give additional hints on the stellar structure of such objects.

4 Conclusions and perspectives

The results presented here have been obtained by ad hoc numerical simulations including N-body and hydrodynamical calculations, atomic and molecular chemistry, star formation, stellar evolution and feedback effects.

First star formation episodes are very ‘bursty’ and efficient metal spreading leads to a rapid transition from Pop III to Pop II-I regime. Primordial Pop III stars dominate cosmic star formation for a relatively short time and only a residual fraction of SFR survives at late times. This makes direct searches of Pop III stars statistically very difficult and implies the need to rely on additional indirect methods to shade light on primordial star forming events.

Among the possible processes that could affect the final results it is worth mentioning the possible existence of primordial supersonic flows originated at decoupling. They could induce homogeneous gas streaming motions on Mpc scales and have impacts on early structure formation, reionization and the lowest-mass dwarf galaxies.

Results, though, are not very sensitive to the assumed values for the critical metallicity, Pop III metal yields or IMF slope.

In these kinds of studies the commonly assumed background framework is the ΛCDM model, however different scenarios are possible. Alternative non-Gaussian models and dark-energy quintessence models have been tested, as well. Despite the details in these cases might change the overall trends are generally recovered.
Also the elementary nature of dark matter might play a role leaving room for warm dark matter in place of cold dark matter [28]. The dumping effects of warm dark matter at small scales seem to be evident in terms of dark-matter structure distributions and shapes, however the implications on the first stars and reionization are less trivial.

Some final considerations are briefly devoted to future perspectives to observe the first stars and galaxies in the next decades. The upcoming James Webb Space Telescope (JWST)\(^1\) is under construction and has been designed to detect faint sources at early times and to study reionization and early galaxy formation. The Square Kilometer Array (SKA)\(^2\) radio telescope will be built in the South hemisphere and will give scientists the possibility to study a large number of topics, among which hydrogen distribution, galaxy formation and radio emission in the first billion years [26]. The Athena\(^3\) space mission will target the hot and energetic Universe by employing high-resolution X-ray spectroscopy. Its capabilities will be exploited to investigate also GRB X-ray afterglows up to redshift \(z \sim 6 - 10\).

In spite of the huge costs for the technical development of such experiments, the scientific return is supposed to be unprecedented.

References


\(^{1}\)http://www.jwst.nasa.gov
\(^{2}\)https://www.skatelescope.org
\(^{3}\)http://www.the-athena-x-ray-observatory.eu


27. Maio, U., Ciardi, B., & Koopmans, L., Advancing Astrophysics with the Square Kilometre Array (AASKA14), **9** (2015)


The Milky Way Hot Baryons and their Peculiar Density Distribution: a Relic of Nuclear Activity

F. Nicastro\textsuperscript{1,2}, F. Senatore\textsuperscript{1}, Y. Krongold\textsuperscript{3}, S. Mathur\textsuperscript{4}, M. Elvis\textsuperscript{2}
\textsuperscript{1} INAF - Osservatorio Astronomico di Roma, Via Frascati, 33, 00078 Monte Porzio Catone (RM), Italy
\textsuperscript{2} Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA
\textsuperscript{3} Instituto de Astronomia - UNAM, Mexico City (DF), Mexico
\textsuperscript{4} Ohio State University, Columbus OH, USA

Abstract

We know that our Galaxy is permeated by tenuous, hot, metal-rich gas. However much remains unknown about its origin, the portion of the Galaxy that it permeates, its total mass, as any role it may play in regulating activity in the Galaxy. In a Letter currently in the press with the ApJ, we show that this hot gas permeates both the disk of the Galaxy and a large spherical volume, centered on the Galactic nucleus, and extending out to distances of at least 60-200 kpc from the center. This gas displays a peculiar density distribution that peaks about 6 kpc from the Galaxy’s center, likely witnessing a period of strong activity of the central supermassive black hole of the Milky Way that occurred 6 Myrs ago. With our study we are also able to update the total baryonic mass of the Galaxy to $M_b = (0.8 - 4.0) \times 10^{11} \, M_\odot$, sufficient to close the Galaxy’s baryon census.
1 Introduction

The visible baryonic mass of the Milky Way amounts to $M_{\text{Obs}}^b \simeq 0.65 \times 10^{11} M_\odot$ \(^1\). The total baryonic plus dark matter mass of our Galaxy, is instead $M_{\text{Tot}} \simeq (1 - 2) \times 10^{12} M_\odot$ \(^2\). This, assuming a universal baryon fraction of $f_b = 0.157$ \(^3\), implies a total baryonic mass of $M_{\text{Pred}}^b \simeq (1.6 - 3.2) \times 10^{11} M_\odot$, between 2.5 and 5 times larger than observed. A large fraction of the baryonic mass of our Galaxy is thus currently eluding detection.

This missing-baryon problem is not a monopoly of the Milky Way: most of the galaxies in the local universe suffer a deficit of baryonic mass compared to their dynamical mass and the problem is more serious at smaller dynamical masses (e.g. \(^4\)), suggesting that lighter galaxies fail to retain larger fractions of their baryons. These baryons could be at least partly hiding under the form of tenuous hot ($\sim 10^6$ K) gas, heated up by recurring episodes of nuclear activity during the galaxy’s lifetime, such as bursts of star formation followed by powerful supernova explosions or accretion-powered ignitions of the central supermassive black hole, which may have powered energetic outflows that pushed material out to large distances from the Galaxy’s center.

Over the past several years, a number of experiments, as well as theoretical works, have attempted to gain insights into the location and mass of the hot medium in our own Galaxy. Our peripheral position in the Galaxy, at about 8.5 kpc from the Galaxy’s center and roughly in the Galaxy’s plane, gives us hope of solving the problem: once a physically motivated density profile is assumed for the hot absorbing medium, the observed column densities (as well as the other observables) will depend critically on the sky position (and distance, for Galactic background targets) of the sources towards which the column densities are measured. This consideration has recently motivated several studies, which have used available spectra of extragalactic targets, with no other selection criterion than being at high Galactic latitudes, to measure OVII column densities and compare them with physically motivated or simple phenomenological density profile models \(^5,6,7\). The results, however, are often contradictory, with estimated total masses of the million degree medium within a 1.2 virial-radius sphere (300 kpc) that strongly depend on the flatness of the assumed density profile, and range from a negligible $M_{\text{Hot}} \simeq 2.4 \times 10^9 M_\odot$ \(^5\) to a significant $M_{\text{Hot}} \simeq 10^{11} M_\odot$ \(^7\).

Here we present an experiment that settles the controversy by adopting a num-
ber of novel and rigourous data analysis techniques and sample selection criteria, and that has recently been published by the ApJ Letter \(^{(8)}\), hereinafter N16).

Throughout this contribution, we refer to all densities and masses in units of \((A_O/4.9 \times 10^{-4})^{-1} \times [Z/(0.3Z_\odot)]^{-1}(f_{OVII}/0.5)^{-1}\), where \(A_O\) is the relative abundance of oxygen compared to hydrogen, \(Z\) is the metallicity and \(f_{OVII}\) is the fraction of OVII relative to oxygen. For easy comparison to other works (e.g. \(^{(7)}\)), we compute hot baryon masses within a 1.2 virial-radius sphere. Errors on best-fitting parameters (and quantities derived from those) are provided at 90% confidence level for a number of interesting parameters equal to \((31-N_{dof})\), where \(N_{dof}\) is the number of degrees of freedom in the fit.

2 Data, Model Components and Procedures

2.1 LGL and HGL Samples

Our LGL+HGL sample differs from those previously used to perform analyses similar to ours (e.g.\(^{(5,9)}\)) in three important ways: (1) for the first time we use simultaneously HGL and LGL samples; (2) our two XMM-Newton Reflection Grating Spectrometer (RGS) samples are complete to a minimum Signal to Noise per Resolution Element in the continuum, SNRE>10 at 22 Å; (3) whenever possible (see below), we remove the degeneracy between column density and Doppler parameter of the instrumentally unresolved OVII lines, by performing a detailed curve of growth analysis (e.g. \(^{(10)}\)).

Our final HGL and LGL samples contain 18 and 13 lines of sight, respectively, leading to a well-defined SNRE-complete observed distribution of 31 OVII \(K\alpha\) EWs and sky positions (see Figure 1 in N16).

HGL OVII absorbers spread over more than an order of magnitude in column densities, from a minimum value of \(N_{OVII} = 0.8^{+1.2}_{-0.5} \times 10^{16}\) cm\(^{-2}\), to a maximum value of \(N_{OVII} = 33^{+480}_{-29} \times 10^{16}\) cm\(^{-2}\). The spread is less extreme for LGL absorbers that span a factor of about 5 in OVII column densities.

2.2 Functions and Fitting Procedure

We model the derived distribution of 31 column densities and sky positions with two most general families of density profiles, i.e.: spherically-symmetric (SS), where the only scale-length parameter is the core-radius \(R_c\) (exponential-
and β-profile models), and Cylindrically-Symmetric (CS), characterized by two different scale-length parameters, the coplanar core-radius $\rho_c$ and the vertical core-height $h_c$ (again, exponential- and β-profile models: see N16 for the analytical expressions adopted). For each functional form, we also allow for the inclusion of an additional parameter ($R_s$ for the SS profiles and $h_s$ for the CS profiles, both in kpc) allowing for a possible offset of the distributions from the Galaxy’s center (SS models) or plane (CS models).

2.3 Halo Extent and Masses

Given the one-dimensional nature of our observables, only a lower limit to the total extent of the volume containing the hot absorbing gas seen against HGL targets can be evaluated in our analyses. We evaluate this limit by stopping the line of sight integration at a line of sight distance $\xi$ where the relative difference between two consecutive values of the column density differs by less than 0.01% (much lower than the typical relative uncertainty on our column density measurements, which is of the order of $\approx 10\%$ in the best cases). Under the assumption of a centrally symmetric halo, the largest of these line of sight distances in the best-fitting HGL models, sets effectively a lower limit to the radial size of the halo, and so its baryon mass. Smaller halos are not allowed by the necessity to accumulate sufficient column density (and emission measure) along the thickest HGL lines of sight. On the other hand, larger halo sizes (and therefore baryon masses) are clearly possible, but not directly measurable through our observables.

2.4 Caveats on Parameter Degeneracy

For the same limitation intrinsic in the one-dimensionality of our observables, the parameters of our models are all degenerate, to some extent. For exponential profiles, the problem is negligible and only a moderate degeneracy is present between the peak density $n_0$ and the scale-distance parameters $R_c$ or $\rho_c$ and $h_c$. For β-like profiles, instead, the scale-distance parameters are often strongly degenerate with the flatness index $\beta$ and, when this happens, is impossible to discriminate statistically between very steep and compact (exponential) or flat and extended profiles. In these cases exponential profiles or β-like profiles with either the flatness index $\beta$ or radial scale distance $R_c$ frozen to some physically motivated value, provides the only non-degenerate solutions
(see N16 Tables and details). In all cases, however, the simultaneous model-
ing of LGLs and HGLs, together with the presence of a radial offset in the
distribution (i.e. best-fitting $R_s > 0$), tend to break the degeneracy between
scale-distance and the index $\beta$. When this happens, the shape of the density
profile can be determined and all model parameters are generally robustly con-
strained to physically reasonable best-fitting values (models A and B in table
2 of N16), and so are the derived minimum extent and mass of the halo.

3 Modeling and Results

First we modeled the HGL and LGL separately. The 18 HGL absorbers are
equally well fitted by both SS and CS models, and the additional degree of com-
plexity introduced by CS models over SS models is not statistically required.
Density profiles are generally steep and, interestingly, the best-fitting profile is
exponential and has an offset-radius $R_s = 5.4^{+0.6}_{-0.4}$ kpc (see details in N16). This
shift in radius would indicate that the hot baryon density in the halo increases
radially from the Galaxy’s center up to its peak value at 5.4 kpc, and then
decreases monotonically towards the virial radius. The implied baryonic mass
is unimportant $M_{Halo} = 3.3^{+4.1}_{-1.4} \times 10^9 \, \text{M}_\odot$ (in full agreement with the mass
derivable from the best-fitting parameters of the spherical-saturated model of
(5); see their table 2).

Unlike the HGL absorbers, for the 13 members of our LGL sample a flat-
ten disk-like CS density profile is statistically greatly preferred to a SS profile
[$\chi^2_{\text{flat}} = 12.8$ for 10 degrees of freedom (dof), versus $\chi^2_{\text{sph}}(dof) = 22.5(11)$].
The LGL absorbers are clearly tracing a disk-like distribution in the Galaxy’s
plane, with best-fitting radial and height scale lengths in excellent agreement
with those of the stellar disk of the Milky Way (11) (see details in N16). The
mass of the hot gas in the Galactic disk is only $M_{\text{Disk}} = 1.4^{+1.1}_{-0.6} \times 10^8 \, \text{M}_\odot$, of
the order of that of the other gaseous components of the disk (11).

The models that best-fit separately our HGL (halo) and LGL (disk) ab-
сорbers, are very different in both their central densities and profiles (Tab. 1 in
N16), and neither of the two can adequately model the column-density distri-
bution of the other. Either a compromise single-component model is needed,
or the two components must be physically distinct.

We then proceeded to model simultaneously and self-consistently all 31
HGL and LGL lines of sight. We tried two alternative families of functions:
(A) a single-component set of models, with all parameters free to vary and the normalizing density peaking at $R = R_s$ (A-type models, hereinafter), and (B) a 2-component set of models in which a parameter-varying SS component with $n = 0$ for $R < R_s$ and $n = n(R)$ for $R \geq R_s$, is added to a flattened disk-component with parameters frozen to the LGL best-fitting values (B-type models, hereinafter). Both sets of models can provide statistically acceptable fits (see Table 2 in N16). In both cases offset radii $R_s > 0$ are statistically preferred (compare A with M4 and B with M3 in Fig. 1a,b, and see N16 for details).

Our two best-fitting models A and B have offset radii $R_s = 5.6^{+0.6}_{-0.6}$ kpc and $R_s = 6.7^{+0.9}_{-1.3}$ kpc, consistent with each other and with the best-fitting value found by fitting the HGL sample only. In model A, $R_s = 0$ is ruled out at a 4-interesting-parameter statistical significance of 14.9σ. Similarly, for our alternative best-fitting model B, $R_s = 0$ is excluded at a 4-interesting-parameter statistical significance of 6.0σ.

Fig. 1a and 1b show that Model A (left 2 panels of Fig. 1a) is able to better reproduce the observed spread of OVII column densities along HGL lines of sight, compared to model B (left 2 panels of Fig. 1b), which however (by construction) reproduces better the observed LGL columns (see N16 for details). By comparison, the alternative two models with $R_s$ frozen to zero, M3 and M4, clearly reproduce HGL and LGL columns less well than the respective models B and A. The actual solution lies probably in between models A and B.

Our best-fitting models A and, to a lesser extent, B also reproduce well the Emission Measure at high Galactic latitudes and towards the Galactic center (see N16 for details).

From our best-fitting models we derive total hot baryon masses in the ranges $M_{Hot}(A) = 0.2^{+0.3}_{-0.1} \times 10^{11} M_\odot$ and $M_{Hot}(B) = 1.3^{+2.1}_{-0.7} \times 10^{11} M_\odot$. These masses are $> 10$ times larger than those obtained by fitting the HGL sample only. This is due to the flatness of the best-fitting density profiles (see N16 for details). These flat profiles imply minimum sizes of the halo of $> 60$ kpc and $> 200$ kpc for Model A and B, respectively.

Adding the hot baryon mass to the visible mass of the Milky Way, gives a total baryonic mass in the range $M_b = (0.8 - 4.0) \times 10^{11} M_\odot$, sufficient to close the Galaxy’s baryon census.
3.1 Discussion

Our analysis indicates not only (1) that both the Galactic plane and the halo are permeated by OVII-traced million degree gas, but also (2) that the amount of OVII-bearing gas in the halo is sufficient to close the Galaxy’s baryon census and (3) that a vast, \( \sim 6 \) kpc radius, spherically-symmetric central region of the Milky Way above and below the 0.16 kpc thick plane, has either been emptied of hot gas (Model B) or the density of this gas within the cavity has a peculiar profile, increasing from the center up to a radius of \( \sim 6 \) kpc, and then decreasing with a typical halo density profile (Model A).

