UNIVERSIDAD NACIONAL DE CÓRDOBA

Facultad de Matemática, Astronomía y Física

PHD THESIS ON ASTRONOMY

MACHINE LEARNING METHODS APPLIED TO COSMOLOGICAL PROBLEMS.

AUTHOR: MARTÍN EMILIO DE LOS RIOS

ADVISOR: DR. MARIANO JAVIER DOMÍNGUEZ

0.1 Keywords.

- Dark-Matter.
- Galaxy Clusters.
- Cosmology.
- Machine Learning.

0.2 Classifications

- **02.70.Rr** General statistical methods.
- **98.65.Cw** Galaxy Clusters.
- **98.80.Es** Observational cosmology.
- **95.35.+d** Dark matter (stellar, interstellar, galactic, and cosmological).
- **95.36.+x** Dark energy.

Los vicios de la humanidad demuestran lo poco que soy.

Abstract.

In this work we present studies with novel machine learning techniques about different cosmological problems. In the first part of the work we present the necessary theoretical framework to understand the astronomical problems that we will address. In Chapter 1 we present an introduction of the standard cosmological model, while in Chapter 2 we present the principal details of the Machine learning techniques that we will use. This kind of algorithms represent a new way of analysing big datasets and to find patterns and relations between the variables that are involved in an specific problem. While these techniques have been applied with a lot of success in technological problems and in other areas of science, we apply it for 2 astronomical issues: the dynamical classification of galaxy clusters (Chapter 3) and the statistical study of the temperature anisotropies of the cosmological cosmic microwave background (Chapter 7). We also present an statistical analysis of the relation between the dynamical status of galaxy clusters and the magnetic fields in their intra-cluster medium (Chapter 4) and a series of individual studies of systems that were previously classify as merging clusters (Chapters 5 and 6). In the last Chapter (8) we introduce a theoretical study of the observational and cosmological implications that may have an interaction between the dark matter particle and the cosmic microwave background photons.

Contents

III

0.1 Keywords. III

0.2 Classifications III

Abstract. **VI**

1 Standard cosmological model. **3**

1.1 Brief introduction about the geometry and evolution of the Universe. . . . 3

1.2 The dark sector. 10

1.2.1 The dark matter. 10

1.2.2 Dark Energy. 23

Bibliography **25**

2 A brief introduction to Machine Learning Techniques. **49**

2.1 Supervised Learning. 51

2.1.1 Most important supervised learning models. 51

2.1.2 Model evaluation and Over-fitting. 62

2.2 Non-Supervised Learning. 65

2.2.1 Most important non-supervised learning models. 66

Bibliography **71**

3 Construction of a catalogue of merging clusters with Machine Learning techniques. **73**

3.1 Merging clusters and their importance in cosmology. 73

3.1.1 A challenge for the standard cosmological model. 75
3.1.2 Constraint in the cross-section of the dark matter particle. 78
3.2 The MeSsI (*Merging Systems Identification*) algorithm. 79
3.2.1 Preparation of the training set using cosmological simulations. . . 80
3.2.2 Estimation of features for the merging clusters sample. 83
3.2.3 Training different machine learning models. 85
3.2.4 Application of the MeSsI algorithm to real catalogues. 94
3.2.5 The case of multiple mergers. 99
3.2.6 Conclusions. 99

Bibliography **101**

4 Statistical studies of galaxy clusters according to their dynamical status. **113**
4.1 Introduction. 113
4.2 Observational data. 114
4.2.1 Galaxy clusters survey. 114
4.2.2 Observations of the Faraday Rotation effect. 116
4.3 Statistical analysis. 118
4.3.1 Density distributions. 122
4.4 Inferred magnetic fields. 123
4.5 Conclusions. 126

Bibliography **128**

5 Individual study of the candidates for merging clusters $A2029/33$ and $A1204$. **135**
5.1 Data Acquisition. 136
5.1.1 Photometric observations. 136
5.1.2 X-ray emission observations. 137
5.1.3 Spectroscopic Data. 137
5.2 Weak lensing analysis. 137
5.2.1 Identification and classification of sources. 138
5.2.2 Shape measurements. 140
5.2.3 Estimation of the individual masses of the substructures. 141
5.2.4 Two-dimensional density distribution. 143

5.2.5 Results. 144
5.3 Dynamic analysis. 144
5.3.1 Two-body model. 149
5.4 Conclusions. 151
5.4.1 $A1204$. 153
5.4.2 $A2029/A2033$ 154

Bibliography **155**

6 Individual study of merging cluster candidate $A267$. 163
6.1 Data Acquisition. 164
6.1.1 Spectroscopic and photometric catalogues. 164
6.1.2 X-ray observations and data reduction. 164
6.1.3 Simulated data. 165
6.2 Dynamic and morphological analysis. 165
6.3 Fossil groups in cosmological simulations. 169
6.4 Conclusions. 172

Bibliography **173**

7 Studies on the angular distribution of the cosmological parameters. 181
7.1 On the temperature power spectrum of the cosmic microwave background. 181
7.2 Analysis of the main components of the temperature power spectra. . . . 185
7.2.1 Unsupervised learning techniques applied to the principal components. 189
7.3 *CosmoML*: Estimation of the cosmological parameters through supervised machine learning methods. 190
7.3.1 Cosmological parameters estimation using CMB + Supernovas. . 198
7.4 Angular distribution of the cosmological parameters. 203
7.4.1 Reconstruction of the power spectra. 204
7.4.2 Estimation of the cosmological parameters. 207
7.4.3 Estimation of the angular distribution of the cosmological parameters in PLANCK data. 208
7.5 Conclusions. 214

Bibliography **219**

8 Study on the interaction between dark matter and the CMB photons. **223**

8.1 Redistribution function for the generalised Sunyaev-Zeldovich effect. . . 224

8.2 Study on the interaction between dark matter and photons. 230

8.2.1 Dipolar dark matter model. 230

8.2.2 Velocity distribution of dark matter particles. 235

8.3 Conclusions. 239

Bibliography **242**

List of Figures

1.1 Energy density vs scale factor for the different fluids. a_{eq} correspond to the time where the densities of the relativistic and non-relativistic matter are equal. Figure extracted from *Modern Cosmology, Scott Dodelson*[49]. 7

1.2 Scale factor evolution in the different cosmological phases. Figure extracted from *Modern Astronomy, Bradley W. Carrol & Dale A. Ostlie.* [37] 8

1.3 Rotation curve of the disk galaxy NGC 6503. Figure extracted from *Begeman et al. 1991* [24] 11

1.4 Power spectra of the cosmic microwave background for different values of $\Omega_{dm,0}$. 13

1.5 Interaction Cross-section between dark matter and baryons as a function of dark matter particle mass for the most important candidates. In red are shown hot dark matter candidates, in pink warm dark matter candidates and in blue cold dark matter candidates. Figure extracted from *Roszkowski et al. 2017* [86] 16

1.6 The 4 basics methodologies used in the dark matter search. *Bauer et al. 2013* [23] 18

1.7 Constraint in the cross-section vs mass space for scalar interactions. The limits found by different experiments are: LUX in solid black lines [10], in grey are plot PandaX-II results [74], in brown lines are shown XENON100 and XENON1T [18, 19] and in violet lines CDMSlite-II [101]. The predicted limits for future projects are: CRESST-III [100] in light-blue, DarkSide G2 [2] in dotted violet lines, DEAP3600 [46] in dotted blue lines, LZ [102] in black dotted line, SuperCDMS and SNOLAB [7] in pink dotted line and, in dotted brown line, XENON1T/nT [17]. We also plot in green the area favoured by the minimal super-symmetric models (MSSM) and in brown the area favoured by the constraint minimal supersymmetric model (CMSSM). In blue shadow areas are shown the events observed by CDMS (blue), CoGenT (grey), CRESST-II (light-blue) y DAMA-LIBRA (light green). Finally, the shadow orange area at the bottom of the plot correspond to the irreducible neutrino background [28]. Figure extracted from *Roszkowski et al. 2017* [86]. 20

1.8 Full-sky map of the γ-ray emission observed with the Fermi spatial telescope FERMI. Figure extracted from http://www.nasa.gov 21

