



FACULTAD DE CIENCIAS

**Search for dark matter production in
association with top quarks in the dilepton
final state at $\sqrt{s} = 13$ TeV**

A thesis submitted in fulfilment of the requirements for the

Degree of Doctor of Philosophy

Written by

Cédric Prieëls

Under the supervision of

Jónatan Piedra Gómez

Pablo Martínez Ruiz del Árbol

June 2020

Abstract

Resumen

Acknowledgements

Acronyms used

SM Standard Model

DM Dark Matter

LHC Large Hadron Collider

CMS Compact Muon Solenoid

ATLAS A Toroidal LHC ApparatuS

CERN European Council for Nuclear Research

CMB Cosmic Microwave Background

ML Machine Learning

MFV Minimal Flavour Violation

WIMP Weakly Interactive Massive Particles

PF Particle Flow

BSM Beyond the Standard Model

Contents

1	Introduction	1
2	The Dark Matter case	4
2.1	At the origins of dark matter	4
2.1.1	Zwicky and the virial theorem	4
2.1.2	Spiral galaxies rotation curves	5
2.1.3	Cosmic Microwave Background (CMB) anisotropies	6
2.1.4	Other evidences	8
2.2	Dark matter properties	8
2.3	Dark matter candidates	8
2.4	Dark matter searches	8
2.5	Production at the LHC	8
2.5.1	The single top production channel	8
2.5.2	The $t\bar{t}$ production channel	8
2.6	Previous results	8
3	The experimental device	9
3.1	The LHC accelerator	9
3.2	The CMS detector	9
3.2.1	Tracker	9
3.2.2	Electromagnetic calorimeter	9
3.2.3	Hadronic calorimeter	9
3.2.4	Muon system	9
3.2.5	Trigger	9
3.2.6	Data aquisition	9
4	Objects reconstruction	10
4.1	Particle Flow algorithm	10
4.2	Leptons reconstruction	10
4.2.1	Electrons	10
4.2.2	Muons	10
4.3	Jets reconstruction	10
4.3.1	B-tagging	10
4.4	Missing transverse energy	10
5	Data, signals and backgrounds	11
5.1	Data samples	11
5.2	Signal samples	11
5.3	Background prediction	11

5.3.1	The main background: $t\bar{t}$	11
5.3.2	Drell-Yan estimation	11
5.3.3	Non prompt contamination	11
5.3.4	Smaller backgrounds	11
5.3.5	Weights and corrections applied	11
6	Event selection	12
6.1	Signal regions	12
6.2	Control regions	12
6.3	Background-signal discrimination	12
6.3.1	Discriminating variables	12
6.3.2	Neural network	12
7	Results and interpretations	13
7.1	Systematics and uncertainties	13
7.2	Results	13
8	Conclusions	14
8.1	Future prospects	14
	Appendices	15
	Bibliography	16

Chapter 1

Introduction

The Standard Model (SM) of particle physics is nowadays the most accepted mathematical model used to describe the elementary particles and three of the four fundamental forces of nature (electromagnetic, weak and strong interactions). This model is quite simple in concept, but has been able to describe most of the phenomena observed in nature so far with an incredible level of precision, and made a lot of predictions that have now been proven to be true, such as the postulate of the Higgs mechanism [1, 2] followed by the discovery of the Higgs boson itself in 2012 [3, 4] by the Compact Muon Solenoid (CMS) and A Toroidal LHC ApparatuS (ATLAS) experiments of the European Council for Nuclear Research (CERN).

However, as accurate as it seems to be, this theory is known to have several shortcomings which require further investigation. Eventual exotic particles which do not fit in the current model could be the sign of new physics and have therefore been extensively searched for over the course of the last decades in order to enhance our understanding of the Universe and all its constituents.

In this context, the first serious Dark Matter (DM) hypothesis was introduced in the 1970s because of gravitational anomalies observed by several astrophysicists, as a way to explain the apparent nonluminous missing mass in the Universe [5]. Indeed, the visible mass in most galaxies appears to be way too low to explain several astrophysical processes, such as the rotation curves of the galaxies [6], which seems to be incompatible with the well established laws of gravitation. Some additional measurements of the gravitational lensing (in the Bullet Cluster, for example [7]) and the anisotropies observed in the CMB [8] can be quoted amongst other evidences for the existence of DM, as explained in details in Section 2.1.