The large value of \( R_s \) in both the scenarios implied by Model A and Model B can be understood in terms of a radially expanding blast-wave or a
shock-front generated in the center of the Galaxy and traveling outwards, so
acting as a piston onto the ambient gas, and compressing the material at its
passage, while pushing it (or a fraction of it) outwards. The central black hole
of our Galaxy, could have played a fundamental role in this (e.g. (12,13,14)),
during a recent period of its activity. Faucher-Gigure & Quataert (2012) study
the property of galactic winds driven by active galactic nuclei, and show that
ergy-conserving outflows with initial velocity $v_{in} > 10000 \text{ km s}^{-1}$, can move
in the ambient medium producing shocked wind bubbles that expand at veloc-
ities of $v_s \approx 1000 \text{ km s}^{-1}$ into the host galaxy. If the observed OVII-bearing
bubble in our Galaxy is tracing one of such shocks generated by our central
supermassive black hole during a period of strong activity then, at a speed of
1000 km s$^{-1}$, the expanding shell would have taken 6 Myrs to reach its current
radius of 6 kpc. Interestingly, (6 ± 2) Myr is also the age estimated for the
two disks of young stars present in the central parsec of our Galaxy that are
thought to be a relic of a gaseous accretion disk that provided fuel for AGN-like
activity of our central black hole about 6 Myr ago (15,16).

4 Acknowledgements

F.N. acknowledges support from the INAF-PRIN grant 1.05.01.98.10.

5 References

References

THE PROBES AND SOURCES OF COSMIC REIONISATION

Francesco Haardt

DiSAT, Università degli Studi dell’Insubria, COMO

Abstract

The reionisation of the all-pervading intergalactic medium (IGM) is a landmark event in the cosmic history of structure formation. Still, despite much recent progress, a coherent description of the thermal state and ionisation degree of the IGM, the repository of most of the baryons across the history of the Universe, remains elusive. Most of our understanding of IGM physics, and its implication for galaxy formation and metal enrichment, depends critically on the properties of the cosmic ionising background. Over the last 15 years detailed models of cosmic reionisation have been proposed and adopted in cosmological studies. Yet, many uncertainties still exist, and background models are now facing hard challenges, when confronted to new observations both in the local Universe and at high redshifts. In this contribution, I review our current understanding of the reionisation era, and the many problems still open.
1 Introduction

Hydrogen, along with the pristine helium nuclei (plus traces of light metals) formed during the primordial nucleosynthesis, remained ionised until the temperature of the Universe dropped below few thousands degrees, $\simeq 350,000$ years after the Big bang. Then H and He recombined, and for most of the cosmic history the neutral intergalactic medium (IGM), the repository of most of the baryons of the Universe, evolved smoothly at a rather slow pace. But two notable exceptions exist: the reionisation of hydrogen and helium. Studies of the so-called Gunn-Peterson absorption in the spectra of distant quasars show that hydrogen was already highly ionised out to redshift $z \sim 6^{1, 2}$, a value recently confirmed by the latest CMB polarisation data $^{3}$. Helium, instead, shows signs of complete second reionisation at much later times, $z \sim 3, ^{4}$ and confirmed by several following studies $^{5}$, though different views have been recently proposed $^{6}$.

Modern observations of the IGM, via analysis of absorption lines in the spectra of distant quasars, have provided several probes and tests of the $\Lambda$CDM paradigm (e.g., a measurement of the power spectrum of matter fluctuations, upper limits to the neutrino masses, an independent measure of baryonic acoustic oscillations $^{7, 8, 9}$). However, despite much recent progress, a coherent description of the thermal state and ionisation degree of the IGM remains elusive. The intensity and spectrum of the cosmic ionising background (hereinafter UVB) is one of the most critically important, yet uncertain, astrophysical input parameters for cosmological simulations of the IGM and galaxy formation, for interpreting quasar absorption-line data, and for deriving information on the distribution of primordial baryons (traced by HI, HeI, HeII transitions) and of the nucleosynthetic products of star formation (CIII, CIV, SiIII, SiIV, OVI, etc.). Perhaps most importantly, the UVB is tied inextricably to the two reionisation events.

The reionisation of the all-pervading IGM, the last phase transition in the history of the Universe, has an astrophysical, rather than cosmological, origin. While it is generally assumed that the IGM is kept ionized by the integrated UV emission from quasars and star-forming galaxies, the relative contributions of these sources as a function of the cosmic epoch are poorly known. Because of the high ionisation threshold (54.4 eV) and small photoionisation cross-section of HeII, and of the rapid recombination rate of HeIII, the double ionisation of
helium is expected to be completed by hard UV-emitting quasars around the peak of their activity, at \( z \sim 2.5 - 3.5 \) \cite{10, 11, 12}, more than one billion years later than the reionisation of HI and HeI. At \( z > 3 \), the declining population of bright quasars is generally believed to make an increasingly small contribution to the radiation background at the Lyman limit (13.6 eV).

It was then suggested that massive stars in galactic or pre-galactic systems may have provided the additional HI ionising flux needed at early times \cite{13, 14, 15, 16, 17, 18, 19, 20, 21}. However, the leakage of Lyman continuum photons from bright galaxies seems to be very modest \cite{22}, and it has been therefore argued that dwarf galaxies (with virial mass below \( \sim 10^{9} M_\odot \)) may produce the dominant contribution to the HI ionising background \cite{23}. State-of-the-art galaxy formation models can match many diverse current observational constraints, both during and after the HI reionisation era, only if a fraction as large as \( \sim 10\% \) of the ionising radiation produced by massive stars can escape into the IGM \cite{24}.

Very recently, however it has been suggested \cite{25} that the low scattering opacity reported by Planck, the almost all-negative detections of Lyman continuum leakage from star forming galaxies, combined to the claim of a significant population of faint AGNs at \( 4 < z < 6.5 \) \cite{26}, all this may point toward a dominant role of QSOs in HI reionisation, contrary to general wisdom.

2 Open Issues

2.1 The Escape Fraction of Lyman continuum radiation

Today, our knowledge of the UVB still suffers of one main limitation concerning the amount of ionising radiation provided by star forming galaxies and QSOs. A simple estimate of the minimum amount of high energy photons needed to complete HI reionisation by \( z \sim 7 \) (one photon per baryon per recombination time \cite{27}), once converted into a minimum dark matter halo mass, would result in the requirement of vigorous star formation in dark matter haloes down to \( 10^{8} M_\odot \) \cite{28}. Similarly, analysing HST/COS data within a large collaboration, we have shown how at low redshift (\( z < 0.2 \)) five times more ionising photons than what predicted by successful reionisation models are required \cite{29}. Such high values of the low redshift UVB can be matched if the escape fraction of HI ionising radiation from galaxies (\( f_{\text{esc}} \)) is as large as \( \sim 10\% \).
The escape fraction of Lyman continuum radiation from star forming galaxies is the most important and less known parameter affecting any estimate of the UVB. Yet, a detailed assessment of the UVB adopting a physically motivated, luminosity-and-redshift dependent escape fraction is mandatory for the next generation of cosmological simulations, and quasar absorption line studies. Current results on $f_{\text{esc}}$ based on numerical simulations are somewhat ambiguous. It was argued that the escape fraction increases with increasing halo mass, in the range $10^{10} - 10^{11} M_{\odot}$, and found values of few percent at best \cite{30}. On the other hand other groups claimed, on the ground of SPH simulations with radiation post processing, an opposite trend, i.e., higher escape fraction in lower mass haloes \cite{31, 32}. Moreover, the numerical values of the resulting escape fractions are in strong disagreement, with generally much lower values.

Questions arise on the dependence of the results upon the different numerical resolutions attained in different simulations. Ultrahigh resolution simulations of dwarf galaxies with an adaptive mesh refinement (AMR) scheme with on-the-fly radiative transfer found a very high, time varying escape fraction, with a time average of $f_{\text{esc}} \approx 0.5$ increasing towards higher halo masses \cite{33}. Regardless the actual value found, the trend toward high masses could be partially attributed at the increase of star formation efficiency with galaxy mass. Note however that others recently showed that the early progenitors of todays dwarfs ($\sim 10^8 M_{\odot}$) possibly had a higher stellar component than their low-redshifts counterparts of similar size \cite{34}. Although different results in terms of $f_{\text{esc}}$ by different studies can be partially due to different numerical resolution employed, a larger role is probably played by the different treatments of the radiative transfer and sub-grid physics assumed by different analysis (such as the impact of stellar winds, supernova feedback, and so on, on gas temperature and ionisation state).

### 2.2 UVB fluctuations

Most studies (in particular, cosmological simulations) treat the ionising background as spatially uniform (but see \cite{35}). During reionisation itself such assumption is obviously wrong, because some regions are exposed to local strong ionising radiation, while others remain neutral with no nearby illuminating sources. Moreover, even within ionised regions, differences in the opacity of
the IGM along different lines of sight induce large spatial variations in the background. These fluctuations are expected to be both larger and more observationally relevant in the case of the helium-ionising background, because the IGM absorbs helium-ionising photons more strongly than hydrogen ionising photons, leading to shorter mean free path and larger fluctuations. Furthermore, quasars are comparatively much rarer than galaxies, and themselves quite variable, implying large random fluctuations in the background. These fluctuations persist even in the post-reionisation Universe, because of source clustering, stochastic fluctuations in the quasar number density, and the clumpy nature of the IGM.

While spatial variations in the HI-ionising background are expected to be small owing to the large mean free path of ionising photons $\sim 200 \text{ cMpc}^{36}$, recent studies indicate a mean free path of $\sim 30 - 50 \text{ cMpc}$ for helium-ionising photons at $z \sim 3^{37, 38, 39}$. Theoretical studies, both analytical $^{40, 41, 42}$, and numerical $^{43}$, predict order of magnitude fluctuations in the HeII ionisation rate. More recently, through a count-in-cell approach to estimate the distribution of UV background photons that included source clustering, it was shown that clustering can account for less than 30% of the variance of intensity fluctuations for a QSO correlation function not exceeding $\sim 15 \text{ cMpc}^{44}$. Moreover, very recently, it was shown how by adopting an homogenous UVB derived from 1-D RT calculations leads to an over-estimate of the HI reionisation redshift $^{35}$. In a work in progress, we include in a straightforward way the effect of patchy reionisation in 1-D UVB models based on the best currently known properties of both sources and sinks of ionising photons $^{47}$.

2.3 Measuring the UVB

A last issue that is worth mentioning concerns the actual measurement of the UVB. There are basically three different ways for probing the UVB. The so-called quasar proximity effect $^{45}$ and the low surface brightness emission from the outskirts of galactic discs $^{46}$ allow intrinsically difficult measurements of UVB because of significant uncertainties. Alternatively, a third route to UVB consists in forcing a match between the (more easily observed) Lyman-\(\alpha\) forest opacity and that from a theoretical model of the IGM whose residual neutral fraction depends upon the UVB. Such last method is largely employed in cosmological studies $^{29}$, still it does depend on few, and sometimes subtle,
underlying assumptions.

Recently, with a group of colleagues, we performed pilot direct observations of the edge-on galaxies UGC 7321, using the Multi Unit Spectroscopic Explorer (MUSE) integral field spectrograph at the Very Large Telescope \(^4\). By pointing at the tip of the HI disk, we aimed at mapping the diffuse H\(_\alpha\) emission arising from recombination of gas that is being photoionised by the local UVB. We detect spatially-extended emission visible at the location of the ionisation front inferred from deep radio HI observations. Through photoionisation calculations, we will constrain the HI photoionisation rate, obtaining a value for the UVB that should be considered an upper limit given that current data cannot directly rule out a contribution of local sources to the ionization budget. These pilot observations show the prospects of pinning down the intensity of the UVB in the era of large integral-field spectrographs at 8m telescopes.

### 3 Conclusions

In this brief essay, I reviewed our current understanding of the Epoch of Reionisation, the last phase transition the baryon component of the Universe experienced along its cosmic history. I mainly focused on few still open issues, such as the assessment of the ionising output from astrophysical sources and the role played by spatial fluctuations of the ionising background.

### References

7. P. McDonald et al., AJ 635, 761 (2005)
9. A. Slosar et al., JCAP 4, 26 (2013)
40. L. Zuo, MNARS 258, 36 (1992)
42. M. Meiksin et al., MNRAS 342, 1205 (2003)
44. V. Desjacques et al., MNRAS 444, 2793 (2014)
47. F. Haardt et al., in prep., (2016)
UNDERGROUND STUDY OF BIG BANG NUCLEOSYNTHESIS
IN THE PRECISION ERA OF COSMOLOGY

Carlo Gustavino
(for the LUNA collaboration)
INFN Sezione di Roma, I-00185 Roma, Italy

Abstract
The key reactions of big bang nucleosynthesis studied with the underground accelerator LUNA are reviewed and their implication in cosmology and particle physics are discussed. In particular it will be shown that the ongoing study of $^2\text{H}(p,\gamma)^3\text{He}$ reaction allows to significantly improve the accuracy of the cosmological baryon density and to constraint the existence of "dark radiation", e.g. sterile neutrinos, hot axions or other relativistic particles.

1 Introduction

In the standard cosmology the expansion rate of the universe is governed by the Friedmann equations:

$$H^2 = \frac{8\pi}{3} G \rho$$  \hspace{1cm} (1)
Were $H$ is the Hubble parameter, $G$ is the Newton’s gravitational constant and $\rho$ is the energy density which, in the early Universe, is dominated by the “radiation”, i.e. the contributions from massless or extremely relativistic particles. The radiation density can be expressed as follows:

$$\rho = \rho_{\gamma} \left[ 1 + \frac{7}{8} \left( \frac{4}{11} \right)^{4/3} N_{\text{eff}} \right]$$

(2)

In this formula $\rho_{\gamma}$ is the photon density and $N_{\text{eff}}$ is the contribution of other relativistic species. Using this formula $N_{\text{eff}} = 3.046$ if only the three known neutrino families are considered. Figure 1 shows the leading processes of BBN nuclear chain. The only free parameter to calculate the primordial abundances in standard BBN is the baryon density $\Omega_b$, which is usually expressed normalized to the black-body photon density as $\eta = n_B/n_{\gamma}$.

Table 1 summarizes both the results of BBN calculations (assuming the ΛCDM model and the $\eta$ parameter derived from CMB experiments) and the results of direct observations of light isotopes. In this table, the primordial $^4\text{He}$ abundance is given in terms of the baryonic mass fraction $Y_p$, while the abundance of the other nuclides is expressed as ratios by number.

The computed $^4\text{He}$ abundance essentially depends on the amount of free neutrons available, therefore its (very small) uncertainty is almost entirely due
Table 1: Calculated and observed abundances of light isotopes derived from standard BBN and from direct astrophysical observations (see text).

<table>
<thead>
<tr>
<th>Isotope</th>
<th>SBBN Theory</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y_p</td>
<td>0.2466±0.0006 [1]</td>
<td>0.254±0.003 [2]</td>
</tr>
<tr>
<td>D/H</td>
<td>(2.61±0.08)×10⁻⁵ [1]</td>
<td>(2.53±0.04)×10⁻⁵ [3]</td>
</tr>
<tr>
<td>³He/H</td>
<td>(1.00±0.01)×10⁻⁵ [4]</td>
<td>(0.9±1.3)×10⁻⁵ [6]</td>
</tr>
<tr>
<td>⁷Li/H</td>
<td>(4.68±0.67)×10⁻¹⁰ [4]</td>
<td>(1.23±0.68)×10⁻¹⁰ [7]</td>
</tr>
<tr>
<td>⁶Li/⁷Li</td>
<td>(1.5±0.3)×10⁻⁵ [8]</td>
<td>≲10⁻² [9]</td>
</tr>
</tbody>
</table>

to the neutron lifetime error. The abundance of helium strongly depends on the expansion rate of the Universe. In fact, faster is the expansion, faster is the cooling. As a consequence, the BBN inset starts earlier, when higher is the amount of neutrons available to form ⁴He. In other words, the amount of helium strongly depends on N_{eff} (see eq. 1 and 2), thus constraining the possible existence of "dark radiation", i.e. extra relativistic species in the early Universe not considered in the ΛCDM model. On the other hand, the primitive abundance of ⁴He derived from observations in HII (ionized hydrogen) regions of compact blue galaxies has a quite large error, limiting the possibility of deriving stringent constraints in cosmology and particle physics exploiting the ⁴He abundance.