1.9 Current constrains in the mass vs cross-section space for the annihilation of dark matter particles. In gold line is shown the results obtained with the analysis of dwarf galaxies with Fermi-LAT [11] data. In black line is plotted the analysis combining Fermi-LAT and MAGIC [69] data. The results obtained with HESS observations [3] are shown in violet lines, while the constraints obtained with Fermi-LAT observations of the Milky-way centre [57] are shown in grey lines. The results obtained by Planck collaboration through the analysis of CMB data [79] are plotted in dotted blue and green lines. The constraints obtained with the joined analysis of 45 dwarf galaxies [40] are shown in dotted brown lines, while the limits obtained by CTA collaboration [4] assuming an NFW profile are shown in dotted pink lines, while the constraints obtained assuming an Einasto profile are shown in dotted red line. In grey dotted line is shown the cross section of the annihilation that correspond to the thermal production of WIMPs particles [99]. Figure extracted from [87] 22

2.1 Percentage of objects well classified as a function of the number of examples present in the training set 50
2.2 *k-nearest neighbour* algorithm. (Left). KNN trained with $k = 1$. (Right.) KNN trained with $k = 15$. Figure extracted from *Hastie et al. 2001* [2]. . 53
2.3 Classification schema using *Support Vector Machine*. 55
2.4 Classification scheme using a *Random Forest*. Figures extracted from *Murphy* [4]. 56
2.5 Classification scheme of an artificial neural network. 59
2.6 Comparison between the classification schemes of traditional machine learning and *Deep Learning* algorithms. 62
2.7 Traditional method to evaluate the performance of a supervised learning algorithm. 63
2.8 Confusion matrix used to evaluate the performance of supervised machine learning algorithms. 64
2.9 Example of ROC curve used to compare different supervised learning methods. 66
2.10 Schema of a two-dimensional principal component analysis. 69
2.11 Example of a facial reconstruction algorithm by varying the number of the principal components used for its reconstruction. Figure extracted from http://www.declanoller.com/tag/pca/ 70

3.1 *Left* X-ray image superimposed to mass contours estimated with weak gravitational lenses studies. *Right* Optical image superimposed to mass contours estimated with weak gravitational lenses studies. Figures extracted from *Clowe et al.* [17] 74
3.2 Schematic geometry of a merging cluster. Figure extracted from *Harvey et al. 2015* [43]. 79
3.3 Scheme of a halo that had a major merger in the snapshot 61 and whose haloes survived as subhaloes until the snapshot 63. 82
3.4 Distribution of merger times of the merging clusters sample. In black line is shown the result found by *Pinkney et al. 1996*[69]. 83
3.5 From left to right and from top to bottom, is shown the result of the application of the iterative Dressler-Shectman test. 86

3.6 Distribution of features for interacting (red) and non-interacting (black) systems. . . . 87
3.7 Distribution of features for interacting (red) and non-interacting (black) systems. . . . 88
3.8 Comparative ROC curve for the different machine learning methods used for the classification of the clusters. . . . 89
3.9 True positives rate (red crosses), false positive rate (blue asterisks), effectiveness (black dots) and normalized length of the final sample (green diamonds) as a function of ν_{crit}. . . . 90
3.10 True positives rate (red crosses) and effectiveness (black points) as a function of the final size of the sample. In black line is shown the number of clusters and the expected rates for a $\nu_{crit} = 0.3$. . . . 91
3.11 ROC curve of *Random Forest* for the galaxy classification. . . . 91
3.12 *Left*: Comparison between the estimated masses with the real ones for the interacting substructures. The mass ratio between the main and the secondary substructure can be seen in colour scale. *Right*: Comparison between the distribution of estimated and the real mass ratios of the merging substructures. . . . 93
3.13 *Left*: Real and observed separation between the centres of the interacting substructures, normalized to the virial radius real. *Right*: Absolute value of the radial velocity separation between the merging substructures, normalized to the real velocity dispersion. . . . 94
3.14 Geometry of the classified merging systems. . . . 95

4.1 Comparison of the main properties of the relaxed and unrelaxed clusters. *Top Left*: Distribution of the cluster radius (R_{200}). *Top Right*: Distribution of the number of member galaxies. *Centre Left*: Distribution of apparent magnitude r_{mag}. *Centre Right*: Distribution of the cluster richness. *Bottom Left*: Distribution of redshifts. *Bottom Right*: Histogram of absolute magnitude R_{mag}. . . . 117
4.2 angular sigma (RMs) . . . 119
4.3 projected sigma (RMs) . . . 120
4.4 Angular Bright Faint . . . 120

4.5 projeted Sigma normalized R200 121

4.6 Angular sigma (RMs) by Han 122

4.7 all density distribution 124

4.8 Density distribution M40 125

5.1 $A1204$: Radial shear profile as a function of the projected distance to the centre of the substructure. In the upper panel graph we show the main component, while the secondary component of $A1204$ is shown in the bottom panel. The solid and dotted lines represent the fitted SIS and NFW profiles respectively. The points and crossings show the tangential and cross ellipticity components of the selected background galaxies, averaged in annular bins, respectively. 145

5.2 $A2029/2033$: Radial shear profile as a function of the projected distance to the center of the substructure. In the upper panel graph we show the main component ($A2029$), while the secondary component ($A2033$) is shown in the bottom panel. The solid and dotted lines represent the fitted SIS and NFW profiles respectively. The points and crossings show the tangential and cross ellipticity components of the selected background galaxies, averaged in annular bins, respectively. 146

5.3 Projected density distribution in the field A1204 obtained with the weak lensing analysis. The scale, marked at bottom left, is given in $h_{70} M_{\odot}/pc^2$. Red contours corresponds to a projected density above 3σ significance level (250 $h_{70} M_{\odot}/pc^2$). X-ray contours are plotted in blue, the contour levels are (3, 5, 7, 9 and 12) times the rms noise. Red dots and green squares are the BCGs positions and the dynamical centres, respectively. Pink and light-blue points are the positions of the galaxies classified as members, for the primary and secondary component of A1204, respectively. 147

5.4 Projected density distribution in the field A2029 and A2033 obtained with the weak lensing analysis. The scale, marked at bottom left, is given in $h_{70} M_\odot / pc^2$. Red contours corresponds to a projected density above 3σ significance level (400 $h_{70} M_\odot / pc^2$). X-ray contours are plotted in blue, the contour levels are (3, 5, 7, 9 and 12) times the rms noise. Red dots and green squares are the BCGs positions and the dynamical centres, respectively. Pink and light-blue points are the positions of the galaxies classified as members, for A2029 and A2033, respectively. 148

5.5 Projection angle (α) as a function of the relative radial velocity difference (V_r) computed according to the equations of motion for A1204 (*left panel*) and A2029/2033 (*right panel*). Solid black curves separate the bound and unbound solutions according to the Newtonian criterion. Blue curves are the solutions of the equations of motions for bound (solid line) and unbound systems (dashed line) and the solutions for each system are marked with open circles, according to the observed V_r marked with the red line, where the gray region corresponds to the uncertainty in these values delimited by the dashed black line. Point curves express the uncertainties in the computed curves considering the errors in the adopted lensing masses. 152

6.1 Distribution of the member galaxies of the substructures identified by `MeSsI`. The color scale shows the redshift of each galaxy. 166

6.2 Results of the Dressler-Shectman test for the $A267$ core. We plot the positions of the galaxies within a box of $R_{200} \times R_{200}$, with the colourbar and size coding their velocities and their δ parameter, respectively. 167

6.3 Phase-space diagram for A267. Red and blue circles represent the two structures found with MCLUST. Dashed lines correspond to constant values of $(|\Delta(v)|/\sigma_1) \times (r/r_{200})$ following *Noble et al. 2013* [20] 168

6.4 *Left*: CMD diagram for A267, where the red, blue and black dots represent galaxies belonging to the main substructure, the infalling substructure and the outskirts, respectively. Black squares enclose galaxies with distances to the cluster's centre $< 0.5 R_{200}$. *Right*: Colour histogram of the structures. 169