As far as we currently know from cosmological measurements, ordinary baryonic matter only constitutes around 5% of the Universe, while DM represents around 26% of the energy density of the Universe (the rest is being considered as dark energy) [9]. Understanding the nature and properties of this new kind of exotic matter is therefore crucial to try and understand the physics in the Universe.

Nowadays, the existence of DM is well established in the physics community, even though it has never been observed directly, since our only evidences so far for its existence come from its large-scale gravitational effects. While its nature is still unknown and extensively studied, one of the best DM candidate is the so-called Weakly Interactive Massive Particles (WIMP)s, predicted to interact both gravitationnaly and weakly with SM particles. This would allow direct and indirect direction of such candidates, used as the driving process of many of experiments over the last decades, trying to find the hint of a possible interaction between standard baryonic particles

and eventual DM particles, or even between several DM particles themselves. Dark matter production through the use of a particle accelerator colliding SM particles together, such as the Large Hadron Collider (LHC), is also a possibility, and will be considered as the main channel towards the eventual detection of this exotic matter throughout this work. The production through colliders is actually able to provide constraints on low dark matter masses as well, in a region where both the direct and indirect searches are less efficient, which makes the LHC a perfect tool to study this kind of Beyond the Standard Model (BSM) physics.

However, observing DM is still extremely difficult, mainly because it barely interacts with ordinary baryonic matter, except through gravity (we have to assume that it does interact with SM at least weakly for the sake of this work though, as we would not be able to discover it as an individual particle if it were not the case). This means that nowadays, all the experiments searching for DM have only been able to put constraints on the DM particle mass and on the interaction cross sections between the dark and standard sectors. Actually, even if the collisions between protons produced by the LHC do have an sufficient amount of energy to produce this kind of particles, we would not expect them to interact with our detector, meaning that a direct detection is out of our hands for now. The eventual presence of such matter has to be inferred from the study of the interaction between SM particles and CMS itself, since a typical DM-like event consists of at least one energetic SM particle produced in association with a large imbalance in the transverse momentum due to the presence of an eventual DM candidate that was able to escape our detection.

In the context of this work, DM is searched for in association with one or two top quarks which play the role of the SM particle allowing us to trigger the event. This is indeed a perfect channel for this kind of searches if we assume that the interaction between the dark and standard sectors respect the principle of Minimal Flavour Violation (MFV), which can be consistently defined independently of the structure of the new physics model [10]. In this case, this interaction should follow the same Higgs-like Yukawa coupling structure as the usual SM baryonic particles. This is an important consequence, since it will be shown in Section 2.2 that this coupling is actually stronger with more massive particles: the heavier the SM particle considered is, the easier it is for this particle to couple with the dark sector. This makes the top quark, the most massive of all the fundamental particles observed by far, is therefore an excellent object to study in this context.

However, this also means that its phenomenology is mostly driven by its large mass and that it decays before hadronization can occur, usually into a W boson and a bottom quark ($\sim 96\%$ branching ratio [11]). The final state of the process we are interested in is then made out of some b jets, leptons and/or quarks and is categorized depending on this number of b jets and on the decay of the W itself. This work will actually be focused on the two leptons final state, also known as the dileptonic channel, mostly since this channel does not have lots of background processes raising to a similar final state, even though its branching ratio is the smallest, as will

be explained in Section 2.5. Additionally, leptons are by reconstruction much cleaner than jets. This means that their identification and momentum calculation is easier to perform, and that the uncertainties associated to these measurements will be smaller.

The LHC has now been running for 10 years, and several similar searches have already been carried out and published in the past by the CMS and ATLAS collaborations, at different center of mass energies. First of all, at 8 TeV, several searches for a pair of top quarks were published by the CMS (in association with DM in the semileptonic [12] and dileptonic [13] final states) and ATLAS collaborations [14]. At 13 TeV, the ATLAS collaboration published on one hand several studies, considering different final states [15, 16, 17]. On the other hand, the CMS collaboration published a few extremely important papers for this study [19, 20]. For the first time in 2019, the results obtained by the single top and $t\bar{t}$ analyses have also been combined and published using the 2016 data [21]. Our main objective is now to repeat and improve this analysis that was performed while considering a larger dataset, globally improving the analysis strategy and including the dileptonic final state for the first time in this combination.