Differently from helium, the error of (D/H)_{obs} is smaller than the theoretical error of (D/H)_{BBN}. This is because recent works in Damped Lyman-Alpha (DLA) systems at high redshifts provide the primordial abundance of deuterium with good accuracy [3], while the theoretical error is affected by the poorly known ²H(p,γ)³He process [10]. As it will be shown in the following, the accurate measurement of the ²H(p,γ)³He cross section allows to measure the η parameter (at the BBN epoch) with about the same accuracy of the one derived from CMB experiments (relative to the universal time of about 380,000 years). The (D/H) value is also sensitive to expansion rate of universe, therefore the deuterium can be exploited to constrain the existence of "dark radiation", in combination with the ⁴He results [3, 5] or with CMB data [1, 10]. Figure 2 shows deuterium abundance as functions of baryon-to-photon ratio and assuming different values of N_{eff}. It is quite clear from this figure that presently the main obstacle to improve the sensitivity in de-
Figure 2: Deuterium abundance as functions of baryon-to-photon ratio. The blue lines indicate abundances for a single value (integer plus 0.046) of \( N_{\text{eff}} \). The red bands indicate the nuclear uncertainty on those yields for \( N_{\text{eff}} = 3.046 \). The green band indicates the observational uncertainty of \((D/H)_{\text{obs}} \) [3].

\[\begin{align*}
\text{Presently the best estimation of the baryon density derived from CMB data is} \quad 100\Omega_{b,0}h^2(\text{CMB}) = 2.22 \pm 0.02 \quad [1], \quad \text{while the comparison of} \quad (D/H)_{\text{BBN}} \quad \text{and} \quad (D/H)_{\text{obs}} \quad \text{provides} \quad 100\Omega_{b,0}h^2(\text{BBN}) = 2.20 \pm 0.04 \pm 0.02 \quad [3]. \quad \text{In these equations} \quad \Omega_{b,0} \quad \text{is the present day baryon density of the universe and} \quad h \quad \text{is the Hubble constant in units of} \quad 100 \text{ km s}^{-1}\text{Mpc}^{-1}. \quad \text{The first error term of} \quad (D/H)_{\text{BBN}} \quad \text{is due to experimental nuclear cross section uncertainties, while the second one is due to the} \quad (D/H)_{\text{obs}} \quad \text{uncertainty} \quad [3]. \quad \text{Combining} \quad ^4\text{He} \quad \text{and} \quad ^2\text{H} \quad \text{calculation and observations} \quad N_{\text{eff}}(\text{BBN}) = 3.57 \pm 0.18 \quad \text{is obtained} \quad [3]. \quad \text{As it will be shown in the following, the uncertainty of} \quad \Omega_{b} \quad \text{and} \quad N_{\text{eff}} \quad \text{can be significantly reduced with the precision measurement of the} \quad ^2\text{H}(p,\gamma)^3\text{He cross section at BBN energies.}
\end{align*}\]

The \((^3\text{He}/H)_{\text{BBN}}\) value has a quite good theoretical error. Unfortunately, the \(^3\text{He} \) observations in our galaxy are affected by large systematics uncertainties [6]. Therefore, \(^3\text{He}\) does not represent a powerful probe to constrain the ΛCDM model.

The error budget of calculated abundance of \((^7\text{Li}/H)_{\text{obs}}\) depends on many nuclear reaction of the BBN reaction chain. The observed abundance is deduced from the strength of its characteristic absorption line at about 680 nm in low metallicity stars in the galactic halo. It is worth to point out the tension between observations and theory, referred in literature as “lithium problem”.

\[\text{275}\]
For what concern the $^6$Li isotope, it is reported in literature a controversial measurement in which the asymmetry of the absorption line of lithium may be due to a sizeable amount of $^6$Li with respect $^7$Li [11]. Even though many of the claimed $^6$Li detections are controversial, for a very few metal-poor stars there still seems to be a significant amount of $^6$Li ("second Lithium problem") [9, 12].

In summary, the BBN theory provides a powerful tool to constrain particle physics and cosmology. Although primordial abundances span many orders of magnitude, observations and theory are fairly in agreement, confirming the overall validity of standard BBN. However, some tension between theory and measurements is apparent, possibly due to physics beyond the Standard Model (see for example [4] and references therein).

Figure 3: Astrophysical S-factor data of the $^3$He$(\alpha, \gamma)^7$Be reaction as a function of the center-of-mass energy. A theoretical curve and the ab initio prediction are also shown.

2 The BBN fusion reactions at LUNA

At BBN energies ($30 \lesssim E_{cm}(keV) \lesssim 300$) the cross sections are very low because of the coulomb barrier between the interacting nuclei, making important to perform the measurements in a low background environment. To this end the LUNA accelerator operates deep underground, at the "Laboratori Nazionali del Gran Sasso" (LNGS), Italy. The Gran Sasso mountain provide a shield against cosmic ray muons providing the reduction of background induced by
cosmic rays of several order of magnitude with respect surface [13]. In nuclear astrophysics the cross section $\sigma(E)$ is often factorized as follows:

$$\sigma(E) = \frac{S(E)e^{-2\pi\eta^*}}{E}$$  \hspace{1cm} (3)

In this formula, the exponential term takes into account the Coulomb barrier, while the astrophysical factor $S(E)$ contains all the nuclear effects. The Sommerfeld parameter $\eta^*$ is given by $2\pi\eta^* = 31.20Z_1Z_2(\mu/E)^{1/2}$. $Z_1$ and $Z_2$ are the nuclear charges of the interacting nuclei. $\mu$ is their reduced mass (in units of a.m.u.), and $E$ is the center of mass energy (in units of keV). In the following are reviewed the most important results obtained by the LUNA collaboration concerning BBN.

2.1 $^3\text{He}(\alpha,\gamma)^7\text{Be}$

The BBN production of $^7\text{Li}$ is dominated by the $^3\text{He}(\alpha,\gamma)^7\text{Be}$ reaction, with subsequent decay of radioactive $^7\text{Be}$ to $^7\text{Li}$. The $^3\text{He}(\alpha,\gamma)^7\text{Be}$ reaction was studied at LUNA by detecting the promptly emitted $\gamma$-rays from the reaction and by measuring the $^7\text{Be}$ activity created in the experiment. Figure 3 shows the LUNA result [14–16] and literature data. Note that only the LUNA data are well inside the BBN energy region. The LUNA result exacerbates the "lithium problem", excluding a nuclear solution to solve the tension between theory and observations.

2.2 $^2\text{H}(\alpha,\gamma)^6\text{Li}$

Standard BBN production of $^6\text{Li}$ is dominated by just one nuclear reaction, $^2\text{H}(\alpha,\gamma)^6\text{Li}$ [17]. The $^2\text{H}(\alpha,\gamma)^6\text{Li}$ S-factor has been recently measured at low energy by LUNA. Figure 4 shows the LUNA result [8] together with previous direct measurements [19, 20] and theoretical calculations [21]. Note that only the LUNA data are well inside the BBN energy region, excluding a nuclear solution to explain the debated overabundance of this isotope in metal poor stars.

2.3 $^2\text{H}(p,\gamma)^3\text{He}$

The $^2\text{H}(p,\gamma)^3\text{He}$ was the first BBN reaction studied by the LUNA collaboration, by using the proton beam produced by the 50 kV pilot accelerator, a
windowless deuterium gas target and a 4π BGO crystal to detect prompt \(\gamma\)s [22]. Although with this accelerator was possible to measure the cross section only up to \(E_{cm} = 23\) keV, this measurement allowed to reduce the uncertainty of deuterium abundance of a factor 3 with respect to previous estimations. Figure 5 shows the data of the \(^2H(p,\gamma)^3He\) reaction in literature. Only a single dataset of \(S_{12}\) is currently available in the relevant BBN energy range, in which the authors state systematic uncertainty of 9% [23]. Figure 5 also shows the behavior of \(S_{12}\) obtained by the theoretical ab initio calculation [24]. This theoretical result is about 20% higher than the fit of experimental data. The existing difference between theory and data let some author to adopt the theoretical curve [5] or the \(S_{12}\) value obtained from measurements [25].

The deuterium primordial abundance is determined by the cross sections of \(p(n,\gamma)^2H\), \(^2H(p,\gamma)^3He\), \(^2H(\alpha,\gamma)^6Li\) and \(^2H(p,\gamma)^3H\) reactions. Table 2 shows the contribution of each reaction to the deuterium abundance error budget [10], that is mainly due to the uncertainty of the \(^2H(p,\gamma)^3He\) cross section. The uncertainty due to this reaction is even larger if ab initio prediction is taken into account. The present LUNA 400 kV facility [26] make possible to extend the measurements up to \(E_{cm} = 266\) keV, i.e. well inside the BBN energy range. Figure 6a shows experimental setup in which a barrel BGO detector with about 70% efficiency is implemented. The angular distribution of photons emitted by
the $^2H(p, \gamma)^3He$ reaction is studied with a Germanium detector facing the gas target, as shown in figure 6b. The angular distribution is inferred exploiting the the doppler effect affecting the energy of emitted $\gamma$'s by the $^2H(p, \gamma)^3He$ reaction. Figure 7 shows the result of a preliminary measurement performed with the HPGe detector. The energy distribution of emitted photons is well in agreement with ab initio calculation.

Figure 5: S-factor data for the reaction $^2H(p, \gamma)^3He$. The red solid curve shows the prediction of recent ab initio theoretical calculation.

Figure 6: a): Scheme of gas target setup and BGO detector. b): Scheme of gas target setup and HPGe detector.
Table 2: List of the leading reactions and corresponding rate symbols controlling the deuterium abundance after BBN. The last column shows the error on the ratio $D/H$ coming from experimental (or theoretical) uncertainties in the cross section of each reaction, for a fixed baryon density $\Omega_b h^2 = 0.02207$.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Rate Symbol</th>
<th>$\sigma_{D/H} \cdot 10^5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(n,\gamma)^2H$</td>
<td>$R_1$</td>
<td>±0.002</td>
</tr>
<tr>
<td>$d(p,\gamma)^3He$</td>
<td>$R_2$</td>
<td>±0.062</td>
</tr>
<tr>
<td>$d(d,n)^3He$</td>
<td>$R_3$</td>
<td>±0.020</td>
</tr>
<tr>
<td>$d(d,p)^3H$</td>
<td>$R_4$</td>
<td>±0.0013</td>
</tr>
</tbody>
</table>

3 Conclusions

The LUNA experiment has measured with high accuracy several leading processes of BBN. For what concern $^4He$, $^3He$, $^7Li$ and $^6Li$ the constraints in cosmology and particle physics are presently limited by the uncertainty of abundances derived from astronomical observation. Instead, direct observations of deuterium abundance [3] and the accuracy of CMB data [25] make the lack of data of the $^2H(p,\gamma)^3He$ reaction the main obstacle to improve the accuracy of $\Omega_{b,0}(BBN)$ and $N_{eff}$. The study of the $^2H(p,\gamma)^3He$ reaction in the BBN energy range is in progress with the LUNA400 facility. The goal is to reach an accuracy of about 3% level, to resolve the 20% tension between data and nuclear calculations and to improve the 9% uncertainty quoted in the literature data.

References

Figure 7: Simulated spectra of the $^2H(p,\gamma)^3He$ reaction, assuming isotropic (green) and ab initio (blue) angular distribution at $E_{cm} = 112.5$ keV. The experimental data (red) are also shown. Data have been normalized to remark the close agreement with the ab initio angular distribution.

Science with the Square Kilometer Array

Luigina Feretti

Istituto di Radioastronomia INAF
Via P. Gobetti n. 101, 40129 Bologna, Italy

Abstract

The Square Kilometre Array project is an international effort to build the world largest radio telescope, with a collecting area of a square kilometre (one million square metres). It will represent a huge advancement in both engineering and research & development. This unique instrument will bring together a wealth of the worlds finest scientists and engineers to answer fundamental questions in physics and astrophysics and change our understanding of the Universe.

1 Introduction

The Square Kilometer Array (SKA) is a global project to build the next-generation interferometric radio telescope operating at metre to cm wavelengths. Whilst 10 member countries are the cornerstone of the SKA (Australia, Canada, China, India, Italy, New Zealand, South Africa, Sweden, Netherlands,
United Kingdom), around 100 organisations across about 20 countries are participating in its design and development. The organisation is managed by the not-for-profit SKA Organisation, which has the headquarters at the Jodrell Bank Observatory, near Manchester in the United Kingdom. World leading scientists and engineers are designing and developing a system which will require supercomputers faster than any in existence, and network technology that will generate more data traffic than the entire Internet.

Figure 1: Artist impression of the SKA antennas. Top: dishes for mid to high frequency observations; Middle: dense aperture arrays for mid frequency observations; Bottom: aperture arrays for low frequency observations (credit SKA Organization/Swinburne Astronomy Productions)
The SKA will use thousands of dishes, each about 15m wide. Two other types of antenna, known as aperture arrays, will also be used to look at very large areas of the sky all at once (Fig. 1). The antennas will be arranged in spiral arms going out from a central core. The mid to high frequency dishes will be located in South Africa and Africa, whereas the low frequency antennas will be located in Western Australia. The two sites have been chosen for their very low-level of man-made radio interference, as well as for ionosphere and troposphere characteristics, connectivity, communication network, and costs.

The SKA will be developed over a phased timeline. Pre-construction development started in 2012 and involves the detailed design, implementation, R&D work, and contract preparation needed to bring the SKA first phase to construction readiness. For SKA Phase 1, Australia will host the low frequency instrument with more than 500 stations, each containing around 250 individual antennas, whilst South Africa will host an array of about 200 dishes, incorporating the 64-dish MeerKAT precursor telescope. Phase 2 will complete the telescope arrays at both sites, and become fully operational in the late 2020s, by which time the SKA will consist of about 2000 high and mid frequency dishes and aperture arrays and a million low frequency antennas.

The SKA will already start conducting science observations in 2020 with a partial array. Even before the SKA comes online, a series of demonstrator telescopes and systems known as pathfinders, as well as precursors located at future SKA sites (Australia SKA Pathfinder ASKAP, MeerKAT, Murchinson Wide Field Array MWA, Hydrogen Epoch of Reionization Array HERA) are already operational or under development, paving the way for the kinds of technology which the SKA will need to pioneer to make the huge data available to scientists.

2 SKA Specifications

The SKA will provide up to 1 million m$^2$ of collecting area distributed over a distance of $\sim 3000$ km on a frequency range from 70 MHz to 10 GHz. The bandwidth is of about 50% of frequency, the expected image dynamic range is up to $10^6$ and the polarization purity is $\sim 30$ db.

**Sensitivity:** SKA will provide an improvement of 100-fold in sensitivity with respect to current best instruments. Fig. 2 plots the point source continuum sensitivity of radio telescopes over the years. The increase has been
exponential and doubling every about three years. This is due both to new instruments and to new technologies. SKA is the next step.

**Survey Speed**: Taking advantage of technology developments in radio frequency devices and digital processing, SKA will achieve a sky imaging capability about 10,000 times faster than the best current imaging radio telescopes. Moreover, it will be able to observe the sky with multiple independent beams.

**Angular Resolution**: With 3000 km max baseline, the SKA will reach the mas angular resolution. This is comparable to the angular resolution currently reached with the Very Long Baseline Interferometry (VLBI) technique, and represents an increase of 1000-fold with respect to traditional interferometers.

Figure 2: *Radio telescope sensitivity versus time, when telescopes were built or after major upgrades, from Ekers 2*)
3 Science Goals

The SKA will be able to conduct transformational science, breaking new ground in astronomical observations. SKA scientists have focused on various key science goals for the telescope, each of which will re-define our understanding of space as we know it. The SKA Science has been recently reviewed and presented in a 2000 pages book that contains 135 chapters and 1200 contributors 1), while the SKA Key Science projects are under definition. Early science observations are expected start in 2020 with a partial array.

At low frequencies (< 200 MHz), the main goal is the study of the early Universe through the redshifted HI 21-cm line. At intermediate frequencies (∼ 1 GHz), studies of pulsars, galaxy evolution, cosmology and magnetism will be carried out. Higher frequencies (> 3 GHz) will be suitable for more studies of magnetism and for searches of molecular gas.

A major breakthrough will be reached through the planned all-sky surveys. In more detail: the HI survey will provide intensity mapping and redshift of $10^9$ objects up to $z \sim 2$; the radio continuum survey will detect $10^{10}$ objects, i.e. $\sim 2 \times 10^5$ sources/deg$^2$, down to a flux level of 100 mJy; the survey in polarization will provide the Rotation Measure (RM) of $10^7$ objects, i.e. 300 sources/deg$^2$ (currently RMs are available for $\sim 1$ source/deg$^2$ 13).

The main topics identified in the Science Book 1) are summarized below.