6.5 Projected density map made by *Okabe et al.* [21] superimposed to the X-ray emission contours. In red circles and green squares are plotted the galaxies belonging to the substructures identified by *MeSsI* in this region. 170
6.6 Results obtained from simulated data. *Upper panel:* Logarithmic of the median mass of the FGs at $z = 0$ related to the redshift of the last major merger event, z_{merger}. *Lower panel:* Median factor of the mass increment after the last merger event as a function of z_{merger}. In blue are shown FG_{12} and in red FG_{14}. 171

7.1 Examples of power spectra computed with `CAMB` for different sets of cosmological parameters. 186
7.2 Example of a full-sky spectrum (black line) superimposed to the estimated bins (red dots). 187
7.3 Power spectrum of the different cosmological models that will be used in the analysis with machine learning methods. 188
7.4 Correlation between the 8 first principal components and the cosmological parameters. 189
7.5 Cosmological models belonging to the same groups found by `mclust` in the $\Omega_m h^2$-Ω_k space. 190
7.6 Comparison between cosmological parameters predicted with supervised learning methods and real ones for all cosmological models of the test set. 193
7.7 Comparison between cosmological parameters predicted with supervised learning methods and real ones for all cosmological models of the test set. 194
7.8 Comparison between the cosmological parameters predicted with the extended feature space and the real ones for all the cosmological models of the test set. 195
7.9 Comparison between the cosmological parameters predicted with the extended feature space and the real ones for all the cosmological models of the test set. 196
7.10 Relative importance of the different features when predicting the cosmological parameters. 197
7.11 Relative importance of the different features when predicting the cosmological parameters.. 198

7.12 Binned Hubble diagram (Modulus distance in function of the redshift). In gray lines are shown the supernovas of the simulated samples and in the black line is highlighted the mean of DM for each bin of z. While in black points are plotted the average of the DM in bins of z for the JLA supernova sample. 200

7.13 Comparison between the cosmological parameters predicted with CMB+Hubble-diagram features and the real ones for all cosmological models of the test set. 201

7.14 Comparison between the cosmological parameters predicted with CMB+Hubble-diagram features and the real ones for all cosmological models of the test set. 202

7.15 Relative importance of the features when estimating the Hubble constant. 203

7.16 Analysed sky areas corresponding to the pixelization scheme of *Healpix* with $NSIDE = 4$. 205

7.17 Comparison between the power spectrum of 2 randomly selected areas of the sky and the correspondent *full-sky* power spectrum. 205

7.18 Architecture Scheme used in the neural network. 206

7.19 Example of the reconstruction of an spectrum of a zone of the sky corresponding to a pixel of the pixelization scheme of *Healpix* with $NSIDE = 4$.207

7.20 Histogram of Δ for all the cosmological models of the test set 208

7.21 Comparison between the cosmological parameters predicted from the reconstructed spectra and the real ones for all the cosmological models of the test set. 209

7.22 Comparison between the cosmological parameters predicted from the reconstructed spectra and the real ones for all the cosmological models of the test set. 210

7.23 Angular map of the contamination mask. 211

7.24 SILC angular map, provided by *Rogers et al.* [12, 13]. 211

7.25 Histograms of the estimated cosmological parameters for the 192 areas of the sky. In black lines are shown the distributions of each simulated map, while in red lines are shown the distributions corresponding to the real data. Also, in vertical line is plotted the average of the cosmological parameter estimated with the full-sky map of each simulation, where the error band is given by the standard deviation of the estimated values. In red vertical line is shown the value estimated by the methods described in section 7.3 for the full-sky map of Planck, where the error band corresponds to the χ statistic. For a better comparison we added, in vertical green line, the value estimated by the Planck collaboration with its respective error [11].213

7.26 Value of the estimated cosmological parameter as a function of the percentage of the sky that is unaffected by the Planck contamination mask in each zone. 215

7.27 Histograms of the estimated cosmological parameters for the 192 areas of the sky with the contamination mask. In black lines are shown the distributions of each simulated map, while in red lines are shown the distributions corresponding to the real data. Finally, the green line shows the distribution found in Planck data adding the value of the χ statistical corresponding to the error detailed in the table 7.4. 216

8.1 Coordinate reference system for the elastic two-body interaction. 225

8.2 Energy redistribution function for the interaction between CMB photons and electrons in the ICM that follow a Maxwell-Boltzmann energy distribution with different temperatures. In dotted lines are plotted the results found using the Boltzmann method, while in lines are shown the results obtained through the traditional radiative transfer method. 228

8.3 Energy redistribution function for the interaction between CMB photons and electrons coming from the annihilation between dark matter particles that follow a non-thermal energy distribution with different temperatures. In dotted lines are plotted the results found using the Boltzmann method, while the results obtained with the traditional radiative transfer method are plotted in solid lines. 229

8.4 Thomson cross-section σ_{th} estimated through the collisional method and normalised to the theoretical value. 230
8.5 Feymann diagrams representative of the interactions between dark matter and photons. 231
8.6 Cross-section for the interaction between photons and dipolar dark matter particles of different masses, estimated through the collisional method. *Left.* Dipolar dark matter with $M = D = 2 * 10^{-16} KeV^{-1}$ and $m_{DM} = 10^6 KeV$. *Right* Dipolar dark matter with $M = D = 2 * 10^{-16} KeV^{-1}$ and $m_{DM} = 10^9 KeV$. 235
8.7 Energy redistribution function for the interaction between photons and dipolar dark matter particles with $M = D = 10^{-9} KeV^{-1}$. *Left* Particles with $m_{dm} = 1e3 KeV$ and with a thermal energy distribution with different temperatures. *Right* Particles with a thermal energy distribution with $K_b T = 100 KeV$ and different masses. 236
8.8 Energy redistribution function for the interaction between photons and dipolar dark matter particles with $m_{DM} = 10^6 KeV$, a thermal energy distribution with $k_B T = 5 * 10^4 KeV$ and different values of M and D. For a better comparison the value of δI_γ was multiply by 10^{16} and 10^{28} for the dark matter with $M = D = 2 * 10^{-13} KeV^{-1}$ y $M = D = 2 * 10^{-16} KeV^{-1}$ respectively. 237
8.9 Velocity distribution of a dark matter population in equilibrium with an NFW halo ($\rho_0 = 10^{14} M_\odot / Mpc^3$, $r_0 = 1 Mpc$) at a distance of $r = 0.3 Mpc$ from the centre. In solid lines are shown the results obtained through the Eddington method, while in dashed lines are shown Maxwell-Boltzmann distributions. 239
8.10 Energy redistribution function produced by the interaction of CMB photons with a population of dipolar dark matter particles that follows a Maxwell-Boltzmann distribution with a velocity dispersion dependent on the radio (See equation 2.6 and 2.7 of *Petač et al. 2018* [6]). For a better comparison the δI_γ value of the dark matter with $m_{dm} = 10^6 KeV$ and $m_{dm} = 10^9 KeV$ was multiply by $5 * 10^{11}$ and 10^4 respectively. 240

List of Tables

3.1 Comparison of initial velocities for different simulations. The velocity V_{2500} is estimated assuming that the cluster moves as a point mass in a free fall path from its initial position to a separation of $2.500kPc$. Table extracted from *Lage & Farrar* [54]. 77

3.2 This table shows the candidates for interaction groups identified in the SDSS catalogue DR7 [1]. In the first column the name of the group under study is presented. From the 2^{nd} to the 5^{th} column we present the estimated mass and positions of the main substructure of each cluster. From the sixth to the ninth column we present the estimated mass and position of the second substructure. Finally in the last column we present the previous works in which each object was studied. Those clusters that have a previous classification as merging clusters are identified with ✓. References: 1 [90],2 [93], 3 [29], 4 [19], 5 [83], 6 [2], 7 [76], 8 [75], 9 [68], 10 [16], 11 [26], 12 [65], 13 [24], 14 [32], 15 [8], 16 [73], 17 [10], 18 [50], 19 [3], 20 [46], 21 [27], 22 [13], 23 [70] 96