After a detailed introduction about DM in general, the experimental device will be detailed in the Chapter 3. This will include a discussion about the LHC itself, along with a complete description of CMS, the detector used to collect the data that will be analyzed throughout this work. This data has been collected during the years 2016, 2017 and 2018 and corresponds to an integrated luminosity of $\sim 138 \text{ fb}^{-1}$, collected during the Run II of operation of the LHC and at a center of mass energy $\sqrt{s} = 13 \text{ TeV}$. In particular, a particular care will be given to the explanation of the Particle Flow (PF) algorithm, used to reconstruct the different objects used and that will be defined in the Chapter 4, while the estimation of the different backgrounds and the selection of interesting events will be detailed throughout the Chapters 5 and 6.

Distinguishing between the signal we are searching for and backgrounds having a much higher cross-section and kinematically really close, such as the SM $t\bar{t}$ without production of DM is not a straightforward task (sometimes a production of missing transverse energy due to the presence of physical neutrinos is even obtained). To isolate the signal and to obtain some discriminations between these kind of processes, an algebraic reconstruction of the event and top-notch Machine Learning (ML) techniques are used in this work, in order to train a network of neurons. The main objective is to make them learn how to combine the discriminating power of a set of input variables in order to create a single output variable describing the probability of a single event to be classified as signal or background. All this process will be detailed in the Section 6.3.

Finally, a statistical interpretation of our data will be performed and different sources of systematic uncertainties will be considered in the Chapter 7. This will allow us to set upper limits on the cross section production value of DM particles in our particular channel and for the simplified models considered in this analysis. The conclusions of this work and some additional future prospects will then finally be presented in the Chapter 8.

Chapter 2

The Dark Matter case

In this chapter the case for Dark Matter (DM) will be presented in Section 2.1, along with a summary of the main evidences, mostly astrophysical, which lead to the introduction of this kind of Beyond the Standard Model (BSM) physics. Then, the main properties expected by such exotic matter in Section 2.2, nowadays most accepted DM candidates, such as the Weakly Interactive Massive Particles (WIMP) briefly introduced previously will be presented in Section 2.3 and the main ways we have to search for this new physics will be shown in Section 2.4.

Finally, the searches performed in colliders such as the Large Hadron Collider (LHC) and our particular channels of interest (DM produced in association with either one or two top quarks) will be detailed in Section 2.5 and the latest similar results and exclusion plots published over the course of the previous years by the ATLAS and CMS collaborations will be displayed in Section 2.6.

2.1 At the origins of dark matter

The origin of the concept of dark matter can be traced back to the 17 and 18th centuries, shortly after Newton's works on gravitation, even though this concept changed a lot over the years. Back then, DM was more considered to be ordinary matter which simply did not emit any kind of electromagnetic radiation and is therefore invisible and dark, but which does have a strong impact in the gravitational point of view because of its mass. It was for example considered in the 20th century to be found in massive astronomical objects able to absorb the light or other objects located behind them, such as black holes.

2.1.1 Zwicky and the virial theorem

Around this time, the first experimental evidences for the existence of dark matter were shown. In 1933, Fritz Zwicky managed to determine the mass of the Coma Cluster using the virial theorem [22], which states that if we consider thousands of stars interacting gravitationally with each other in a galaxy, the average kinetic energy of any given star can be directly related to the average gravitational potential energy of this star. Mathematically, the virial theorem can be written in the Equation 2.1, where \vec{r}_i is the position of the mass m_i and \vec{F}_i is the vectorial

force applied to this mass.

$$\sum_i \frac{1}{2} m_i v_i^2 = -\frac{1}{2} \sum_i \vec{r}_i \cdot \vec{F}_i \quad (2.1)$$

By applying Newton's third law and the universal law of gravitation, the virial theorem can be written for a particular system of gravitationally interacting masses as Equation 2.2, where the universal constants of gravitation $G = 6.67 \cdot 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}$.

$$\sum_i m_i v_i^2 = \sum_i \sum_{i < j} \frac{G m_i m_j}{r_{ij}} \quad (2.2)$$

By assuming a spheric distribution of the masses in the Coma Cluster and by knowing from astronomical observations and calculations the radius of cluster, its total mass and the mean square velocity of stars within this cluster, Zwicky then managed to compute the average mass to light ratio of the galaxies in this cluster and concluded that the value obtained was around 500 times larger than the mass previously estimated by Edwin Hubble, who simply considered the number of visible galaxies in this cluster. One plausible explanation for this discrepancy is to introduce the concept dark matter, which contributes to the mass of the cluster without increasing the galactic luminosity.