**Cosmology and Dark Energy:** The new frontier of cosmology will be led by 3-dimensional surveys of the large-scale structure of the Universe, based on the all-sky surveys mentioned above. This will constrain non-gaussianity of primordial fluctuations, measure the matter dipole with high precision for a comparison with the matter dipole, analyze the Barion Acoustic Oscillation (BAO) signal to a level competitive with optical surveys, and address the nature of Dark Energy. SKA will revolutionize cosmology, in combination with future optical/infrared surveys such as Euclid and Large Synoptic Survey Telescope 9).

**Cosmic dawn and the Epoch of Reionization:** SKA will allow the direct detection of the 21-cm signal during the epoch of cosmic reionization into the preceding dark ages, at $z \sim 6-28$, thus probing the neutral intergalactic medium that pervaded the Universe during and prior to the formation of first galaxies 7). Imaging of the 21-cm line provides the most direct means to study the earliest galaxies and the evolution of large scale structure.
Fundamental Physics with pulsars: SKA will provide a 10-fold increase in the number of known pulsars and milli-second pulsars, including binary systems, and possibly discover rare and exotic objects. With pulsars being strongly self-gravitating bodies and very high-precision clocks, timing observations of these objects will allow strong-field experiments, in particular tests of general relativity, alternative theories of gravity, and extreme magnetic fields \(^8\). Thousands of milli-second pulsars will form a pulsar timing array, which will be able to detect nano-Hz gravitational waves (GW) and the expected GW background due to coalescing binary super-massive black holes \(^5\).

Transient Universe: Radio transients are both the sign and signature of the most extreme phenomena in the Universe, like exploding stars, compact object mergers, black holes, ultra-relativistic flows \(^3\). SKA will be a superb facility for the study of transients, in particular the detection of new objects and the fast response to transient alert. Of great interest are currently the Fast Radio Bursts (FRB), very short and apparently coherent bursts that may be tracers of new astrophysical phenomena (see Burgay, this conference).

Continuum Universe: Radio continuum surveys are a common tool to address several scientific areas like galaxy and cluster evolution, black holes, star formation history, accretion process, transients, cosmology, etc. \(^11\) In a study of the microJy and nanoJy sky, and source population multiwavelength properties, Padovani \(^10\) argued that most sources from currently planned all-sky surveys will have a radio counterpart, with the likely exception of the optical ones, and in selected areas. On large areas of the sky, and at lowest flux levels, radio sources detected with SKA will have no counterparts, therefore the radio information will be the only one available, thus of paramount importance.

Magnetism: Magnetic fields are present in astrophysical objects and in the intergalactic medium over vastly different scales and strengths, however many of their properties are still poorly known. Crucial questions are how the fields originated and which is their impact on life-cycles of the structure and object evolution. Surveys of polarization properties of the sky will allow the characterization of the Galaxy magnetic field, and magnetic fields on a large range of scales with high precision \(^6\), while deep small surveys and targeted observations will supply detailed information on selected objects.

Cradle of Life: The combination of sensitivity and resolution of SKA will allow imaging of proto-planetary disks and probing the formation of habi-
itable planets. SKA will also detect the presence of heavy molecules that are the building blocks of life \(^4\), and contribute to SETI studies. SKA can detect airport radars at \(\sim 30\) light years, thus surveying about 1000 stars, many with planets, and therefore probe signals from potential external civilizations.

**Hydrogen Universe:** The detailed imaging of neutral hydrogen (HI) in external galaxies, in our Galaxy and in the local group will allow investigations of the structure and kinematics of galaxies, of AGN-driven outflows, of the formation, growth and evolution of galaxies in different environments \(^12\). In addition to studies of the HI emission, the analysis of intervening HI absorption will be crucial to trace the gas at low optical depth, up to very high redshifts.

**Synergy with other projects:** For all the topics given above, combining and/or cross-correlating SKA results with multifrequency data from other projects (VLBI, ALMA, LSST, Euclid, CTA, LIGO, LISA, CMB projects, ultra high energy cosmic rays detectors, neutrino experiments, etc.) will represent a powerful tool to improve our knowledge in physics and astrophysics.

### 4 Conclusions

From challenging theory of relativity, looking at how the very first stars and galaxies formed just after the big bang, understanding the nature of a mysterious force known as dark energy, to exploring the vast magnetic fields which permeate the cosmos and investigating the formation of life, the SKA will truly be at the forefront of scientific research. Moreover, with such a sophisticated and powerful instrument, greatest serendipitous discoveries will likely be even richer and beyond prediction \(^14\).

Complete information on the SKA project and updates on its progresses can be found at the project web site https://www.skatelescope.org.

### 5 Acknowledgements

The author wishes to acknowledge financial support from PRIN INAF 2014.
References


Recent results obtained at LHC show deviations from predictions of the Standard Model. Therefore it is appropriate to remind about results obtained in cosmic ray experiments earlier at the same energies (in the center of mass system). In this paper, the comparison of LHC and cosmic ray data is fulfilled and various possibility of their explanation is considered.

1 Introduction

Now TeV energy region is being intensively investigated in LHC experiments and various deviations from predictions of the standard model are observed. Some of them are confirmed in further experiments, others are not confirmed. But it is interesting to note that, as a rule, the results obtained in nucleus-nucleus interactions (f.e. so-called imbalanced events \(^1\)), sharp increasing of
secondary particle multiplicity with energy \(^2\)) are confirmed. Results obtained in proton-proton interactions (2 TeV resonance \(^3\)), excess of missing energy and lepton transverse momentum \(^4\)) are not always confirmed. In this connection it is necessary to remind that TeV energy region is investigated in cosmic rays more than 60 years, and many unusual results were obtained (i.e. alignment, penetrating cascades, Centauros \(^5\), various deviations in EAS development \(^6\), excess of muon bundles, increasing with energy \(^7, \, 8\)). The last effect was observed at accelerator detectors, too, firstly in LEP detectors \(^9, \, 10\), then at LHC detectors \(^11\). For explanation of different unusual events various theoretical models were proposed, but none of them can explain all observed events and phenomena. In papers \(^12, \, 13, \, 14\) the model of production of quark-gluon matter blobs with large orbital momentum was proposed, which allows explain all observed experimental data.

2 The backgrounds and requirements to new model

The necessity of development of a new model for description of all unusual experimental data is caused by a large amount and contradictoriness of various requirements for their explanation.

The first of them is the cross section. In accelerator experiments, rather rare events (with cross section of nb and even pb) can be measured due to a large intensity of beams. In cosmic ray experiments, the values of the cross section of the order of mb (sometimes \(\mu b\)) can be detected only due to a low intensity of CR flux. Since all observed unusual phenomena have a threshold behaviour (they are detected at TeV energies in the center of mass system, which correspond to PeV energy region in cosmic rays) the simplest way of their explanation is a production of some heavy particles (or states of matter). But in this case geometrical cross section

\[ \sigma = \pi \lambda^2 = \pi / m^2, \]  

will be very small (\(\sim 10^{-34} \text{ cm}^2\) at \(m \sim 1 \text{ TeV}\)). Therefore a transition from point-like quark-quark interactions to multi-quark and gluon interactions is required. In this case, some blob of quark-gluon matter is produced and geometrical cross section will be

\[ \sigma = \pi R^2, \]
where $R$ is an effective size of QGM blob, which cannot be less than one nucleon size, and cross section will be of the order of mb.

The second important point is the energy in the center of mass. Usually even for nucleus-nucleus interactions the energy in the center of mass system is calculated for target of nucleon mass. But for collective interaction of many quarks and gluons the target mass can be more than nucleon mass. Therefore instead of

$$\sqrt{s} = \sqrt{2m_N E}$$

we must use

$$\sqrt{s} = \sqrt{2m_c E},$$

(3)

where $m_c$ – some compound mass, which in the first approximation can be determined as

$$m_c = nm_N.$$  

(4)

The third point is connected with an orbital momentum. Its value and significance strongly increase for collective interactions. As was shown in 15), the value of the orbital momentum is proportional to $\sqrt{s}$ and calculations 16) showed that its value can reach $\sim 10^4$. At such large orbital momentum $L$, a big centrifugal barrier must appear:

$$V = L^2/2mR^2.$$  

(5)

Its value will be large for light quarks ($u$ and $d$), but small for heavy quarks (i.e. for top-quarks). And though top-quarks are absent in interacting nuclei, but suppression of decays of QGM blob into light quarks gives a time for top-quark pair production in boiling quark-gluon matter and their fly out from the blob. This process decreases the energy $\sqrt{s}$, and correspondingly orbital momentum and centrifugal barrier are sharply decreased and the rest part of the blob decays into light quarks. The decay of top-quarks:

$$t(\bar{t} \rightarrow b(\bar{b}) + W^+(W^-)).$$  

(6)

In their turn, $b$-quarks give jets or can decay into $c$-quarks. $W$-bosons decay into leptons ($\sim 30\%$) and into hadrons (mainly pions) $\sim 70\%$. These changes of interaction model allow explain all unusual experimental results obtained both in accelerator and cosmic ray experiments.
3 Explanation of unusual results of accelerator experiments

In nucleus-nucleus interactions two undoubted results were obtained. The first, so-called unbalanced events in which dijet symmetry is violated \(^1\). In the frame of the considered model these events can be explained very easily. The production and consecutive decay of \(t\)-quark gives \(b\)-quark and \(W\)-boson. Interaction of \(b\)-quark gives jet and the decay \(W\)-boson into pions (\(\sim 10\)) forms an unbalanced event. Really at the decay of \(t\)-quark kinetic energies are distributed as

\[
T_b \approx 65 \text{ GeV} \quad \text{and} \quad T_W \approx 25 \text{ GeV}.
\]

If to take into account fly-out energy of \(t\)-quark, \(T_b\) can be more than 100 GeV and ATLAS experiment’s picture will be obtained.

The second, the sharp increasing of secondary particle multiplicity in heavy nuclei interactions \(^2\). It is important to underline that this result is obtained for energy in center of mass system, which is calculated for nucleon mass \(\sqrt{s_{NN}}\). But in frame of the considered model this mass must be larger and LHC results allow evaluate it. In fig.1 possible positions of experimental point are shown.

![Figure 1: Charged particle multiplicity according to accelerator experiments \(^2\).](image)

---

\(^1\) Explanation of unusual results of accelerator experiments.

\(^2\) Explanation of unusual results of accelerator experiments.
There are two limiting positions. The upper limit can be evaluated from condition, that \( \sqrt{s_{AA}} \) for nucleus-nucleus interaction cannot be more than \( \sqrt{s_{NN}} \) for pp-interaction. In this case a number of nucleons in QGM blob

\[
\sqrt{\frac{n}{N}} = \sqrt{\frac{s_{AA}}{s_{NN}}} < \frac{50}{3.5} \approx 14.
\]

This case corresponds to the central collision of nuclei, and target mass is equal to mass of interacting nucleus (~ 200). The lower limit can be evaluated if to assume that the energy dependence of AA-interaction on energy is the same as for pp-interaction. In this case \( \sqrt{s_{AA}} \) will be about 20 TeV and correspondingly \( \sqrt{\frac{n}{N}} \) will be about 6, and the total number of nucleons in a blob ~ 36. In this case a target mass is equal to about 1/6 of total target nucleus mass.

It is interesting to compare obtained values with the simplest geometrical models. At low energies nuclei can be considered as spheres and their average region of intersection is shown in fig. 2, left.

![Figure 2: Region of intersection of two spheres (left) and two disks (right).](image)

The volume of spherical segment is

\[
V = \pi h^2 (3R - h)/3. \tag{7}
\]

On average, \( h = R/2 \) and volume

\[
V = \pi \frac{R^3}{4} (3R - \frac{R}{2})/3 = \frac{4}{3} \pi R^3 \left( \frac{5}{8} \right) \approx 0.156 V_{\text{sphere}}.
\]

Two volumes will be about 0.31 \( \approx 1/3 \). At very high-energies, nuclei can be considered as flat disks and correspondingly the region of intersection will be equal to two flat segments (fig. 2, right).

Area of each will be equal to

\[
S = h(6a + 8b)/15. \tag{8}
\]

For picture in fig. 2, right: \( b = R \), \( h = R/2 \), \( a = 2R\sqrt{3}/2 = R\sqrt{3}, \)
\[ S = \frac{R}{2} (6R\sqrt{3} + 8R) / 15 = R^2 \left( 3\sqrt{3} + 4 \right) / 15 \approx \pi R^2 \frac{0.2}{15} \approx 0.2 S_{\text{disk}}. \]

For two segments, the area will be equal to \( \approx 0.4 \) of full nucleus target area. Both values 0.3–0.4 lie in experimental interval 0.17–1.

Very interesting results were obtained using LHC detectors for investigations of cosmic ray muons. Due to a high spatial resolution, these detectors can register muon bundles with big multiplicity. In fig. 3 the results of such investigation in ALICE are presented \(^{11}\).

![Figure 3: Atmospheric muon multiplicity distribution in ALICE detector \(^{11}\).](image)

A remarkable excess of bundles with high multiplicity (more than 100) was detected. This excess cannot be explained in the frame of traditional process of muon generation in decays of various mesons (\( \pi, K, D \) etc.). The similar results were obtained earlier at LEP detectors (ALEPH \(^9\), DELPHI \(^10\)). Unfortunately, in experiments at accelerator detectors there is no possibility to evaluate the energies of primary particles which are responsible for the appearance of muon bundles with high multiplicity. Such possibility has the experimental complex NEVOD-DECOR \(^7\) (see the next part of the paper).

4 Explanation of CR unusual events

Firstly, it is necessary to underline that cosmic rays consist mainly of nuclei (\( \sim 60\% \)), if consider their energies per particle (see tab.1). Usual opinion
that cosmic rays consist mainly of protons (~90%) is based on calculations of secondary particle flux in the atmosphere in which a large contribution give leading particles, for that energy per nucleon is important. For EAS measurements, full primary particle energy must be taken into account. Results of EAS investigations show that above the “knee” the part of nuclei is increased. So in cosmic rays most part of interactions are nucleus-nucleus interactions and less part of proton-nucleus interactions.

Table 1: Composition of cosmic rays at low energies.

<table>
<thead>
<tr>
<th>Particles</th>
<th>Z</th>
<th>A</th>
<th>Energy per nucleon</th>
<th>Energy per nucleus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protons</td>
<td>1</td>
<td>1</td>
<td>92%</td>
<td>40%</td>
</tr>
<tr>
<td>α-particles</td>
<td>2</td>
<td>4</td>
<td>7%</td>
<td>21%</td>
</tr>
<tr>
<td>Light nuclei</td>
<td>3-5</td>
<td>10</td>
<td>0.15%</td>
<td>3%</td>
</tr>
<tr>
<td>Medium nuclei</td>
<td>6-10</td>
<td>15</td>
<td>0.7%</td>
<td>18%</td>
</tr>
<tr>
<td>Heavy nuclei</td>
<td>≥11</td>
<td>32</td>
<td>0.15%</td>
<td>18%</td>
</tr>
</tbody>
</table>

The explanation of various unusual events observed in cosmic rays was given elsewhere 12). In this paper the problem of the excess of muon bundles (so-called “muon puzzle”) will be considered only. A serious advancement in investigations of the dependence of muon bundle intensity on primary particle energy was done in experiment NEVOD-DECOR 7). In this experiment primary particle energy was evaluated by measurements of zenith angle dependence. In fig.4 the results of inclined muon bundle detection are presented.

From this figure it is seen that at increasing of zenith angle the number of muon bundles is increased in comparison with theoretical calculations at respective zenith angle. Though there is no direct dependence between zenith angle and primary particle energy for each individual event, in general such dependence exists. In fig.5 the results of simulations of the contribution of different primary energies into production of events with fixed local muon density at various zenith angles are presented 7).

As one can see from fig.4, the increasing of local muon density starts at energies more 10^{16} eV, which correspond several TeV energies in center of mass system. In principle, this increasing up to energy 10^{17} eV can be explained if CR mass composition becomes heavier. But above 10^{17} eV such explanation is impossible. The further increase of excess of muon bundle number was measured in Pierre Auger Observatory (fig.6 8)).
Figure 4: Experimental (points) and calculated (curves) local muon density spectra for different zenith angles; arrows indicate effective primary energies.

Figure 5: Contribution of various primary energies into muon bundles flux for fixed local muon density in dependence on zenith angle.
In the model of QGM blobs, the solution of “muon puzzle” is the following. For production of QGM blob not only high temperature (energy) but high quark-gluon density is required. Therefore firstly such blobs will be generated in interactions of heavy particles (iron nuclei) with nuclei of atmospheric atoms. Then, with increasing of primary particle energy, in interactions of more light nuclei. The last will be protons. Of course nucleus interactions have big fluctuations and clear separation of contribution of various nuclei is impossible. But in general the dependence on nucleus mass exists.