3.3 This table shows the candidates for interaction groups identified in the WINGS catalogue [15]. In the first column the name of the group under study is presented. From the 2^{nd} to the 5^{th} column we present the estimated mass and positions of the main substructure of each cluster. From the sixth to the ninth column we present the estimated mass and position of the second substructure. Finally in the last column we present the previous works in which each object was studied. Those clusters that have a previous classification as merging clusters are identified with ✓. References: 1 [19], 2 [68], 4 [32], 3 [73], 5 [91], 6 [49], 7 [85], 8 [52], 9 [36], 10 [79], 11 [78], 12 [55], 13 [31] 97

3.4 This table shows the candidates for interaction groups identified in the HeCS catalogue [77]. In the first column the name of the group under study is presented. From the 2^{nd} to the 5^{th} column we present the estimated mass and positions of the main substructure of each cluster. From the sixth to the ninth column we present the estimated mass and position of the second substructure. Finally in the last column we present the previous works in which each object was studied. Those clusters that have a previous classification as merging clusters are identified with ✓. References: 1 [93], 2 [29], 3 [32], 4 [9], 5 [71],6 [28], 7 [94], 8 [84], 9 [53], 10 [4], 11 [80], 12 [58] 98

4.1 Correlation between galaxy and AGN surveys in different photometric bands. 116

5.1 Specifications of the observational data used in the study of each cluster. In column (1) we detailed the ID of each system. In columns (2), (3) and (4), the coordinates and redshift of the centres of the systems are detailed. Finally, in columns (5), (6), (7) and (8) we present the specifications of the observational data used. († Average *seeing* in the mosaic.). 136

5.2 Main results of the weak gravitational lens analysis. The ID of each system is presented in the 1^{st} column. In the columns (2), (3) and (4) we show the angular coordinates and redshift of the centres used for the analysis. In column (5) we show the geometric factor. In the columns (6) and (7) we shown the results of the fitting for the SIS profile, the velocity dispersion and M_{200} respectively. while in the columns (8), (9) and (10) the results for the fitting of an NFW profile are shown: c_{200}, R_{200} and the estimated mass M_{200} respectively. 144

5.3 Solutions for the 2-body model. Columns: (1) Cluster identification. (2) Solution class. Resultant values of α (3), R (4), R_m (5), V (6) and the computed probability P (7) for each solution. 151

6.1 Main characteristics of the substructures identified within R_{200}. 167

7.1 Studied intervals for each cosmological parameter. 187

7.2 Relative importance of the principal components of the spectra. 188

7.3 Results of the performance (χ) of the different machine learning methods when estimating the cosmological parameters. 194

7.4 Results on the performance (χ) of the *Support Vector Machine* method when estimating the cosmological parameters in the different frames of the sky corresponding to the pixelization scheme of `Healpix` with $NSIDE = 4$. 208

Chapter 1

Standard cosmological model.

1.1 Brief introduction about the geometry and evolution of the Universe.

Cosmology is the study of the origin and evolution of the Universe. Taking into account the temporary and spatial scales involved, in order to study the evolution of the Universe is necessary to use the general relativity framework. In this theory the space-time structure is represented with a differential manifold plus a metric tensor that fulfil Einstein field equations:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R - g_{\mu\nu}\Lambda = 8\pi G T_{\mu\nu} \tag{1.1}$$

where $R_{\mu\nu}$ is the Ricci tensor, $g_{\mu\nu}$ is the metric tensor, Λ is the cosmological constant, G is the Newton constant and $T_{\mu\nu}$ is the energy-momentum tensor that depends on the energy and matter distribution of the universe.

The standard cosmological model is based on the assumption that the Universe is isotropic and homogeneous at large scales (cosmological principle). This principle is supported by lot of observations (galaxy distributions at large scale, cosmic microwave background, etc. [78, 92]). Taking into account this assumptions and resolving the Einsteins equations, the Universe can be described by the Friedmann-Robertson-Walker (FRW) metric:

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = dt^2 - a^2(t)[\frac{dr^2}{1-Kr^2} + r^2(d\theta^2 + sen^2(\theta)d\phi^2)] \tag{1.2}$$

where $\{x^\mu\}_{\mu=0,1,2,3}$ are the coordinates of the space-time ((t,r,θ,ϕ) in the second equality), $a(t)$ is the scale factor and K is a constant that specify the curvature ($K = 0$ flat curvature, $K < 0$ open curvature and $K > 0$ closed curvature).

If we assume that the matter of the Universe can be represented, in a first approximation, as a perfect fluid with an state equation $p = \omega\rho$, and taking into account the symmetries introduced by the cosmological principle, the energy-momentum tensor is:

$$T_{\mu\nu} = \textit{diag}(\rho, -p, -p, -p) \tag{1.3}$$

where ρ is the total matter density and p is the pressure. Putting 1.3 and 1.2, in 1.1, we can obtain the Friedmann equations that determine the evolution of the scale factor

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3} \tag{1.4}$$

$$\frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{K}{a^2} = 4\pi G(\rho - p) + \Lambda \tag{1.5}$$

where dot indicate derivation with respect to the cosmic time t. Replacing 1.4 in 1.5, we obtain:

$$H^2(t) \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{K}{a^2} + \frac{\Lambda}{3} \tag{1.6}$$

where $H(t)$ is the Hubble parameter. Deriving 1.6 and combining with 1.4 we can obtain the energy conservation equation

$$\dot{\rho} + 3H(\rho + p) = 0 \tag{1.7}$$

and, taking into account the state equation for a perfect fluid

$$\rho \propto a^{-3(1+\omega)} \tag{1.8}$$

The last cosmological observations [78, 83], suggest that the universe is compound for different fluids, and so, the total energy density can be expressed as

$$\rho = \sum_{i=1}^{N} \frac{\rho_{i,0}}{a^{3(1+\omega_i)}} \tag{1.9}$$

where $\rho_{i,0}$ is the density of each fluid in the present time, ω_i is the state equation of each fluid and where the scale factor is normalized in order to have $a_0 = a(t_0) = 1$, where t_0 is the present time. Taking into account these definitions, the equation 1.6 can be expressed as

$$H^2(t) = \frac{8\pi G}{3} \sum_{i=1}^{N} \frac{\rho_{i,0}}{a^{3(1+\omega_i)}} + \frac{K}{a^2} \tag{1.10}$$

It is worth to note that the cosmological constant can be described as a prefect fluid with an state equation $p_\Lambda = -\rho_\Lambda$.

After that, using equation 1.10, we can estimate K in function of the densities of the different matter and energy fluids.

$$K = H^2(t) \left(\frac{8\pi G}{3H^2} \sum_{i=1}^{N} \frac{\rho_{i,0}}{a^{3(1+\omega_i)}} - 1 \right) \tag{1.11}$$

Defining the critic density ρ_c as the total density that the universe have to had in order to be flat ($K = 0$).

$$\rho_c = \frac{3H^2}{8\pi G} \tag{1.12}$$

Taking into account this definition, we can define the density parameter Ω_i [38] for each fluid *i* as:

$$\Omega_i \equiv \frac{\rho_i}{\rho_c} \tag{1.13}$$

For completeness we define the density parameter for the curvature as:

$$\Omega_K = \frac{K}{H^2} \tag{1.14}$$

and so, the equation 1.11 is reduced to

$$\sum_{i=1}^{N} \Omega_i + \Omega_\Lambda + \Omega_K = 1 \tag{1.15}$$

Finally, taking into account all these definitions, the evolution of the scale factor (Equation 1.6) is:

$$H^2(t) = H_0^2 \left(\sum_{i=1}^{N} \frac{\Omega_{i,0}}{a^{3(1+\omega_i)}} + \frac{\Omega_{K,0}}{a^2} + \Omega_{\Lambda,0} \right) \tag{1.16}$$

where subscript 0 indicates the parameters values at present time t_0.

From this equation we can infer that the evolution of the universe is determined by the proportions and properties of the different fluids that make up the universe.

Is known that the universe is composes basically for 3 different types of matter. Taking into account the state equation of each fluid individually, we can study the evolution of its density and determine the fluid that dominates in each cosmological phase.

- Non-Relativistic Matter: The state equation for non-relativistic matter is $\omega = 0$, and so taking into account the equation 1.8, the density is $\rho \propto a^{-3}$. This result can be understood because the matter density drops out as a^{-3} due to the expansion of the universe.