Zwicky's results were actually quite controversial since they were based on statistical calculations relying on different hypotheses not always justified, such as the fact that the galaxies in the cluster must be gravitationnaly bound with each other and they were actually proven to be overestimated later on [23].

2.1.2 Spiral galaxies rotation curves

Despite being controversial and slightly off, Zwicky's results were soon followed by a series of additional astronomical observations leading to the same conclusion, the possible existence of nonluminous matter in all the galaxies, dark matter. The most famous of these results is the study of the observed and expected rotation curves of the stars within spiral galaxies in the 1970s [6].

According to this study, if we assume that we can apply Newton's universal laws of gravitation at the galactic scale, then the stars within this kind of galaxies should rotate with a velocity depending on the radius to the galactic center given by the Equation 2.3, where $M(r)$ accounts

for the total mass encountered within a radius r .

$$v_{\text{rotation}}(r) = \sqrt{\frac{GM(r)}{r}} \quad (2.3)$$

At first approximation, one can assume that most of the mass within this kind of galaxies comes from the inner core, meaning that, at large radius, the velocity of individual stars is expected to decrease proportionally to $\frac{1}{\sqrt{r}}$.

However, observations made by Vera Rubin and her team in the early 1970s with a new spectrograph designed the velocity curves of spiral galaxies with a degree of accuracy never reached before, did not confirm these expectations [24]. Indeed, according to these results, from a given value of the radius, the velocity curve appears to be flat instead of decreasing as expected. These results are illustrated in Figure 2.1.

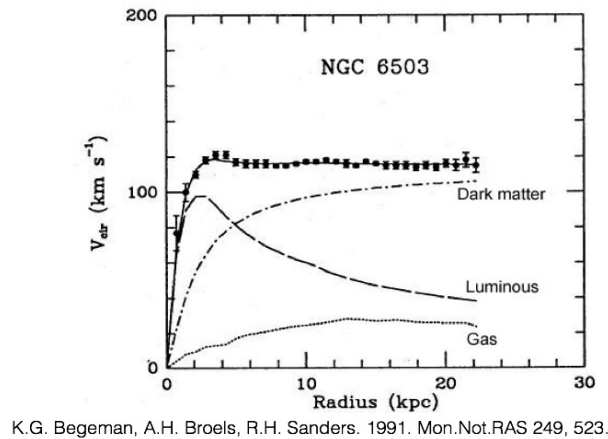


Figure 2.1: Expected and observed rotation curves of the galaxy NGC 6503 [6]. The black dots correspond to the data and the *Luminous* line corresponds to the expected rotation curve.

2.1.3 Cosmic Microwave Background (CMB) anisotropies

The CMB is a mostly uniform background of primary radio waves emitted when the Universe became transparent around 380.000 years after the Big Bang and was discovered accidentally in the 1940s [25]. Studying it is extremely important, as it is actually made out of the oldest and cleanest electromagnetic radiation we can find in the Universe.

Precise measurements of this radiation are actually critical in many different fields of physics, since any proposed model of the Universe must be able to explain this radiation, its temperature and anisotropies. Recent measurements determined that the CMB can be considered as emitting a thermal black body spectrum at a temperature of $(2.72548 \pm 0.00057)\text{K}$ [26]. However, we now know that this temperature is actually not constant as some small anisotropies (at the 10^{-5} level) depending on the value of the angular angles of observation can be observed.