Increasing number of QGM blobs with energy and corresponding number of $W$-bosons, which mainly decay into pions (on average, $\sim 20$) with not large energy increases the multiplicity of muons bundles.

5 Conclusion

Investigations of TeV energy region gave many interesting experimental results. Proposed model of production of QGM blobs with large orbital momentum allows explain all new phenomena observed in this energy region.
Acknowledgements

Authors thank R.P. Kokoulin for fruitful discussions and the help in preparation of this paper. This work was supported by Ministry of Education and Science of the Russian Federation (government task).

References

Accelerator dark matter searches: comparison with direct and indirect searches

Mario Antonelli

Laboratori Nazionali di Frascati - INFN
Via E. Fermi 40 Frascati

Abstract

Results from searches for dark matter performed at modern accelerators will be reviewed. Comparison with results obtained in direct and indirect search experiments will be described in the context of specific frameworks.

1 Introduction

Gravitational effects on astrophysical scales supported by the large-scale structure of the Universe and measurements of the cosmic microwave background are consistent with the existence of Dark Matter particles (DM) in Nature. However, the nature of DM particles is still unclear. Although the Standard Model (SM) has achieved great success after the LHC discovery of a 125 GeV Higgs boson, it cannot provide any suitable candidate for the cold DM. Therefore, the existence of DM gives a hint of new physics beyond the SM (BSM).
The most popular and attractive candidates for DM are the so-called weakly interacting massive particles (WIMPs), which can be thermally produced in the early Universe and naturally give the correct observed DM relic density. Because of their weak interactions with SM particles, WIMPs with masses of $O(10^2)$ GeV are expected to be produced at high energy colliders. Therefore, searching for such DM particles is a very important task for: (i) collider experiments; (ii) direct detection experiments in shielded underground detectors; (iii) for indirect detection experiments with satellites, balloons, and ground-based telescopes looking for signals of DM annihilation.

Very little is known about the properties of the DM particle(s) as well as their interaction with the known SM particles. Many models have been proposed falling into three main distinct classes:
(i) Models where the DM may be the only accessible state to our experiments. In such a case, effective field theory (EFT) allows us to describe the DM SM interactions mediated by all kinematically inaccessible particles in a universal way.
(ii) The large energies accessible at the LHC call into question the validity of the EFT approximation. Models characterized by the presence of additional states mediating the DM particle interactions with the SM, as well as the DM particle itself, have been proposed.
(iii) Complete DM models close this gap by adding more particles to the SM, most of which are not suitable DM candidates. The classical example is the Minimal Supersymmetric SM (MSSM), with R-parity conservation, where the natural DM candidate is the lightest neutralino.

2 Collider searches

Collider searches are performed with dedicated analyses optimised to detect the signatures predicted by the DM models.

2.1 Searches in the contest of EFT DM models

Search for pair-produced DM particles in collisions of SM particle is typical of EFT inspired models. Since DM particles escape from the detector without energy deposit, an additional energetic SM particle (jet, photon, W/Z ...) is required to reconstruct the signature with missing transverse energy MET.
The latest ATLAS and CMS analyses have been optimised to search for monojet, monophoton, "monoW", and "monoZ" with MET signatures. No significant excesses have been found with respect to the SM expectations and hence constraints on the energy scales of effective contact operators describing DM interactions with SM particles have been set in the contest of EFT models. Fig. 1 shows the ATLAS results expressed in upper limit on the DM-nucleon scattering cross section as a function of the DM mass, $m_\chi$, and compare to those obtained with direct and indirect searches.

It can be seen that collider searches have better sensitivity with respect to direct searches to lower value of DM masses. Indirect searches are instead effective at high DM masses.

The complementarity of the three approaches holds for various kind of operators of the effective lagrangian (DM cooling to SM particle type). Fig. 2 shows the expectation on the DM searches for a representative, generation-independent, couplings of a spin-1/2 dark matter particle $\chi$ with quarks $q$,
gluons $g$, and leptons $\ell$ (including neutrinos) EFT, given by

$$\frac{1}{M^2} \bar{\chi} \gamma^\mu \gamma^5 \chi \sum_q \bar{q} \gamma_\mu q + \frac{\alpha_S}{M^3} \bar{\chi} \gamma^\mu \gamma^5 C^a \chi \sum_q \bar{q} \gamma_\mu q + \frac{1}{M^2} \bar{\chi} \gamma^\mu \chi \sum_\ell \bar{\ell} \gamma_\mu \ell . \quad (1)$$

Figure 2: Dark matter discovery prospects in the $(m_\chi, \sigma/\sigma_{th})$ plane for current and future direct detection, indirect detection, and particle colliders for dark matter coupling to gluons, quarks, and leptons, as indicated.

The annihilation cross section is normalized to the value $\sigma_{th}$, which is required for a thermal WIMP to account for all of the dark matter in the Universe.

As long as the mediators interacting with both DM and SM particles are so heavy that they can be integrated out, the EFT approach is valid. However, it has many shortages and would break down in many cases, as pointed out in several recent works. It is well known that the EFT fails when the typical momentum transfer involved in the reaction is comparable to the mediator mass. This means that in order to safely use the EFT, the scale of BSM physics should be much higher than the collision energy, otherwise mediators may be directly produced in collisions. Besides, the EFT approach is invalid for a mediator with a width comparable to or larger than its mass [16, 23]. Furthermore, unitary and perturbativity conditions could also set constraints on the validity of the EFT. In a more appropriate approach, so-called simplified models, mediators with moderate masses are introduced to connect DM particles to SM particles. In principle, the simplified model approach can be mapped into the EFT approach, as studied in the limit of heavy mediators. On
the other hand, it can be realized as particular cases of UV complete models, and has been widely used in supersymmetry studies. For a mediator with a moderate mass, full kinematics and topologies of DM signatures at colliders can be studied in details. Furthermore, collider constraints and reaches could also be easily compared with those from direct and indirect detection experiments in specific simplified models. Fig. 3 shows the 95% CL exclusion regions in $M_{\text{med}}$-$m_{\text{DM}}$ plane for searches performed by the CMS experiment.

Figure 3: 95% CL exclusion regions in $M_{\text{med}}$-$m_{\text{DM}}$ plane for different MET based DM searches from CMS in the lepto-phobic AV and V models. It should be noted that the exclusion regions and relic density contours in this plot are not applicable to other choices of coupling values or models.

2.2 Searches in the contest of MSSM

Supersymmetry (SUSY) represents an appealing possibility for Beyond the Standard Model Physics; its discovery would help provide answers to many of the preeminent questions in particle physics, astrophysics, and cosmology. The most general supersymmetric extension of the SM, the MSSM, requires 105 parameters to describe SUSY breaking, including flavour generation, in addition to the parameters of the SM. Searches in this contest are performed in terms of simplified models and searched signatures are inspired by predictions from SUSY breaking models or by the request to reduce the quadratic divergence of the higgs mass radiative correction that lead to light stop and sbottom.
Search for squarks and gluinos

Squarks ($\tilde{q}$) and gluinos ($\tilde{g}$) are expected to be produced in pairs ($\tilde{q}\tilde{q}$, $\tilde{q}\tilde{g}$, $\tilde{g}\tilde{g}$) in many of the R-parity SUSY models. Some of the most simple decays are to neutralinos ($\tilde{\chi}_i^0$) $\tilde{q} \rightarrow q\tilde{\chi}_i^0$, $\tilde{g} \rightarrow g\tilde{\chi}_i^0$ or to charginos ($\tilde{\chi}_i^\pm$) $\tilde{q} \rightarrow q\tilde{\chi}_i^\pm$, $\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_i^\pm$ with the charginos decaying via a $W^\pm$ to neutralinos. Depending on the SUSY spectrum below the gluino/squark mass the chain of decays can be long, including 3-4 decay steps. Hence these decays are characterised by large jet multiplicities and leptons in the final states. The latest result from the ATLAS collaboration is summarized in Fig.4 in the contest of MSUGRA/CMSSM model.

![Figure 4: Exclusion limits at 95% CL for 8 TeV analyses in the $(m_0, m_{1/2})$ plane for the MSUGRA/CMSSM model with the remaining parameters set to $\tan(\beta) = 30$, $A_0 = -2m_0$, $\mu > 0$. Part of the model plane accommodates a lightest neutral scalar Higgs boson mass of 125 GeV. This plot is from: 1507.05525 (Summary of the searches for squarks and gluinos using $\sqrt{s} = 8$ TeV pp collisions with the ATLAS experiment at the LHC).](image)

Searches for third generation of squarks

In Natural SUSY the superpartner of the top quark, $\tilde{t_1}$, is expected to have a mass below 1 TeV. Squarks of the 3rd family (stops and sbottoms) are expected...
to be light, because for low masses the top loop diagrams contribution to the
Higgs mass can be cancelled without introducing an excessive amount of fine-
tuning. Depending on the mass of the stop the following decays could be
dominant: $\tilde{t}_1 \to t\tilde{\chi}_1^0$, $\tilde{t}_1 \to bW\tilde{\chi}_1^0$, $\tilde{t}_1 \to bff'\tilde{\chi}_1^0$ or $\tilde{t}_1 \to c\tilde{\chi}_1^0$. The searches
are designed such that they cover all the possible decays of the stop into a
neutralino LSP and make use of advanced techniques for reconstructing the
decay products. For the latter case searches include topologies in which the
W is boosted and reconstructed as a single jet. Summary of limits from these
topologies are shown in Fig. 5.

Figure 5: Summary of the dedicated ATLAS searches for top squark (stop) pair
production based on $13 \text{ fb}^{-1}$ of pp collision data taken at $s = 13 \text{ TeV}$. Exclusion
limits at $95\% \text{ CL}$ are shown in the stop neutralino mass plane. The included
decay modes, all considered having $BR = 100\%$ are: $\tilde{t}_1 \to t\tilde{\chi}_1^0$, $\tilde{t}_1 \to bW\tilde{\chi}_1^0$, $\tilde{t}_1 \to bff'\tilde{\chi}_1^0$ and $\tilde{t}_1 \to c\tilde{\chi}_1^0$. The dashed and solid lines show the expected and
observed limits, respectively, including all uncertainties except the theoretical
signal cross section uncertainties.

Electroweak sector searches

Direct production of charginos ($\tilde{\chi}^{\pm}$) and neutralinos ($\tilde{\chi}_i^0$), considering the $\tilde{\chi}_1^0$
as LSP, is expected to have a clean signature due to the presence of SM bosons
and leptons in the decay chains. Depending on the mass splitting a heavier neutralino can decay to the LSP via a Z or the SM Higgs, and the charginos via a W. The final states in direct pair production of charginos and neutralinos are characterised by a large lepton multiplicity (≥ 2). Depending on the decay, a dilepton invariant mass compatible with the Z mass can be vetoed or required.

The $\tilde{\chi}^+_1$ and $\tilde{\chi}^0_2$, assumed to be produced in pairs, decay to the LSP as follows: $\tilde{\chi}^+_1 \rightarrow W^± (\rightarrow l\nu)\tilde{\chi}^0_1$ and $\tilde{\chi}^0_2 \rightarrow Z (\rightarrow ll)\tilde{\chi}^0_1$. The summary of all electroweakinos ATLAS searches can be seen in Fig. 6.

![Figure 6](image-url)

Figure 6: The 95% confidence level exclusion limits on C1C1, C1N2, and N2N3 production with either SM-boson-mediated or slepton-mediated decays, as a function of the C1, N2, N3 and N1 masses. The production cross-section is for pure wino C1C1 and C1N2, and pure higgsino N2N3.

3 Conclusion

Direct SUSY and monojet searches, dark matter constraints, as well as Higgs and flavour physics data, provide complementary information on the pMSSM parameters. To highlight the interplay between the different observables, in this section we pre-impose the constraints from Higgs physics, flavour physics and relic density.
In Figure 7, we show the fraction of excluded pMSSM points in the neutralino scattering cross section with matter vs. neutralino mass parameter plane, considering simultaneously constraints from SUSY direct searches and monojet searches. The limits from LUX are also displayed in the same plane. We first see that the LHC searches are complementary to dark matter direct detection, and probe different parts of the parameter space. In particular, LHC probes regions well below the LUX limits, and monojet searches nicely increase the LHC discovery potential in the regions where the neutralino scattering cross section with matter is very small.

Figure 7: Fraction of pMSSM points excluded by the SUSY direct searches (left panel), in addition to the monojet searches (right panel), in the neutralino-proton scattering cross section vs. neutralino mass.

References


NEW PROJECTS ON DARK PHOTON SEARCH  

Venelin Kozhuharov  
*University of Sofia, 5 J. Bourchier Blvd, 1164 Sofia, Bulgaria*  
*INFN - LNF, Via E. Fermi 40, 00044 Frascati (RM), Italy*  
Mauro Raggi  
*INFN - LNF, Via E. Fermi 40, 00044 Frascati (RM), Italy*  
Paolo Valente  
*INFN - Sezione di Roma, P.le Aldo Moro, 2 - 00185 Roma - Italy*

Abstract

Despite the great success of the Standard Model of particle physics the nature of Dark Matter still remains unclear. Recently, the idea of the existence of a hidden sector coupling only weakly with the ordinary matter was revitalized and gained popularity. A simple mediator between the hidden and the visible sector could be a vector particle of a new gauge interaction, the so called dark photon. Numerous of activities were initiated to probe its parameter space. The present results and the foreseen experiments aimed to search for dark photons in few directions are reviewed and discussed.

1 Introduction

The Standard Model (SM) of particle physics provides a unique description of almost all phenomena in the microworld. However, the unexplained nature
of Dark Matter indicate the necessity of its extension to a more fundamental theory incorporating also the cosmological observations. A large variety of theoretical models exist.

From the experimental point of view the best path to follow is to address few of the observed “smoking guns” indicating a possible discrepancy between the SM prediction and the results from the data. At present there are still unexplained anomalies with the annually modulated excess of signal observed by DAMA/Libra \(^1\); the positron and antiproton excess in the cosmic rays \(^2\), \(^3\), \(^4\), \(^5\); the three sigma discrepancy in the anomalous magnetic moment of the muon \(^7\); the failure to explain the \(^8\)Be anomaly \(^6\) within nuclear physics. All those measurements might indication the presence of new undiscovered so far particles.

An elegant model to address the possible existence of new degrees of freedom is the concept of a dark sector (DS) of particles coupled only weakly with the Standard Model particles. The origin of the interactions could be due to the presence of a mediator - an object carrying both SM and DS quantum numbers or SM fields possessing a relatively small charge under any of the gauge symmetries in the Dark Sector. In that sense the mediator provides a so-called portal to the DS \(^8\), \(^9\). One of the simplest possible realization that is ultra violet safe (i.e. does not introduce new scale in the lagrangian) is by employing a vector gauge field, the so-called Dark Photon \(A'\) (DP), which interacts weakly with the SM fermions

\[
\mathcal{L} \sim g' q_f \bar{\psi}_f \gamma^\mu \psi_f A'_{\mu},
\]

where \(g'\) is the universal coupling constant and \(q_f\) are the corresponding fermion charges. The term in equation 1 could be effectively realized also through kinetic mixing of the massive DP with the ordinary photon. In this picture the interaction of the SM particles with the dark photon will be described by two parameters - \(\epsilon \sim g' q_f\) and the dark photon mass \(m_{A'}\). The coupling parameter \(\epsilon\) could be flavour dependent giving rise to different leptophobic or leptophilic DP models \(^10\).

Two different scenarios depending on the phenomenology in the dark sector could be identified. In the case when no new light degrees of freedom \(\chi\) in the dark sector exist the dark photon will decay to SM particles only, with \(\text{Br}(A' \rightarrow e^+ e^-) = 100\%\) for \(1\) MeV < \(m_{A'}\) < 210 MeV. If, however, \(m_{\chi} < m_{A'}/2\) then the decay \(A' \rightarrow \chi\chi\) will be dominant since it is not
suppressed by the small value of $\epsilon$. In the latter scenario the observables will depend on two additional parameters, the coupling strength in the dark sector $\alpha_D$ and $m_{\chi}$. These two scenarios, so called “visible” and “invisible”, result in very different experimental signature. In the visible case a narrow resonant might be observed in the dilepton or, in general, in the diparticle invariant mass spectrum, while in the invisible case the existence of a DP could present itself through the search for missing mass or missing energy.