- Relativistic Matter: For the relativistic matter the state equation is $\omega = 1/3$, and so the density is $\rho \propto a^{-4}$. We can interpret this result if we take into account that the relativistic matter density drops out proportional to a^{-3} due to the expansion of the universe, but it has an extra contribution proportional to a^{-1} because the photons loose energy due to the cosmological redshift effect. The cosmological redshift can be estimated if we take into account the photon propagation through null space-time geodesic of the FRW metric. Comparing the wavelength of the photons in the emission (λ_e) and reception point (λ_r) it can be seen that they have an energy loss ($\lambda_e < \lambda_r$) proportional to the growth of the scale factor:

$$1 + z = \frac{\lambda_r}{\lambda_e} = \frac{a_r}{a_e} \tag{1.17}$$

- Cosmological constant Λ: The cosmological constant has an state equation $\omega = -1$, and so, its density is constant $\rho = cte$.

Studying the temporal evolution of the density of this 3 fluids we can identify 3 cosmological phases. At the beginning ($a \to 0$) the universe was dominated by the relativistic matter, after that, because the relativistic matter density drops out as a^{-4} and the non-relativistic matter density drops out as a^{-3} their density became equal at (t_{eq}) at a characteristic scale $a_{eq} = a(t_{eq})$. After t_{eq} the evolution of the universe became dominated by the non-relativistic matter density. Finally, taking into account that the density of Λ is constant, the evolution of the scale factor at late time is dominated by this parameter. This behaviour can be seen in figures 1.1 y 1.2 .

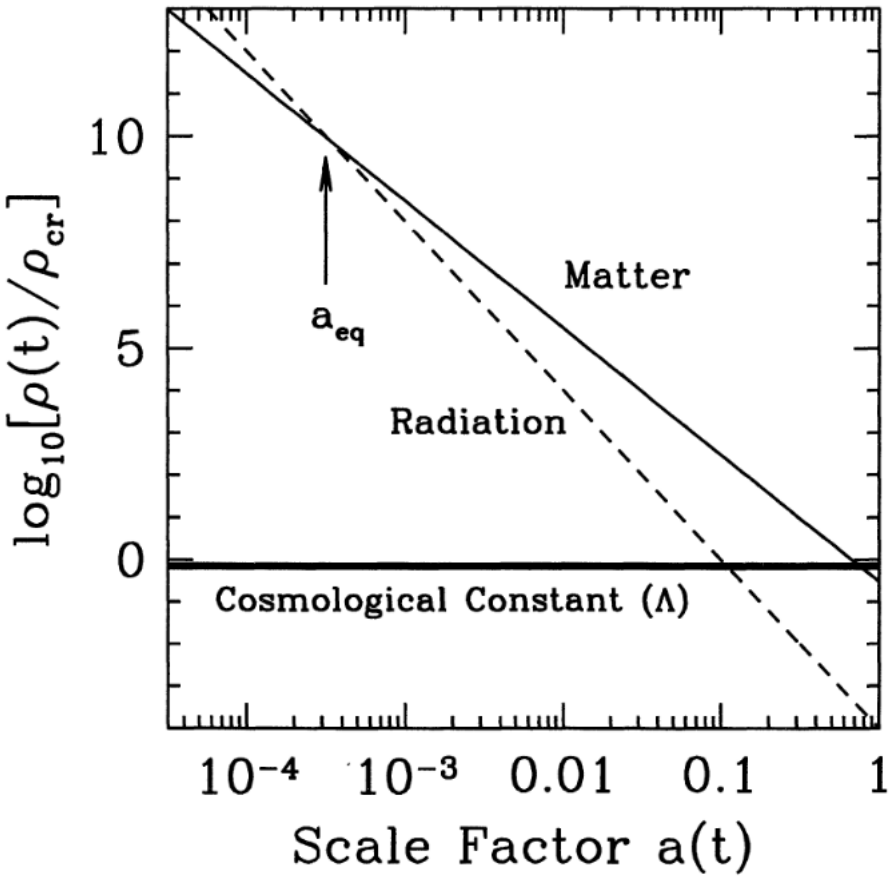

Figure 1.1: Energy density vs scale factor for the different fluids. a_{eq} correspond to the time where the densities of the relativistic and non-relativistic matter are equal. Figure extracted from *Modern Cosmology, Scott Dodelson*[49].

In the last decade, thanks to numerous observations in different electromagnetic bands (Cosmic microwave background [78], supernovas Ia [76, 83], etc.), we achieve very high

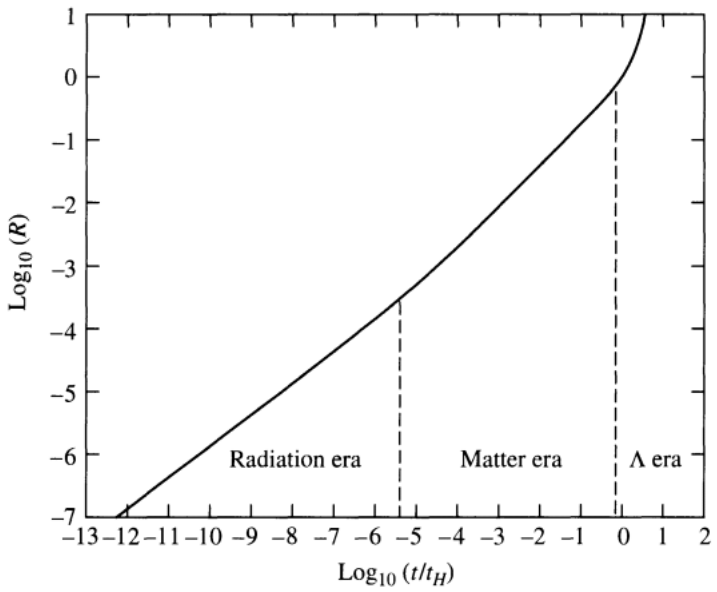

Figure 1.2: Scale factor evolution in the different cosmological phases. Figure extracted from *Modern Astronomy, Bradley W. Carrol & Dale A. Ostlie.* [37]

precision measurements of the different cosmological parameters. The standard cosmological model set down that we live in a flat universe ($K = 0$), with $\Omega_{m,0} = 0.321 \pm 0.013$, $\Omega_{\Lambda,0} = 0.679 \pm 0.013$ and $H_0 = 66.88 \pm 0.92$ [80].

It is worth to note that although this model explain in a very precise way lot of observations, there are still some contradictions, inconsistencies and observations that are difficult to explain. One of the more important and more studied problems is the inconsistency between the Hubble constant estimated trough the cosmic microwave background ($H_0 = 66.88$ [80]) and the one estimated with local Supernovas Ia calibrated trough cepheids stars ($H_0 = 73.24$ [83]).

It is worth to note that this measurements were done with independent data sets that comes from very different physics, and so, the may have different systematic errors. The cosmic microwave measurements is interpreted with the kinetic theory in an expansive universe (Boltzmann-Einstein equation) applied to the early universe (See Chapter 7). Although this is a very well known theory, that was apply successfully in different problems, in order to interpret the results of it application to the early universe is necessary to understand all the contamination sources that affects the photons that comes from the CMB. On

the other hand, the measurement of the Hubble constant made with supernovas Ia information uses very complex models for the calibration of the supernovas explosions, with cepheid stars. While all the physics that intervenes in this explosions makes that the models have a lot of free parameters, the supernovas are in the near universe ($z < 0.4$), and so, are not so affected from foreground contamination as the photons that comes from the CMB.

Different solutions have been proposed to explain this problem [30, 47, 52], but until now there is no conclusive answer.

1.2 The dark sector.

As was specified in section 1.1 according to the standard cosmological model we live in a flat universe ($K = 0$), with $\Omega_{m,0} = 0.321 \pm 0.013$, $\Omega_{\Lambda,0} = 0.679 \pm 0.013$ and $H_0 = 66.88 \pm 0.92$ [80]. Where $\Omega_{m,0} = 0.321$ is the combination of baryonic matter [1] $\Omega_{b,0} = 0.0494 \pm 0.0004$ and dark matter $\Omega_{dm,0} = 0.269 \pm 0.004$.