We see these fluctuations projected over a 2D sphere, and it is therefore natural to introduce at this point Laplace's spherical harmonics, $Y_{lm}(\theta, \phi)$, a complete set of orthogonal functions defined on a sphere and defined by a few parameters such as l , the multipole, representing a given angular angle in the sky ($l=100$ corresponds to $\sim 1^\circ$) and m , the number of poles, such as $-l \leq m \leq l$ [27]. The temperature fluctuations, whose value depend on the two usual spherical angles θ and ϕ can then be expanded using these generic functions, according to the Equation 2.4.

$$\frac{\Delta T}{T}(\theta, \phi) = \sum_{l,m} a_{lm} Y_{lm}(\theta, \phi) \quad (2.4)$$

The information about the anisotropies can be extracted from the values of the a_{lm} coefficients, from which we can obtain the values needed to represent the power spectrum of the CMB, according to the Equation 2.5, and from which most of the cosmological information of the CMB can actually be derived.

$$D_l = \frac{l(l-1)C_l}{2\pi} = \sum_m |a_{lm}^2| \quad (2.5)$$

Interestingly enough, this spectrum is directly affected by the value of the density of the dark matter in the Universe. Doing a multi-parameters fit on the observed data in this plot is then able to give us directly the energy density of baryonic Ω_b and dark Ω_χ matter. Today's most precise measurements have been obtained using the Planck satellite, and lead to the determination of these two quantities: $\Omega_b h^2 = 0.02226 \pm 0.00023$ and $\Omega_\chi h^2 = 0.1186 \pm 0.0020$ [28] and the actual power spectrum obtained is represented in the Figure 2.2.

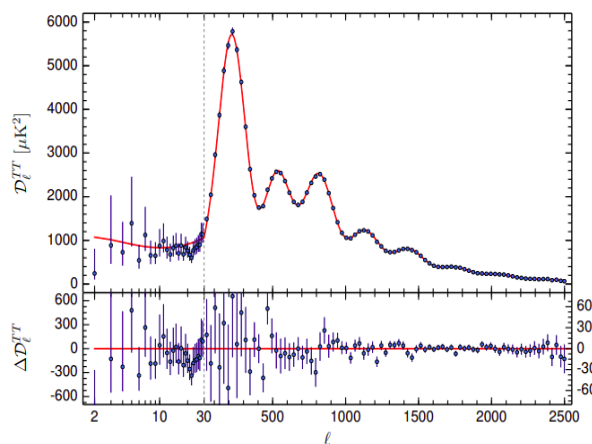


Figure 2.2: Power spectrum of the CMB obtained by Planck, representing the fluctuations of the temperature of this radiation with respect to the angular angle of observation.