2 Visible dark photon decays

Currently, most of the experiments addressing the existence of an $A'$ performed “peak” searches in the $e^+e^-$ invariant mass spectrum. This approach requires a precise spectrometer providing measurement of the electron momentum with high precision. The production of the $A'$ could be either through a bremsstrahlung process (in electron-on-target experiments) or through the $e^+e^- \rightarrow \gamma A'$ at electron-positron colliders.

Two new dedicated to the $A'$ search experiments are planned in the near future - the HPS experiment at the JLaB and the MAGIX experiment at Mainz. Both of them will exploit an electron beam impinging on a thin target.

The HPS experiment will measure the momentum of $A'$ products using a silicon vertex tracker placed inside a dipole magnet. A downstream lead tungstate calorimeter will serve for fast energy measurement and triggering. The silicon tracker is made of six dual sensor layers and will allow to address the dark photon decays in two different regions of its parameter space. For large values of $\epsilon^2$ ($\epsilon^2 > 10^{-7}$) $A'$ decays promptly and the event selection will be based on the $e^+e^-$ invariant mass reconstruction. The dominant background originates from the internal pair conversion of the bremsstrahlung photon into $e^+e^-$ pair. When $\epsilon^2 \leq 10^{-8}$ the finite lifetime of $A'$ will result in the reconstruction of vertices displaced from the interaction point inside the target. This channel is particularly interesting since it could be background free, at the expense of a low signal yield. The projected sensitivity for HPS is $\epsilon^2 < 3 - 4 \times 10^{-7}$ for $20$ MeV $\leq m_{A'} \leq 300$ MeV and $2 \times 10^{-8} \leq \epsilon^2 \leq 2 \times 10^{-10}$ in the region $20$ MeV $\leq m_{A'} \leq 200$ MeV.

The MAGIX experiment is planned to operate at the new energy-recovering superconducting accelerator at Mainz, MESA, which provides an electron beam.
with energy up to 155 MeV and 1 mA beam current. The target will be accomplished as a windowless supersonic gas jet with high density \((10^{19}/\text{cm}^2)\). The A’ decay products, \(e^+\) and \(e^-\), will be detected in a double arm spectrometer with planned resolution of \(\delta p/p = 10^{-4}\). The projected sensitivity indicates reach of \(\epsilon^2 \sim 10^{-8}\) for masses \(10\ \text{MeV} \leq m_{A'} \leq 50\ \text{MeV}\).

3 Invisible dark photon decays

The mass spectrum of the particles in the dark sector is in general ambiguous and nothing prevents the existence of light states. Such a scenario is relatively difficult to probe due to the impossibility to perform a complete reconstruction of the final state when \(A'\) decays to \(\chi\). Thus it is important to start with an initial state that can be fully described. The annihilation of \(e^+\) with \(e^-\) is one of the possibility while another is the usage of monoenergetic beam and search for missing energy taken away by the produced dark photons.

3.1 Missing mass technique

The missing mass technique relies on the complete reconstruction of the annihilation process

\[ e^+ + e^- \rightarrow \gamma A' \]  

through the measurement of the energy and the direction of the recoil photon. In positron on target collisions the missing mass squared is then computed as

\[ M_{\text{miss}}^2 = (P_e + P_{\text{beam}} - P_\gamma)^2 \]  

where usually the electron is considered to be at rest \((P_e = (m_e, 0, 0, 0))\). The cross section for the production of \(A'\) is enhanced with respect to the ordinary \(e^+e^- \rightarrow 2\gamma\) process by a factor \(\delta^{14}\), especially for \(m_{A'}\) close to the center of mass energy of the interaction. The dominant background processes are listed in table 1 and originate from the bremsstrahlung emission in the field of the target nuclei and from the \(e^+ + e^- \rightarrow \gamma\gamma\) annihilation. Usually, the measurement of the energy of the recoil photon provides enough resolution to suppress the \(e^+ + e^- \rightarrow \gamma\gamma\) background.

Three experiments, PADME at the DAΦNE Linac at LNF-INFN \(15\), MMAPS at Cornell \(18\), and at VEPP3 at Novosibirks \(17\), are planning to exploit this technique. They share similar design properties.
Table 1: Dominant background contributions to the missing mass technique

<table>
<thead>
<tr>
<th>Background process</th>
<th>( \sigma (E_{\text{beam}} = 550 \text{ MeV}) )</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>( e^+ e^- \rightarrow \gamma \gamma )</td>
<td>1.55 mb</td>
<td></td>
</tr>
<tr>
<td>( e^+ N \rightarrow e^+ N \gamma )</td>
<td>4000 mb</td>
<td>( E_{\gamma} &gt; 1 \text{MeV}, \text{on carbon} )</td>
</tr>
<tr>
<td>( e^+ e^- \rightarrow \gamma \gamma \gamma )</td>
<td>0.16 mb</td>
<td>( E_{\gamma} &gt; 1 \text{MeV}, \text{CalcHEP}^{16} )</td>
</tr>
<tr>
<td>( e^+ e^- \rightarrow e^+ e^- \gamma )</td>
<td>188 mb</td>
<td>( E_{\gamma} &gt; 1 \text{MeV}, \text{CalcHEP} )</td>
</tr>
</tbody>
</table>

PADME experiment at LNF-INFN, shown in fig. 1 will use 550 MeV positron impinging on a 100 \( \mu \)m thick active target made of policrystaline diamond. The recoil photons from the process \( e^+ + e^- \rightarrow \gamma A' \) will be detected by a ring-shaped BGO crystal calorimeter, located 3 m downstream providing energy and position information. The non-interacted positron beam will be deflected outside the acceptance of the calorimeter by a dipole magnet. Three sets of plastic scintillator detectors will serve to register the charged particles and will provide an efficient veto for the bremsstrahlung background. In addition, a Cherenkov detector, placed along the undeflected beam axis, will help with the suppression of the three photon annihilation background. The complete setup is located in vacuum to diminish the possible beam-residual gas interactions.

The approach of VEPP3 is to use an internal hydrogen gas target placed at the Novosibirsk storage ring, with 500 MeV \( e^+ \). Both the VEPP3 and the MMAPS experiments aim to use CsI crystals from CLEO for their calorimeters, which are shown to provide \( \sigma_E/E = 3\% \) for 180 MeV positrons. In order to be able to operate in parallel with the ongoing VEPP3 activities an extention of the existing beam line is proposed. This would allow to place a charged particle veto detector to suppress the background. MMAPS uses a berillium target and will profit from the higher beam energy - up to 5.3 GeV. This extends MMAPS sensitivity region to masses \( m_{A'} \sim 74 \text{ MeV} \).

A comparison of the characteristics and the performance of the PADME, MMAPS, and VEPP3 experiments is shown in table 2. PADME will be the first experiment to run and possible upgrades of the DAΦNE linac could extend its sensitivity down to the region of \( \epsilon^2 \sim 3 - 5 \times 10^{-7} \). The VEPP3 experiment requires an approval of beam line modification while the Cornell MMAPS experiment is in process of identification of finding.
3.2 Missing energy technique

In the shower development inside a calorimeter, the dark photons could be created through an A-strahlung process by any of the secondary particles. They could account for a large part of the undetected energy of the primary particle and their existence may manifest itself in the form of a large fraction of events with missing energy. This technique relies on the precise knowledge of the response of a calorimeter to an electromagnetic shower development. It is exploited by the recently approved NA64 experiment operating at CERN SPS 19). Synchrotron radiation tagged 100 GeV $e^-$ beam with momentum determined in a magnetic spectrometer impinges on an electromagnetic calorimeter (ECAL) acting as an active beam dump. Events with less than 50 GeV energy in the ECAL and no signal in the veto and in the downstream hadron calorimeter are considered as signal. In the hypothesis of zero background observation the NA64 experiment could set limits down to $\epsilon^2 \sim 10^{-8} - 10^{-6}$ for A' masses in the region 10 MeV - 100 MeV, assuming $10^{10}$ electron events.

3.3 Scattering of dark matter particles

Another possible indication of the presence of dark photons decaying into invisible particles $\chi$ could be the direct observation of the $\chi$ scattering in a low
Table 2: Comparison between the experiments exploiting the missing mass technique

<table>
<thead>
<tr>
<th></th>
<th>PADME</th>
<th>MMAPS</th>
<th>VEPP3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Place</td>
<td>LNF</td>
<td>Cornell</td>
<td>Novosibirsk</td>
</tr>
<tr>
<td>Beam energy</td>
<td>550 MeV</td>
<td>Up to 5.3 GeV</td>
<td>500 MeV</td>
</tr>
<tr>
<td>$M_{A'}$ limit</td>
<td>23 MeV</td>
<td>74 MeV</td>
<td>22 MeV</td>
</tr>
<tr>
<td>Target thickness [$e^{-}/cm^2$]</td>
<td>$2 \times 10^{22}$</td>
<td>O$(2 \times 10^{23})$</td>
<td>$5 \times 10^{15}$</td>
</tr>
<tr>
<td>Beam intensity</td>
<td>$8 \times 10^{-11}$ mA</td>
<td>$2.3 \times 10^{-6}$ mA</td>
<td>30 mA</td>
</tr>
<tr>
<td>$e^+e^- \rightarrow \gamma\gamma$ rate [s$^{-1}$]</td>
<td>15</td>
<td>$2.2 \times 10^6$</td>
<td>$1.5 \times 10^6$</td>
</tr>
<tr>
<td>$\epsilon^2$ limit (plateau)</td>
<td>$10^{-6}$</td>
<td>$10^{-6} - 10^{-7}$</td>
<td>$10^{-7}$</td>
</tr>
<tr>
<td>Time scale</td>
<td>2017-2018</td>
<td>?</td>
<td>2020 (ByPass)</td>
</tr>
<tr>
<td>Status</td>
<td>Approved</td>
<td>Not funded</td>
<td>Proposal</td>
</tr>
</tbody>
</table>

noise detector. This approach is similar to the beam dump technique widely used in the past to search for milicharged particles\(^9\). The $A'$ are produced, through $A'$-strahlung and then decay to $\chi\chi$ pair. The BDX experiment\(^{20}\), proposed to take place at JLaB, will use a CsI (TI) detector shielded by an active vetoing system to suppress the cosmogenic background. Studies indicated that the sensitivity down to $\epsilon^2 \sim 10^{-9}$ for $m_{A'} = 50$ MeV could be reached. However this depends on the extra parameters $m_\chi$ and $\alpha_D$.

4 Conclusions

Present searches for new vector particle $A'$ aim to cover the visible and the invisible final state scenarios in parallel. While on the visible side many results appeared in the last decade the parameter space for $A'$ decaying predominantly to non Standard Model particles is being addressed just recently. The new activities in this direction can be summarized in fig. 2. The PADME and the NA64 experiments are approved and are expected to provide interesting results by 2020, at the same time few unique projects are in preparation.

5 Acknowledgements

The authors would like to thank the organizers for the invitation to present at the Vulcano workshop. VK acknowledges support from INFN, LNF-INFN under contract LNF-SU 70-06-497/07-10-2014 and University of Sofia science...
Figure 2: Recent perspectives for invisible dark photon searches.

fund under grant FNI-SU N39/2016.

References

5. AMS-02 Collaboration, Talks at the “AMS Days at CERN”, 15-17 april 2015
9. T. Spadaro, these proceedings.
15. M. Raggi, V. Kozhuharov, P. Valente et al., The PADME experiment Technical Proposal.
also talk at “Dark Sectors Workshop”, 28-30 Apr., SLAC (2016).
20. M. Battaglieri et al. [BDX Collaboration], arXiv:1406.3028
M. Battaglieri et al. [BDX Collaboration], arXiv:1607.01390 [hep-ex].
NUCLEUS–NUCLEUS INTERACTIONS AND THEIR APPLICATION IN MEDICINE

G. Battistoni

INFN, Sezione di Milano, Italy

Abstract

The use of charged hadrons (protons and nuclei) in cancer therapy is one of the most successful cases of application of nuclear physics to medicine. The physical advantages in terms of precision and selectivity, combined with the biological properties of densely ionizing radiation, make charged particle therapy an elective choice in a number of cases. Hadron therapy is in continuous development and nuclear physicists can give important contributions to this discipline. The physics of proton–nucleus and nucleus–nucleus interactions plays a fundamental role but there are still important uncertainties. In this work some of the basic elements of Charged Particle Therapy will be summarized in connection to the relevant aspects of the underlying nuclear physics.
1 Introduction

Charged Particle Therapy (CPT in the following), or hadron therapy, is an innovative cancer radiotherapy based on nuclear particles (protons, neutrons and light ions) for treatment of early and advanced tumors \(^1\). The original proposal dates back to 1946, when Robert Wilson proposed the therapeutic use of protons for treating cancer \(^2\). Proton therapy has now become an advanced clinical modality, and CPT with heavier ions (generally \(^{12}\text{C}\)) is now becoming more and more attractive. The clinical interest in hadron therapy resides in the fact that it delivers precise treatment of tumors, exploiting the characteristic shape of the Bragg curve of charged hadrons, \(i.e.\) dose deposition as a function of depth of traversed matter, exhibiting a sharp peak (the Bragg peak) at the end of the particle range. As compared to the standard X-ray radiotherapy, accurate and efficient irradiation of the tumor can be obtained reducing the dose to the surrounding healthy tissues, thus achieving less complication probability. Especially for heavy ions, an increased biological effectiveness in killing cancer cells can also be obtained, making this approach very interesting in a number of cases and in particular for radio-resistant tumors. After a rather long period in which hadron treatments were exclusively delivered in research laboratories, today CPT has grown into an advanced, cutting-edge clinical modality. According to recent statistics \(^3\), more than 130,000 patients worldwide have been now treated with charged hadrons (about 10% with carbon ions), and the number of clinical centers dedicated to CPT is now rapidly increasing. Nuclear physics is still playing a fundamental role to help CPT to reach in practice the high level of precision which would be in principle attainable. The Nuclear Physics European Collaboration Committee has dedicated its 2014 report to the contribution of nuclear physics to medicine \(^4\) where a comprehensive review of the key issues in CPT can be found. Here we limit ourselves to a few selected issues. In Section 2 we summarize the basic principles of CPT, while Section 3 will be focused on the relevant nuclear physics for CPT. Section 4 will be dedicated to real time monitoring techniques based on the exploitation of nuclear interactions.
2 The Basic Facts of Charged Particle Therapy

Therapeutic beams are accelerated by cyclotrons (especially for proton therapy) or synchrotrons. The useful energy is determined by the amount of material that has to be penetrated in patients. Typically from 3 to 27 g/cm$^3$ and this roughly corresponds to 50 ÷ 250 MeV for protons and 60 ÷ 400 MeV/u for $^{12}$C ions. As an example, Fig. 1 shows the synchrotron in operation at CNAO (Pavia, Italy) and capable of accelerating different particles and nuclei, from protons to $^{16}$O.

![CNAO synchrotron](image)

Figure 1: The CNAO synchrotron.

CNAO is designed for a fully active dose distribution system. This means that the tumor is ideally divided into “slices”, i.e. in regions that are reached by particles of the same energy. The energy is varied by the synchrotron so as to choose the slice and within the slice the beam moves horizontally and vertically thanks to the finely laminated scanning magnets. Each slice is thus irradiated by “painting” it with a pencil beam. Typical currents can be as high as $10^{10}$ protons/s or $10^8$-$10^9$ $^{12}$C ions/s.

Charged particles loose energy primarily by inelastic collisions with the atomic electrons, resulting in ionization and atomic excitation, while the amount of energy lost due to Coulomb interactions with the material nuclei is instead very small. For charged particles other than electrons the mean ionization en-
ergy loss (or electronic stopping power) can be described by the Bethe-Bloch equation: the growing energy loss with decreasing particle velocity causes the characteristic Bragg peak. The Bragg peak is not perfectly sharp due to energy loss fluctuations (range straggling) and energy spread of the accelerated beams. An example of actual longitudinal dose deposition in water is given in Fig. 2, where we show a comparison of experimental data and simulation for different proton beam energies as measured at CNAO.

Figure 2: Comparison of experimental data (dots) and simulation (continuous lines) for ionization energy loss in water at different proton beam energies.