From the different components of the universe we only understand the physics of the baryonic matter and the radiation (relativistic matter), that only represent the $\approx 5\%$ of the total energy in the present time. The other 95% is composed by 2 unknown forms of matter which physics is not completely understood, the dark matter and the dark energy.

Although we do not know the properties of these fluids, we do know some general characteristics that we can deduce from different cosmological observations. These observations allows us to put constraints in the space-parameters (masses, coupling constants, velocities, etc.) where these particles may live, and in some cases, make stringent predictions that can be contrasted with observations.

In this section we summarise the principal properties of dark matter and dark energy. We also analyse the different candidates that comes from extension to the standard particle model and from modified gravities that have been proposed for dark matter particle.

1.2.1 The dark matter.

Observational evidences.

The dark matter was first postulated by Fritz Zwicky [108] in 1933 as an explanation to the big relative velocities found in the galaxies of the Coma cluster. From then, numerous independent observations at different scales, indicates that there is more matter than the one that we can infer from the light that we observe.

- Galactic scale:

  One of the most important evidences of the dark matter is the flattening of the rotation curves of disk galaxies (figure 1.3).

[1] In astronomy is common to call baryonic matter to all non-relativistic fluids formed by standard model particles and so, for example, the electrons are considered baryonic matter although strictly they are not baryons.

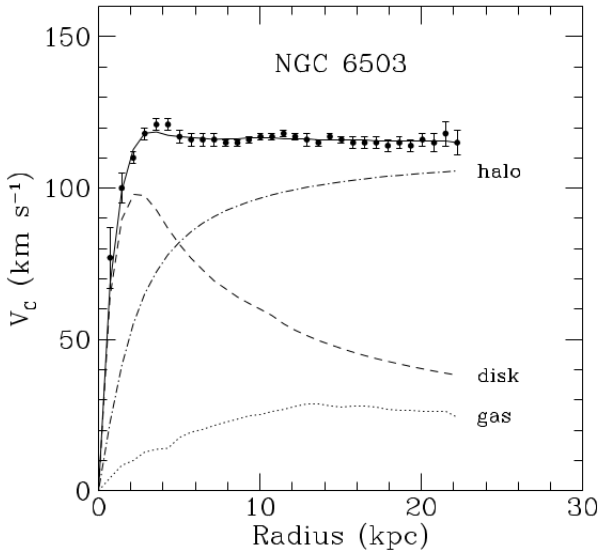

Figure 1.3: Rotation curve of the disk galaxy NGC 6503. Figure extracted from *Begeman et al. 1991* [24]

Assuming that disk galaxies follows the virial theorem, we can established a relation between the mass of a system at a certain distance r from the centre and its rotation velocity:

$$M(r) \propto v^2 r/G \tag{1.18}$$

where G is the Newton constant and v is the rotational velocity, that can be measured, for example, from the HI (neutron hydrogen) emission at $21cm$. If all the mass of a galaxy is associated with their light emission, then, after the radii where we stop seeing light, the enclosed mass would be constant, and so, the rotational velocity would fall down as $v^2 \propto 1/r$. Nevertheless, as can be seen in figure 1.3, the rotation curve became constant ($v = cte$), implying that $M \propto r$ beyond the radii where the galaxy stop emitting light. This suggest the presence of a dark matter halo that extents beyond this radii.

It is worth to note that, through the analysis of gravitational lenses, they where able to found evidence of the presence of dark matter in elliptic galaxies (strong lensing [67] and weak lensing [61].)

- Galaxy Clusters scales:

  The first evidence of the dark matter in galaxy clusters are the studies of Zwicky. He compare the mass of the Coma cluster, estimated using the virial theorem, with the luminosity, and found that the mass-to-light ratio was 10 times bigger than the mass-to-light ratio of galaxies [108, 109]. He conclude that this cluster should have more mass than the one that we observe through light emission. Nowadays the mass of galaxy clusters can be estimated through different methods, as the application of the virial theorem to the radial velocity distribution of galaxies, weak and strong lensing effect, density profiles of X-ray emission from the hot gas in the intra-cluster medium (ICM), etc. Comparing the results of these techniques it can be found that most of the matter of galaxy clusters is in form of dark matter [26].

  Another evidence of the existence of dark matter in galaxy clusters is the offset between the mass distribution (inferred with gravitational lensing analysis) and the X-ray emission from the hot gas in the ICM in merging clusters ([43], [44], [72]). In chapter 3 we present a detailed analysis of the importance of merging clusters in modern cosmology.

- Cosmological scales:

  Using the cosmic microwave background (CMB) information we can put stringent constraints on the dark matter and baryonic matter densities. Taking into account the latest data from Planck telescope [80] they inferred the following parameters:

$$\begin{aligned} \Omega_{b,0}h^2 &= 0.0222 \pm 0.0002 \\ \Omega_{dm,0}h^2 &= 0.1197 \pm 0.0022 \\ \Omega_{\Lambda,0} &= 0.685 \pm 0.013 \\ H_0 &= 67.31 \pm 0.96 \end{aligned}$$

  where $\Omega_{b,0}$ is the baryonic density, $\Omega_{dm,0}$ is the dark matter density, $\Omega_{\Lambda,0}$ is the density of the cosmological constant and H_0 is the Hubble constant.

In order to estimate these parameters you have to fit the power spectra of temperature fluctuations of the CMB. As it can be seen in figure 1.4, the power spectra is very sensitive to the content of dark matter, and so, you can use it to estimate it.

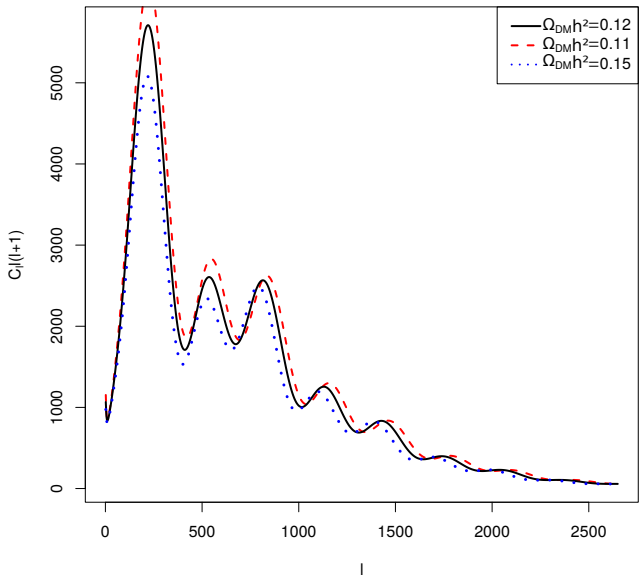

Figure 1.4: Power spectra of the cosmic microwave background for different values of $\Omega_{dm,0}$.

Dark matter general properties.

At the present time there a lot of candidates to be the dark matter particle, and so, is necessary to know the fundamental properties that it should have. It is also important to have in mind all the observations to which it has to be consistent. Taking all this into account we can list 6 properties that a particle must fulfil in order to be the dark matter [103].

1. Does it match the appropriate relic density?

   A viable dark matter candidate has to be an stable particle with a medium lifetime τ bigger than the universe age ($\tau \approx 4.3 * 10^{17} s$) guaranteeing its survival until the present time. At the same time, it must have a production mechanism that fulfil the wright values of the relic densities ($\Omega_{DM} \sim 0.1197$).

2. Is it cold?

The temporal evolution of the matter density perturbations depends on the microscopic properties of the dark matter. In the standard model, after the matter-radiation equality epoch (a_{eq}), the dark matter density perturbations start growing, and the baryons start going down to the potential wells forming the seeds where the galaxies will grow. Due to the non-collisional behaviour of the dark matter, there exist a characteristic minimum scale, that depends on the dark matter velocity (a bigger velocity a bigger scale), where the particles can cluster. All the observations suggest that the dark matter must be cold, i.e. it must have non-relativistic velocities before a_{eq}, in order to be able to cluster in galactic scales.