2.1.4 Other evidences

Bullet Cluster

Gravitationnal lensing

2.2 Dark matter properties

2.3 Dark matter candidates

2.4 Dark matter searches

2.5 Production at the LHC

2.5.1 The single top production channel

2.5.2 The $t\bar{t}$ production channel

2.6 Previous results

Chapter 3

The experimental device

3.1 The LHC accelerator

3.2 The CMS detector

3.2.1 Tracker

3.2.2 Electromagnetic calorimeter

3.2.3 Hadronic calorimeter

3.2.4 Muon system

3.2.5 Trigger

3.2.6 Data acquisition

Chapter 4

Objects reconstruction

4.1 Particle Flow algorithm

4.2 Leptons reconstruction

4.2.1 Electrons

4.2.2 Muons

4.3 Jets reconstruction

4.3.1 B-tagging

4.4 Missing transverse energy

Chapter 5

Data, signals and backgrounds

5.1 Data samples

5.2 Signal samples

5.3 Background prediction

5.3.1 The main background: $t\bar{t}$

$t\bar{t}$ reconstruction

5.3.2 Drell-Yan estimation

5.3.3 Non prompt contamination

5.3.4 Smaller backgrounds

5.3.5 Weights and corrections applied

Chapter 6

Event selection

6.1 Signal regions

6.2 Control regions

6.3 Background-signal discrimination

6.3.1 Discriminating variables

6.3.2 Neural network

Chapter 7

Results and interpretations

7.1 Systematics and uncertainties

7.2 Results

Chapter 8

Conclusions

8.1 Future prospects

Appendices

Bibliography

- [1] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons", Phys. Rev. Lett. 13, pp. 321-323, 1964
- [2] P. W. Higgs, "Broken symmetries and the masses of gauge bosons", Phys. Rev. Lett. 13, pp. 508-509, 1964
- [3] S. Chatrchyan et al., "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", Phys. Lett. B716, pp. 30-61, 2012
- [4] G. Aad et al., "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", Phys. Lett. B716, pp. 1-29, 2012
- [5] V.C. Rubin, W.K. Ford and N. Thonnard, "Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R=4kpc) to UGC 2885 (R=122kpc)", Astrophysical Journal 238, pp. 471-487, 1980
- [6] K.G. Begeman, A.H. Broeils and R.H. Sanders, "Extended rotation curves of spiral galaxies - Dark haloes and modified dynamics", Monthly Notices of the Royal Astronomical Society, vol. 249, issue 3, ISSN 0035-8711, 1991
- [7] A. Robertson, R. Massey and V. EkCMBTemperaturee, "What does the Bullet Cluster tell us about self-interacting dark matter?", Monthly Notices of the Royal Astronomical Society, vol. 465, issue 1, 2017
- [8] J.B. Muñoz, C. Dvorkin and A. Loeb, "21-cm Fluctuations from Charged Dark Matter", Phys. Rev. Lett. 121, 121301 (2018)
- [9] A. Natarajan, "A closer look at CMB constraints on WIMP dark matter", Phys. Rev. D85, 2012
- [10] G. D'Ambrosio G.F. Giudice, G. Isidori and A. Strumia, "Minimal Flavour Violation: an effective field theory approach", Nucl.Phys. 645, pp 155-187, 2002
- [11] M. Tanabashi et al., Particle Data Group, Phys. Rev. D98, 030001 (2018)
- [12] CMS Collaboration, "Search for the production of dark matter in association with top-quark pairs in the single-lepton final state in proton-proton collisions at $\sqrt{s} = 8$ TeV", JHEP, vol. 6 121, 2015
- [13] CMS Collaboration,, "Search for the Production of Dark Matter in Association with Top Quark Pairs in the Di-lepton Final State in pp collisions at $\sqrt{s} = 8$ TeV", CMS-PAS-B2G-13-004, 2014
- [14] "Search for dark matter in events with heavy quarks and missing transverse momentum in pp collisions with the ATLAS detector", Eur. Phys. J. C (2015) 75:92

- [15] ATLAS Collaboration, "Search for the Supersymmetric Partner of the Top Quark in the Jets+Emiss Final State at $\sqrt{s} = 13$ TeV", ATLAS-CONF-2016-077
- [16] ATLAS Collaboration, "Search for top squarks in final states with one isolated lepton, jets, and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector", ATLAS-CONF-2016-050, 2016
- [17] ATLAS Collaboration, "Search for direct top squark pair production and dark matter production in final states with two leptons in $\sqrt{s} = 13$ TeV pp collisions using 13.3 fb^{-1} of ATLAS data", ATLAS-CONF-2016-076, 2016
- [18] ATLAS Collaboration, "Search for dark matter produced in association with bottom or top quarks in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector", Eur. Phys. J. C 78 (2018) 18
- [19] CMS Collaboration, "Search for dark matter produced in association with heavy-flavor quark pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV", Eur. Phys. J. C (2017) 77: 845
- [20] CMS Collaboration, "Search for dark matter particles produced in association with a top quark pair at $\sqrt{s} = 13$ TeV", Phys. Rev. Lett. 122, 011803 (2019)
- [21] CMS Collaboration, "Search for dark matter produced in association with a single top quark or a top quark pair in proton-proton collisions at $\sqrt{s} = 13$ TeV", JHEP, vol. 03 141, 2019
- [22] F. Zwicky, "Die Rotverschiebung von extragalaktischen Nebeln", Helvetica Physica Acta , vol. 6, pp. 110-127, 1933
- [23] S. Van den Bergh, "The early history of dark matter", Dominion Astrophysical Observatory, 1999
- [24] V.C. Rubin, W.K. Ford, "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions", The Astrophysical Journal 159, p. 379, 1970
- [25] A. A. Penzias, R.W. Wilson, "A Measurement of Excess Antenna Temperature at 4080 Mc/s", Astrophysical Journal 142, pp. 419-421
- [26] D.J. Fixsen, "The temperature of the cosmic microwave background", Astrophysical Journal, 2009
- [27] R. Tojeiro, "Understanding the Cosmic Microwave Background Temperature Power Spectrum", 2006
- [28] Planck Collaboration, "Planck 2015 results. XIII. Cosmological parameters", A& A 594, A13 (2016)