Highly ionizing particles, such as fully ionized nuclei, give rise to an energy deposition in matter with a higher spatial density with respect to protons (and photons). Most of the induced secondary electrons deposit the dose in the center of the primary tracks, within a typical radius of the order of nanometres. This leads to a larger probability of producing complex DNA damages, difficult to repair, resulting in an increased killing capability of cells. This is expressed by means of the concept of Relative Biological Effectiveness (RBE) which is defined as the ratio between the absorbed dose of a reference radiation (typically X-rays) and that of the test radiation (for example heavy ions) required to produce the same biological effect. Typically, RBE is determined considering the dose needed to achieve a 10% survival probability of cells in the irradiation. RBE is a very complicated radiobiological concept, which depends on several factors and variables. While in proton therapy a single RBE factor can be
applied throughout the entire radiation field (1.1, i.e. they are in average 10% more effective than photons), the situation for the mixed field in heavier ion therapy is much more complex, since RBE is ion-dependent and varies along the ion path.

3 Nuclear Physics and Particle Therapy

Several nuclear processes are relevant in hadron therapy. Inelastic interactions are responsible for beam attenuation along the longitudinal profile, while elastic scattering, especially in the case of proton therapy, contributes to the transversal profile of dose distribution. Fragmentation of both projectile and target is probably one of the most relevant processes to be studied in detail, since it affects the attenuation of primary beam and the biological effect. Indeed, the most frequently occurring nuclear reactions are peripheral collisions where both beam particles and target may lose one or several nucleons or clusters of nucleons. Those emitted from the projectile fragmentation appear forward peaked in the laboratory frame due to the high velocity of the projectile. The projectile-like fragments continue traveling with nearly the same velocity and direction, and contribute to the dose deposition until they are completely slowed down or undergo further nuclear reactions. Neutrons and clusters from target-like fragments are emitted isotropically and with much lower velocities. The particles ablated from the fireball cover the range between the projectile and target emission. Nuclear fragmentation reactions lead to an attenuation of the primary beam flux and the build-up of lower–Z fragments with increasing penetration depth. These fragments are responsible of the tail beyond the Bragg peak, as shown in Fig.3.

These lighter fragments in general have a RBE factor different from that of primaries, even for the same delivered dose, and Treatment Planning programs for heavy-ion therapy must take into account these effects and their validation against experimental data is mandatory. In order to study such secondary fragments, many experiments have been performed with thick targets made of water or tissue-equivalent materials. The most recent measurements of this kind were performed for carbon ion collisions with water (12, 13, 14, 15). Devoted experiments aimed to measure nuclear cross-sections have also been designed (16).

Accurate modeling of all the mentioned processes is one of the most im-
Figure 3: Bragg curve as a function of depth in water for a 400 MeV/n carbon beam. The points represent experimental data and the solid line represents Monte Carlo calculation (FLUKA code 8, 9, 10). The calculated dose contribution from primary $^{12}$C ions and secondary fragments is shown 11.

important contributions of nuclear physicists to the discipline of CPT. The level of accuracy that present Monte Carlo codes exhibit in the description of nuclear reaction models and the level of accuracy in the comparison of predictions with experimental data are encouraging, but there is still ample room for improvement. However, the amount of available experimental data is not enough to provide a complete benchmarking.

4 Nuclear Interactions and Real Time Monitoring

There are uncertainties on the position of the dose release in CPT treatments which are due to different factors, such as quality and calibration of the Computed Tomography (CT) images or possible morphologic changes occurring between CT and each of the several irradiation sessions, operated in different days, that compose a treatment in CPT. Finally also patient mis-positioning and organ motion during the treatment itself can be sources of uncertainty. All these effects can add up to give an overall uncertainty of the order of few millimeters 18). A real time monitoring procedure can therefore increase the quality assurance of a CPT treatment. Nuclear reactions experienced by the
primary and its possible fragments can be exploited to achieve the goal of in-vivo range monitoring. A discussion of range verification methods and of related physics can be found in the literature 19). Three are the main nuclear processes that can yield a radiation suited for this purpose: production of $\beta^+$ emitters nuclei, gamma de-excitation of nuclei and charged particle production in inelastic interactions. A sketch of the possible useful reactions is shown in Fig. 4. In order to make use of these processes the comparison of measured and pre-calculated distributions of secondary particles is needed. This is a further motivation to push for a continuous upgrade of available Monte Carlo models.

![Diagram of nucleon-nucleus reaction](image)

Figure 4: Sketch of possible nucleon-nucleus reaction in proton therapy (above) and in ion therapy (bottom 19).

5 Conclusions

The contribution of nuclear physics to the development of CPT is still fundamental. The study of nuclear fragmentation is one of the most important issues and further work is needed both at experimental level and in model development. Furthermore new nuclear species are under study as therapeutic beams, $^4$He and $^{16}$O being the most important 20). Important developments are in progress also on the technological side in order to build specific particle detection systems to be used for real time range monitoring.
References

2. R.R. Wilson, Radiology, 47 487 (1946).
3. http://ptcog.web.psi.ch/ Particle Therapy Co-Operative Group (PTCOG)
4. NuPECC, Nuclear Physics for Medicine, NuPECC Report (2013), ESF.
RECENT HIGGS BOSON RESULTS FROM THE LHC

Roberto Di Nardo

CERN, Rt de Meyrin CH-1211 Geneve 23 Switzerland

Abstract

A collection of Higgs boson property measurements from ATLAS and CMS collaboration based on LHC Run1 dataset is presented. In addition, recent results on Higgs physics from LHC Run2 are summarized.

1 Introduction

The ATLAS and CMS collaborations announced on July 4th 2012 the discovery of a new particle in the SM Higgs searches. The properties of this new particle have been studied by both collaborations, confirming its identity as a Higgs boson with a mass of about $m_H = 125$ GeV. Its decays, production modes and couplings have been studied, and its spin and parity properties have been tested and compared with SM expectations. In the SM, the Higgs boson is produced at LHC predominantly via gluon-fusion process ($\sigma \sim 19.5$ pb at $\sqrt{s} = 8$
TeV) followed by vector boson fusion (σ ∼ 1.6 pb at √s = 8 TeV), associated production with a vector boson (σ ∼ 1.1 pb at √s = 8 TeV) and associated production with a t\bar{t} pair. At m_H = 125 GeV, the main Higgs boson decay channels are into b\bar{b} (branching fraction 57.7%), τ\tau (6.3%), WW∗ (21.5%), ZZ∗ (2.6%) and γγ (0.23%). The production cross sections increase by about a factor 2.2 at √s = 13 TeV compared to the 8 TeV, with exception of the associate production with a t\bar{t} pair, which experience a gain of almost a factor 4.

2 Run1 Highlights

2.1 Higgs boson mass and width measurements

Both ATLAS and CMS performed a mass measurement of the Higgs boson exploiting the channels with the best mass resolution (1-2%): the H → γγ and H → ZZ∗ → 4\ell(\ell = e, \mu). The mass measurements of each experiment as well as the combination, are shown in Figure 1. The combined mass of the Higgs boson is m_H = 125.09 ± 0.21(stat) ± 0.11(syst) 1). The dominant systematic uncertainty is related to the scale uncertainty on the electron, photons and muons. The compatibility of the four measurements is tested using the likelihood ratio and the resulting p−value is 10%.

In the SM, the predicted width for a m_H = 125 GeV Higgs boson is Γ_H = 4.07 MeV. A direct upper limit of Γ_H < 1.7 GeV at 95% CL on the total width

Figure 1: Summary of Higgs boson mass measurements from the individual analyses of ATLAS and CMS and from the combined analysis.

as the combination, are shown in Figure 1. The combined mass of the Higgs boson is m_H = 125.09 ± 0.21(stat) ± 0.11(syst) 1). The dominant systematic uncertainty is related to the scale uncertainty on the electron, photons and muons. The compatibility of the four measurements is tested using the likelihood ratio and the resulting p−value is 10%.

In the SM, the predicted width for a m_H = 125 GeV Higgs boson is Γ_H = 4.07 MeV. A direct upper limit of Γ_H < 1.7 GeV at 95% CL on the total width
has been set by CMS using the invariant mass distribution of the diphoton and four leptons final states. CMS also set a lower limit of $\Gamma_H > 3.5 \cdot 10^{-9}$ MeV at 95% CL by measuring the lifetime of the Higgs boson in the ZZ channel. An indirect constraint can be obtained by the ratio between on-shell and off-shell production of the Higgs boson under the strong assumption that the ratio of coupling constants remains invariant at the low and high $m_{VV}$ values. The indirect limit from ATLAS is $\Gamma_H < 22.7$ MeV (33 MeV expected) at 95% CL 2) while from CMS is $\Gamma_H < 13$ MeV (26 MeV expected) at 95% CL 3).

2.2 Measurement of the coupling properties of the Higgs boson

Both the experiments published at the end of Run1 analyses targeting to the measurement of the Higgs boson production rate in several combinations of production and decay modes 4, 5). All these measurements have been combined together to probe the coupling properties of the Higgs boson 6). More than 600 experimental categories contribute to the combination. Each category can receive contributions from different processes and can bring information about different couplings. Different signal strength parametrizations can be chosen with the limitation that single analysis channels are always only sensitive to the product of production cross sections and the decay branching ratios. Only combining them, the different production and decay modes can be partially disentangled. The signal strength modifier parametrization measure the observed signal yield relative to the SM expectation and it is defined as $\mu_i^f = (\sigma_i \cdot BR^f)/(\sigma_{i,SM} \cdot BR_{SM}^f) = \mu_i \cdot \mu^f$ where $\mu_i$ and $\mu^f$ are multiplicative scale factors to the SM expectations from signal events produced via process $i$ and decaying to final state $f$. Figure 2 shows the results of fits using two signal strength parametrizations, in which either all production cross sections are fixed to their SM values 2(a) or alternatively all decay BRs are fixed to their SM value 2(b). The compatibility with the SM expectations is at the level of 60% and 24% respectively. Under the above assumptions, a significance of more than 5$\sigma$ is obtained for the VBF production mode and $\tau\tau$ decay, and the evidence of associate production of the Higgs boson with a vector boson ($W$ or $Z$) is also observed. It is also possible to scale all production and decay modes with a single signal strength parameter $\mu$ and, in this case, the combined ATLAS-CMS measurement is

$$\mu = 1.09 \pm 0.07(stat)^{+0.09}_{-0.08}(syst)$$  (1)
Figure 2: Best fit results for the production (a) and decay (b) signal strengths for the combination of ATLAS and CMS data. The results from each experiment are also shown. The error bars indicate the 1σ (thick lines) and 2σ (thin lines) intervals. (c) Fit results for two parametrizations allowing BSM loop couplings (κV ≤ 1 or BR_{BSM}=0).

In addition, the results can be interpreted in a LO framework 7) (κ framework) in which coupling modifiers are introduced to parametrize possible deviations from the SM predictions of the Higgs boson couplings to SM bosons and fermions. Since the width of the Higgs boson is unknown, an additional assumption needs to be done. Two possible constraints are either κV ≤ 1 or BR_{BSM}=0. The former is satisfied by several theoretical extensions of the SM while the latter does not allow for invisible or undetected decays or modifications to BR which are not measured directly. The two results are shown in figure 2(c).

2.3 Spin and Parity studies

In the SM the Higgs boson has spin zero and even CP value (J^{PC} = 0^{++}) and a deviation from this will be a sign of BSM physics. The J = 1 is excluded from the decay in two photons. Various spin and parity models have been studied by using the angular and kinematic distributions in decays of the Higgs to dibosons without taking into account the measured rates to minimize the model dependence of the results. The ratio of profiled likelihoods for the SM and the alternative hypothesis is used as test statistics, its distribution for the SM and alternative hypothesis is evaluated on pseudo-experiments and the CLs is used
to assess the level of exclusion of the alternative model. The analyses combine the information from $H \rightarrow ZZ^* \rightarrow 4\ell$ and $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$, and in the case of the ATLAS experiment also $H \rightarrow \gamma\gamma$. Figure 4 shows the distribution of the test statistics for the SM Higgs boson and alternative models for ATLAS and CMS. The alternative non-SM hypotheses with are excluded at more than 99% confidence level.

3 First results with 2015 LHC Run2 collision data.

During 2015 both ATLAS and CMS experiments collected their first data at $\sqrt{s} = 13$ TeV, 3.2 fb$^{-1}$ and 2.8 fb$^{-1}$ respectively. Given the still limited amount of data, the sensitivities for the SM Higgs boson analyses were still lower with respect to Run1. On the other hand, the searches for new heavy particles started to be already competitive with respect to the Run1 results thanks to the gain in the parton luminosities.

3.1 SM Higgs into dibosons

Figure 4(a) shows the Higgs boson cross-section measurement performed by ATLAS as function of the center of mass energy of the LHC proton proton collisions obtained with a combination of the $ZZ$ and $\gamma\gamma$ final states. The statistical uncertainty on the 13 TeV measurements is still the dominant uncertainty. A 3.4$\sigma$ significant measurement (assuming SM yields) was expected but...
the observation yield to 1.4σ. The compatibility with the SM has been quantified at the 1.3σ level. Also CMS performed first measurements using collision data at √s = 13 TeV. Figure 4(b) shows the distribution of m_4l measured by CMS while figure 4(c) the fiducial cross section measurement performed in the H → ZZ* → 4ℓ channel as a function of √s (11).

3.2 Heavy resonance searches in the diboson final states

Several ATLAS and CMS analyses have been designed to exploit the big potential of the first LHC collisions at √s = 13 TeV in the search for BSM physics in the diboson final states like: sequential Standard Model (Z' and W', J = 1), Randall-Sundrum graviton (RS G*, J = 2), Bulk RS graviton (Bulk G*, J = 2), HVT Model, extended Higgs sectors (J = 0). The strategy common to most of the analyses and consists in the search for an excess in the m_{VV} spectra over a smooth background, which is usually fit using functions or predicted with MC simulations. In addition, for resonances decaying into W and Z bosons, boosted topologies of the jets coming from the hadronic vector boson decays are extensively investigated. In all the searches, no significant excess was found, with the exception of the resonance search in the γγ final state. ATLAS experiment observed the largest deviation from background only hypothesis at about m_{γγ} = 750 GeV with global (local) significance of 2.0σ(3.9σ) for a J = 0 hypothesis with Γ_X=45 GeV and of 1.8σ (3.6σ) for a J = 2 hy-

Figure 4: Combined H → ZZ* → 4ℓ and H → γγ cross section measured by ATLAS as function of √s. (b) CMS distribution of the four-lepton invariant mass measured with 2.8 fb⁻¹ of data at √s = 13 TeV and (c) the measured fiducial cross section as a function of √s in the H → ZZ* → 4ℓ channel.
The compatibility with the background-only hypothesis as a function of the assumed mass $m_X$ and relative width $\Gamma_X/m_X$ for the ATLAS analysis, optimized for a spin-0 resonance search, is shown in Figure 5(a). CMS experiment observed the largest deviation from background only hypothesis at about $m_{\gamma\gamma} = 760$ GeV with global (local) significance of $< 1.0 \sigma$ (2.8\(\sigma\) for $J = 0$ and 2.9\(\sigma\) for $J = 2$) with a preference for a narrow signal $^{13}$). These results are shown in figure 5.

Figure 5: (a) Compatibility with the background-only hypothesis as a function of the assumed mass $m_X$ and relative width $\Gamma_X/m_X$ for the ATLAS analysis optimized for a spin-0 resonance search. (b) Observed $m_{\gamma\gamma}$ spectra by CMS for the events with both photons in the ECAL barrel detector. (c) Observed CMS background-only $p$-value for spin-0 narrow resonances as a function of $m_{\gamma\gamma}$ from the combined analysis of the 8 and 13 TeV data. The results for the 8 and 13 TeV data sets are also shown separately.

4 Conclusions

The LHC Run 1 physics program has been very successful and both ATLAS and CMS started their era of the Higgs boson properties measurements after its discovery in summer 2012. The results of these studies, performed by the two experiments using about 25 fb$^{-1}$ of proton-proton collision data at $\sqrt{s} = 7$ and 8 TeV, are consistent with the Standard Model expectations within the uncertainties.

The LHC Run2 started in 2015 at $\sqrt{s} = 13$ TeV and about 3 fb$^{-1}$ have been delivered to ATLAS and CMS. Preliminary results on the Higgs boson physics
at this center of mass energy have been obtained by the two experiments and, with the $\sim 30 \text{ fb}^{-1}$ expected at the end of 2016, it will be possible to improve the precision of most of the current Higgs boson related measurements. Although no high-mass Higgs boson nor other exotic resonance has been found yet, the collected luminosity at $\sqrt{s} = 13$ TeV during 2015 allows already to set limits more stringent than those obtained in Run1. The 750 GeV mild excess in the $\gamma\gamma$ channel is quite intriguing and, already with the data expected by August 2016, ATLAS and CMS experiments will be able to determine the origin of this excess observed with 2015 data.