Another consequence of the col dark matter model (CDM) is the hierarchical structure formation. In this model, the smaller structures are form earlier and then, trough their mergers, the bigger structures are set up. On the other hand, if the dark matter would be hot (HDM) the structure formation would be the other way around. At earlier times the bigger structures would be formed and then, trough fragmentation, the smaller structure would be formed. This behaviour can be seen in the density perturbation power spectra where for a hot dark matter it has an exponential decay for large k, i.e. for small scales. It is worth to remark, that all the observational evidences [31, 75] and the cosmological simulations [96], points to a cold dark matter model.

3. Is it neutral?

   Although there are some candidates that have an electric charge (*CHAMPs* [45, 48], *SIMP* [98]) there are some studies that put stringent constraints that discard most of them [103]. Taking this into account, the standard model establishes that the dark matter particle must be neutral or have a very small cross-section $\sigma_{\gamma-dm} < 2.2510^{-6}\sigma_{th}(\frac{m_{dm}}{GeV})$ [97], where σ_{th} is the Thompson cross-section and m_{dm} is the dark matter particle mass.

4. Is it consistent with Big Bang Nucleosynthesis?

   The big bang nucleosynthesis theory predict the relic abundance of the chemical elements produced during the first 3 minutes after the big bang. In the standard model, the nucleosynthesis depends on the ratio between baryons and photons ν_b and in the expansion rate of the universe. Taking this into account, any new particle

that modifies this factors will affect the abundance of the primordial elements and may contradict the observations.

5. Does it leave stellar evolution unchanged?

   Some dark matter candidates can be generated in the hot plasma inside stars and, due to the non-collisional nature, may escape without interact with the other particles. This effect may cause a significant loose of energy modifying the star evolution. It can also be the case that the annihilation of dark matter particles produce an extra source of energy comparable to the nuclear reactions ([88], [93] and [53]), modifying the evolution of the stars and the properties of pop III stars ([84],[41] and [95]).

6. Is it compatible with constraints on self-interactions?

   Taking into account the latest observations, the dark matter must have a very small self-interaction cross-section σ.

   - *Gnedin y Ostriker* [56], shows that for $(0.3 \leq \sigma/m \leq 10^{14})cm^2/g$, the dark matter galactic haloes inside clusters would evaporated in a smaller time than the universe age.
   - *Natarajan et al.* [71], studying gravitational lenses of galactic haloes exclude values of $\sigma/m \geq 42cm^2/g$
   - An upper limit $\sigma/m < 0.1cm^2/g$ was obtained by *Arabadjis et al.* [20] comparing the density profiles of real and simulated haloes.
   - *Hennawi y Ostriker* [60], exclude values of σ/m lower than $0.02cm^2/g$ taking into account the masses of supermassive black holes in the centre of galaxies.
   - *Markevitch et al.* [70], studying the density distributions of dark matter and hot gas in the bullet cluster, put an upperlimit of $\sigma/m < 1cm^2/g$.
   - *Harvey et al.* [59] analysing statistically 72 merging clusters put an upperlimit of $\sigma/m < 0.47cm^2/g$

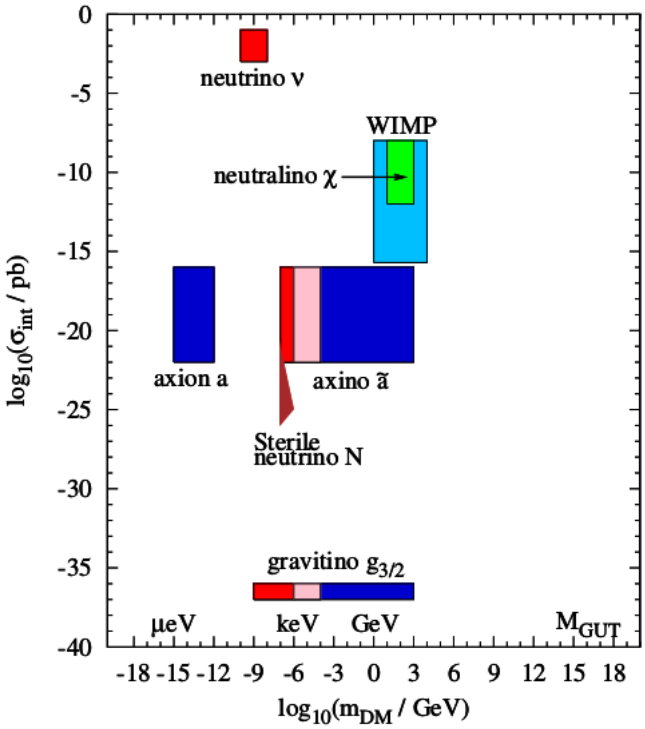

Figure 1.5: Interaction Cross-section between dark matter and baryons as a function of dark matter particle mass for the most important candidates. In red are shown hot dark matter candidates, in pink warm dark matter candidates and in blue cold dark matter candidates. Figure extracted from *Roszkowski et al. 2017* [86]

Dark matter particle candidates.

From the first observational evidences of the existence of the dark matter numerous candidates has been proposed. In figure 1.5 it can be seen the principal candidates order by their mass and interaction cross-section with standard model particles by which it may be detected. As it can be seen both parameters extend several magnitude orders making the *zoo* of candidates very wide.

One of the most important candidates are the WIMPS (Weakly interacting massive particles) that are predicted by some extensions of the standard model of particles. This hypothetical particles have masses around ($m_{DM} \sim GeV$) and only interact through the

weak interaction with standard model particles [86]. If you take into account the production mechanism predicted by the particle physics theory, you can obtained the correct relic abundance measured with cosmological observations. This fact is known as the *WIMP miracle* and is the reason why these candidates are one most studied ones. The WIMPs candidates include particles that comes from supersymmetric theories like the neutralinos and the sneutrinos [51] and particles that comes from gravity theories on higher dimensions like Kaluza-Klein [66] theory.

Besides WIMPS, another very studied candidates that come from extensions of the standard model of particles are the axions and the sterile neutrinos.

- Axions: This particle was postulated in order to resolve the CP symmetry violation problem [2]. The latest observations put stringent constraints in the mass of the axion ($m_{axion} \leq 0.01eV$) and it is expected that they interact very weakly with standard model particles, and so, they would no be in thermal equilibrium in the early universe. Although there is not an unique estimation of the relic density of the axions because you need to make some assumptions about the production mechanism, there is a mass range where the axions fulfil all the requirements and are a viable dark matter candidate [85]. It is worth to note that because of their mass axions may be considered hot dark matter, their properties make that they behave as cold dark matter for the structure formation.

- Sterile Neutrinos: Sterile neutrinos are hypothetical particles that were proposed to explained the results of the latest experiments with neutrinos. The sterile neutrinos have no charge, have $1/2$ spin and only interact with standard model particles through gravity. They were also proposed to explain an emission line in the γ-ray spectra of some extragalactic sources ([32] y [35]).

Experiment for the detection of dark matter.

The diversity of dark matter candidates imply that we must look for signals in very different ways and with different methods. Nowadays there are 4 basics methodologies to look for dark matter signals (Figure 1.6): Direct detection, that look for the signals of

[2]The CP symmetry is the composition of the C symmetry that says that physics laws have to be the same if we interchange the positive and negative charges. On the other hand, the P symmetry says that the physics laws should not change if we invert the spatial coordinates.

the interactions between dark matter and standard model particles, indirect detection, that look for the signals produced by dark matter annihilation or decay into standard model particles, particle collider experiments, that look for dark matter particles produced in the collisions of standard model particles and astrophysical experiments that studied how the dark matter properties may affect different astronomical objects (Section 1.2).