References

6. G. Aad et al. [ATLAS and CMS Collaborations], JHEP 1608 (2016) 045
10. G. Aad et al. [ATLAS Collaboration], ATLAS-CONF-2015-069
11. V. Khachatryan et al. [CMS Collaboration], CMS-PAS-HIG-15-004
12. M. Aaboud et al. [ATLAS Collaboration], JHEP 1609 (2016) 001
FLAVOR PERSPECTIVES WITH THE LHCb UPGRADE

P. Campana - Laboratori Nazionali dell’INFN, Frascati (Italy)

Abstract

The completion of Run 1 of the CERN Large Hadron Collider has seen an unprecedented number of precise measurements done by LHCb in the flavour sector. The performance of a LHCb upgraded detector is discussed, together with the experimental challenges and the physics reach opportunities.

1 Introduction

The primary goal of the experiments at the CERN Large Hadron Collider (LHC) is to answer fundamental questions in particle physics. During Run 1, focussing on flavour physics in the forward region, the LHCb experiment operated at lower instantaneous luminosity and collected a total of 3 fb$^{-1}$. LHCb has proven to be a true general purpose forward spectrometer, with measurements of the $W$, $Z$, and top cross sections that will provide an additional
handle in reducing PDF uncertainties. However, it is in the domain of flavour physics where the experiment made the most significant impact. A highlight is the observation of the extremely rare $B_s^0 \rightarrow \mu\mu$ decay at 6σ significance in a combined CMS and LHCb analysis 1). In the CP violation sector, LHCb has measured the CKM angle $\gamma$ with the best single experiment precision, obtained from time integrated analyses of $B^+ \rightarrow Dh^+$ and $B^0 \rightarrow DK^{*0}$ decays 3). The CP violating phase $\phi_s$ in $B_s \rightarrow J/\psi K^+K^-$ and $B_s \rightarrow J/\psi\pi^+\pi^-$ processes has been determined 4). CP violation in the charm sector has been probed to the level of $10^{-4}$ 8), together with a precise measurement of the $D^0\bar{D}^0$ mixing parameters 9).

With limited statistical significance, LHCb has observed anomalies with respect to SM predictions, that attracted theoretical interest. These are the angular observables in $B^0 \rightarrow K^*\mu\mu$ decays 2), the ratio of $B^0 \rightarrow D^*\tau\nu$ to $B^0 \rightarrow D^*\mu\nu$ decays 5), and the exclusive determination of $|V_{ub}|$ CKM matrix element using the $\Lambda_b \rightarrow p\mu\nu$ decay 6). Further theoretical and experimental efforts are needed to confirm or disprove these effects.

LHCb measurements of charm and onia central exclusive production in the forward region have been made, and spectroscopy studies have proved fruitful. Two new penta-quark charmonium resonances, $P_c^+(4450)$ and $P_c^+(4380)$, were unambiguously identified 10), the resonance nature of the tetra-quark $Z(4430)$ 11), and the quantum numbers of the $X(3872)$ particle determined 12).

2 LHCb Detector Upgrade

LHCb has planned an upgrade for 2021 15), with two crucial improvements: the capability to exploit a 5 times higher instantaneous luminosity ($\sim 2 \times 10^{33}$ cm$^{-2}$ s$^{-1}$), and a readout to acquire events at the LHC bunch crossing rate ($\sim 40$ MHz). The upgrades enable the improvement of the precision of most of its key measurements substantially during LHC Run 3 by the end of 2023, planning to integrate $\sim 50$ fb$^{-1}$ by 2030. This upgrade does not need HL-LHC, as the machine already allows LHCb to increase the peak luminosity. Minimal modifications are needed to the LHCb interaction region to be compliant with the HL-LHC operation. The collaboration is preparing to upgrade the detectors and the readout architecture to enable the transition of the trigger rate from the current 1 MHz to 40 MHz. A general layout of LHCb detector upgrades is...
reported in Figure 1.

Events at a pile-up of $\sim 6$ are very crowded and contain hundreds of particles from several primary vertices in the envelope of the effective LHCb pp collision zone which is a few cm long. A high granularity vertex detector (VELO) is required to handle these events. Therefore the current VELO will be replaced 16), with $50 \times 50 \mu m^2$ silicon pixels to provide a full 3D reconstruction of tracks in space, and to limit the fake tracks.

The driving choices for the upgrade of the tracking system 17) are a uniform and high granular readout over a large area ($\sim 100 \text{ m}^2$ per layer), a fast signal formation, and radiation hardness. A new approach, applied on a large scale for the first time, will be used: 2.5 m long modules made of 0.25 mm diameter scintillating fibres along the $y$ vertical direction, with an $x$ coordinate resolution of $\sim 70 \mu m$. Fibres show excellent timing resolution and signals are contained in 25 ns. Fibre readout is performed with Silicon Photo Multipliers (SiPMs), sensitive to a single photon, that have a very high quantum efficiency. The temperature and voltage of the SiPMs must be controlled for a stable operation, while low temperature conditions must be met (below $-40^\circ C$) to mitigate radiation damage.

A large data rate will be produced by the front-end electronics of the upgraded LHCb, up to 3 TB/s. To reduce the cost of the DAQ and computing infrastructures, a mix of state-of-the art technology solutions are under development 19): new generation of FPGA with enhanced computing and memory resources for the DAQ system, high rate network to provide bandwidth for data collection, and an Event Filter Farm (EFF) with massive processing capacity provided also by graphic processing units (GPU). The goal is to convey all the information from sub-detectors related to the same event into a processing unit, where the High Level Trigger (HLT) software will analyse and select the events. Due to the large cross sections and the high trigger efficiencies, the output data rate will be significant. Data will be written to disk at a rate of 2 to 5 GB/s, with an average event size of $\sim 100$ kB, which is challenging in term of computing resources for end user analysis. The collaboration is studying and testing new offline models in Run 2, also considering future developments in information technology.

Other systems 18) are refurbishing or modifying their layout to comply with the 40 MHz readout. The calorimeter and the muon systems will keep the
current detectors but will upgrade part of the electronics, the silicon tracking system located between the vertex detector and the magnet will have new upgraded silicon strip sensors with a better granularity, while the two RICH detectors will replace the photon sensors with multi-anode PMs with a new front-end electronics.

3 LHCb Detector Performance with High Luminosity

The reconstruction of vertices, the tracking, the particle identification, and the triggering at a quality comparable with Run 1 and 2, represent the main challenges for LHCb at high luminosity.

The VELO reconstructs displaced vertices, and plays a significant role in the trigger and tracking. At trigger level it identifies tracks with high impact parameter (IP) and represents the first handle to reduce the background from minimum bias events. An excellent secondary vertex resolution is vital to resolve the fast \( B_s^0 \) oscillations. Key performance parameters include the occupancy, the hit resolution, the track reconstruction efficiency and the decay time resolution. The occupancy in the pixel detectors reaches a maximum of 0.12 % at the innermost sensor. The single hit resolution is \( \sim 12 \mu \text{m} \), and the
pattern recognition has a reconstruction efficiency of $\sim 99\%$ for tracks with $p > 5$ GeV. The 3D tracking capability of a pixelated layout is superior to the current system. The IP resolution has a strong dependence on the $p_T$ of the track, while the primary vertex (PV) position resolution depends on the number of tracks belonging to the vertex. The upgrade performance is very similar to the present one. As a figure of merit, at $p_T \sim 1$ GeV a 3-dimensional IP resolution of 50 $\mu$m is obtained, see Figure 2 (left), while standard deviations of 10 and 50 $\mu$m are observed in the $x$ and $z$ coordinate of the PV resolution, respectively. A decay time resolution of $\sim 40$ fs is obtained for the $B_0^\pm \rightarrow J/\psi KK$ channel, equivalent to the current performance. The VELO is radiation hard up to 50 fb$^{-1}$, without visible effects on performance.

The most valuable tracks for physics analysis are those (defined as *long*) which are reconstructed in the VELO and in detectors before (UT stations) and after (T stations) the magnet. *Long* tracks have excellent spatial resolution and a precise momentum measurement due to the combined information of the track slopes before and after the magnet. Moreover, there are tracks (*downstream*) measured only in the UT and in T stations, reconstructing long-lived particles which decay outside the VELO ($K^0_S$ mesons and $\Lambda$ baryons). In the upgrade, the track reconstruction efficiency at high luminosity is evaluated to be 99.6% for *long* tracks with $p > 5$ GeV, a value nearly identical to the current one. Thanks to the 3D reconstruction in the VELO, the ghost rate is 2.5%, reduced by a factor 2 with respect to the current one. For *downstream* tracks the situation is currently less satisfactory: the reconstruction efficiency in the upgraded detector is lower by about 10% (in absolute value) with respect to the current one and a higher ghost rate is observed. The lack of $y$-segmentation in the fibre tracker is responsible for a part of the inefficiency, while the new algorithm is not yet optimal. A re-design of the inner zone is currently being investigated, introducing a $y$-segmentation in a small area around the beam pipe, where the inefficiency is localized. The momentum resolution of the upgraded LHCb tracking system is very similar ($dp/p \sim 0.004$) to the current one, see Figure 2 (right), and constant over a large momentum interval (up to $p \sim 50$ GeV).

Particle identification (PID) is a critical aspect of the LHCb upgrade. Most of the physics reach is determined by the identification of hadrons, muons, electrons and photons. The effort has been focused on maintaining or ame-
iorating the performance at high luminosity. The detectors involved in this challenge are the RICH, the CALO and the MUON systems.

The RICH is essential for the study of hadronic final states, precision measurements of CP violation and rare decays of $b$ and $c$ hadrons. Multibody final states such as $B^0_s \rightarrow \phi \phi$, $B^0$ and $B^0_s$ decays to $K^+K^-$ and $\pi^+\pi^-$ would have severe background without $K$ separation. Moreover, PID is crucial in $b$ tagging and also in the identification of protons from decays of heavy baryonic states. The changes of geometry (RICH1) and of photon sensor (RICH1 and 2) provide better yield and angular resolution. The occupancy determines the PID quality, and the most illuminated region of RICH1 shows a manageable 24%, which is 30% smaller than achieved with the old geometry at high luminosity. The PID performance is defined in terms of the efficiency of identifying a true $K$ as a $K$ and the mis-identification probability of a true $\pi$ to be tagged as a $K$ or heavier particle. Simulations give a 2 % $\pi$ mis-identification efficiency, when a 90 % $K$ efficiency is requested, comparable with the current performance. The higher luminosity will lead to a degradation of energy and position measurements for calorimeter objects and to an increase of mis-identification for muons. To mitigate these effects, new algorithms have been developed. The inner muon detectors will be better protected from particles originating from

Figure 2: (Left) 3D resolution for the impact parameter (IP) as a function of $p_{T}^{-1}$, for upgraded (red squares) and current (black squares) VELO layouts. The light grey histogram shows the relative population of $b$-hadron daughter tracks. (Right) Momentum resolution as a function of $p$ for upgraded (blue dots) and current (black dots) tracking layouts.
the beam pipe by increasing the shielding. In the upgrade, at a selection efficiency of 90%, mis-identification values of $\sim 1\%$ and of $\sim 3\%$ are obtained, respectively, for electrons and muons with $p > 10$ GeV, close to the current performance.

The acquisition and the trigger-less software processing of the 40 MHz rate events from LHC represent the keystone of the entire LHCb upgrade. The limitations in CPU budget, determined by the size of the Event Filter Farm (EFF), and the output bandwidth, which is constrained by online computing resources, represent the challenges. The all-software trigger offers maximal flexibility in designing selections, lowers the $p_T$ thresholds in hadronic events, increasing the efficiency for multi-body final states and for charm events. The gain in efficiency has been studied for several channels and depends strongly on the allowed bandwidth; the lower the bandwidth, the more stringent must be the cuts to stay within allowed limits. A conservative approach is adopted, with an accepted HLT output at 25 kHz. An example of a channel with boosted efficiency is $B_s \to \phi \phi$, with an increase of a factor 4 in yield. Considerable gains are obtained for semi-leptonic $B$ decays, up to $\sim 50\%$, depending on the charm decay final state configuration, for three-body charmless $B$ decays and for hadronic open charm $B$ decays (e.g. $B^+ \to D^0[K_s\pi\pi]K^+$), up to a factor of 2 to 3.

4 Physics Prospects

B factories and LHC data demonstrated the success of the Cabibbo-Kobayashi-Maskawa (CKM) paradigm. All the measurements agree in a highly precise and profound way and the quark flavour sector is well described by the CKM mechanism. The gain in statistical precision from high luminosity operation will impact a large set of variables, measured by the dedicated flavour experiment, LHCb. LHCb data from Run 2 ($\sim 5$ fb$^{-1}$), will increase the sample by a factor 4 with respect to the present one, and should confirm or discard the anomalies mentioned above.

The decays $B_{d,s} \to \mu \mu$ are a special case amongst the electroweak penguin processes, as they are chirality-suppressed in the SM and are most sensitive to scalar and pseudoscalar operators. Therefore, in the SM these decays are exceedingly rare, with predicted decay rates for $B^0_s$ and $B^0$ of $3.6 \times 10^{-9}$ and $1.1 \times 10^{-10}$, respectively, and $\sim 10\%$ theoretical uncertainty 20). New
physics can produce significant enhancements in the decay rates. The ratio of branching fractions for $B^0_s$ and $B^0$ is known with better theoretical uncertainty and also provides a stringent test of Minimal Flavour Violation (MFV) models. In Run 1 CMS and LHCb have observed the decay rate of the $B^0_s \rightarrow \mu \mu$ channel finding it in good agreement with the SM, and significantly restricting the available phase space for BSM theories, in particular for MSSM models. Prospects for more precise measurements are interesting, although they will still be dominated by the experimental uncertainty. Studies have been performed for LHCb \cite{23}, showing that the determination of the ratio of branching ratios has 35% accuracy with $\sim 50$ fb$^{-1}$.

The CP-violating phase $\phi_s$ arises in the interference between the amplitudes of $B^0_s$ mesons decaying via the $b \rightarrow c\pi s$ tree diagram to CP eigenstates directly and those decaying after oscillation. In the SM, a global fit to experimental data leads to an expected value of $\phi_s = (0.0363 \pm 0.0013)$ rad, see Figure 3 (right). Non-SM particles could contribute to $B^0_s$ oscillations and a measurement of $\phi_s$ different from the prediction would provide unambiguous evidence for new physics. The $B^0_s \rightarrow J/\psi KK$ and $B^0_s \rightarrow J/\psi \pi \pi$ decays are the channels providing the best sensitivity and at high luminosity the final experimental precision should allow the observation of changes as small as a factor of two with respect to the SM with 3 $\sigma$ significance. The current LHCb experimental uncertainty on $\phi_s$ is 0.050 rad and an extrapolation to 50 fb$^{-1}$ shows that a value of 0.008 rad can be reached.

Another important example for the improvement in precision is the determination of the angle $\gamma$, which is one of the least well known parameters.
of the CKM matrix. It has negligible theoretical uncertainties (∼10⁻⁶) and serves as the reference point for comparison with γ value obtained from loop decays (like Δm_d and sin 2β) and to improve the precision of global CKM fits and the corresponding limits on new physics contributions. Combining several independent decay modes is the key to achieving the ultimate precision, with time-independent channels (B → DK) or with time-dependent ones (B_s → D_s K), involving only tree amplitudes. Further information can come from charmless decays, although penguin pollution makes theoretical control more difficult. Given its particle identification capabilities, LHCb is the unique experiment at the LHC which can measure γ, and at high-luminosity a global fit over the various channels could provide an estimated precision below 1°. BELLE II at KEKB with a data sample of 50 ab⁻¹ could provide a comparable accuracy.

5 Conclusion

An exciting program of measurements is planned with Run 3 (2021-2023), that will see the start of the data taking of an upgraded LHCb detector operating at an instantaneous luminosity a factor 5 higher than in previous runs. The start of HL-LHC operation is foreseen for 2027 and the LHCb experiment will be ready to take data also during high luminosity runs.

6 Acknowledgments

The author wholeheartedly thanks the organizers of the Conference for the kind invitation and for providing an outstanding nice location.

References

5. LHCb Collaboration, PRL 115 (2015) 111803
7. LHCb Collaboration, PRL 113 (2014) 151601
8. LHCb Collaboration, JHEP 04 (2015) 043
9. LHCb Collaboration, PRL 111 (2013) 251801
23. LHCb Collaboration and Bharucha A et al., EPJ C73 (2013) 2373.