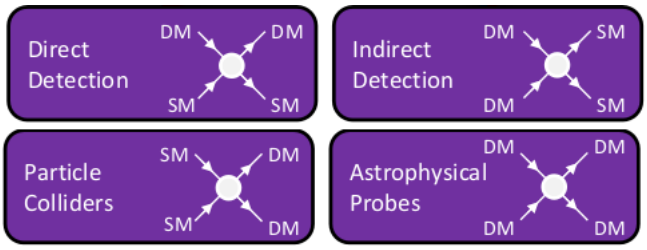

Figure 1.6: The 4 basics methodologies used in the dark matter search. *Bauer et al. 2013* [23]

- Direct Detection:

  It is expected that the dark matter is present all around the universe, and so, the earth is moving trough the dark matter halo of our galaxy. Taking into account the local dark matter density that we expected to have around the earth ($\rho_{DM} = 5 * 10^{-25} g/cm^3$]), we may detect the signals produced by the interactions between the dark matter particles and the standard model particles on earth. As was specify before, this interaction would be very weak, and so, it would be very difficult to detect. Another problem is that there is a lot of contamination sources like, γ-rays, β particles, muons and neutrons that are very difficult to distinguish from a real signal. Taking all this into account, the dark matter detectors are settled more than $2km$ below earth in order to reduce the contamination.

  Because it is not clear how the dark matter will interact with normal matter, it is common to study 2 cases: scalar interactions that do not depend on the spin and vector interactions that depends on the spin.

  Nowadays there are a lot of dark matter direct detection experiments, namely:

  - Cryogenic Dark Matter Search (CDMS) : This experiment is an array of semiconductors, settled in U.S.A., that measures temperatures up to milikelvin.

They were able to measure 3 events that were interpreted as the signal of an interaction between a dark matter particle with a mass of $8.6\ GeV/c^2$ and a cross-section of $\sigma = 1.9 * 10^{-41} cm^2$ [39]. Nevertheless, these results were not confirm by another independent experiment [5, 6].

- DAMA/LIBRA : This experiment look for signals with scintillator detectors [25]. They look for signals with an annual modulation that should be caused by the translation movement of the earth around the sun. They were able to measure a signal with an annual modulation and with a very high signal-to-noise ratio for more that 20 years. Nevertheless the interpretation of this signal as the interaction of a dark matter particle with standard model particles is inconsistent with the results found by another experiments [89].
- XENON : This detector is settled in Italy and uses liquid Xenon as a target. It have a sensitivity to detect a particle with a cross-section up to $2 * 10^{-45} cm^2$ and with a mass of $65 Gev/c^2$. They do not found any signal, discarding almost all the parameters explored by another experiments (see figure 1.7) and, in particular, refuting the results found by CDMS and DAMA/LIBRA ([16] y [13]).
- CoGeNT: This detector look for signals produced by the scattering between dark matter particles and germanium nuclei. They were able to measure a signal with an annual modulation consistent with a dark matter particle with a mass of $7-10\ GeV$ [1], nevertheless this signal only have a confidence level of 2.8σ and was not confirmed by another independent experiments in the same mass range.
- CRESST-II: This detector look for signals produced by the scattering between dark matter particles and $CaWO^4$ crystal nuclei [14]. They were able to measure an excess of events in the mass range of $10\ GeV$ and $25\ GeV$ with a confidence level of 4.2σ y 4.7σ respectively. Nevertheless, after that they demonstrate that this excess were produced by contamination sources that were not taking into account previously [15, 68].

In figure 1.7, we show the parameter space for an scalar interaction (spin independent) and the zones of exclusion found by different experiments.

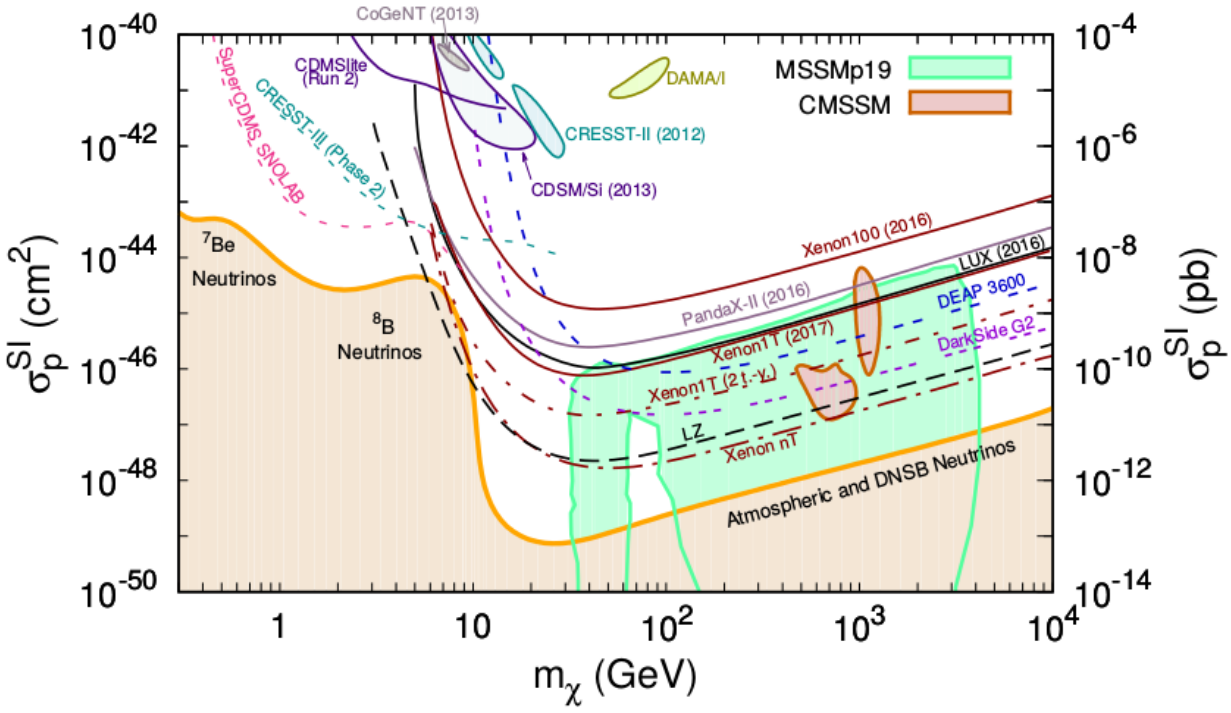

Figure 1.7: Constraint in the cross-section vs mass space for scalar interactions. The limits found by different experiments are: LUX in solid black lines [10], in grey are plot PandaX-II results [74], in brown lines are shown XENON100 and XENON1T [18, 19] and in violet lines CDMSlite-II [101]. The predicted limits for future projects are: CRESST-III [100] in light-blue, DarkSide G2 [2] in dotted violet lines, DEAP3600 [46] in dotted blue lines, LZ [102] in black dotted line, SuperCDMS and SNOLAB [7] in pink dotted line and, in dotted brown line, XENON1T/nT [17]. We also plot in green the area favoured by the minimal super-symmetric models (MSSM) and in brown the area favoured by the constraint minimal supersymmetric model (CMSSM). In blue shadow areas are shown the events observed by CDMS (blue), CoGenT (grey), CRESST-II (light-blue) y DAMA-LIBRA (light green). Finally, the shadow orange area at the bottom of the plot correspond to the irreducible neutrino background [28]. Figure extracted from *Roszkowski et al. 2017* [86].

- Indirect Detection:

  The indirect detection methodology look for signals produced by the annihilation and the decay of dark matter particles into standard model particles. It is worth to note that, as was specify before, is expected that the dark matter is stable or have a

medium lifetime longer that the universe age, then it is expected that the signal from decay will be negligible. The signals from annihilation could be very energetic photons (γ-ray or X-ray), neutrinos or cosmic rays (positrons, electrons, etc.) and their energy depends on the properties of the dark matter. The flux of the these radiation will depends on the annihilation rate that is proportional to the square of the dark matter density $\Gamma_A \propto \rho^2_{DM}$ [27]. This is why it is expected that the signal will be strong towards the galactic centre where the dark matter density is bigger.

Among the most important projects for indirect detection is the γ-ray Fermi spatial telescope. This telescope was able to observed the Milky-way centre (figure 1.8), some Milky-way satellites and the Coma clusters, and to established constraints that exclude some dark matter candidates ([55] and [104]). It is important to remark also the telescopes HESS (*High Energy Stereoscopic System*) [8], VERITAS [62], MAGIC [12] and CTA (*Cherenkov telescope array*) [36], which are designed to detect the *Cherenkov* radiation produced by the interaction between γ-rays and the earth atmosphere. In figure 1.9 we show the current constraints obtained by the different telescopes. We also show the predicted limits for future experiments.

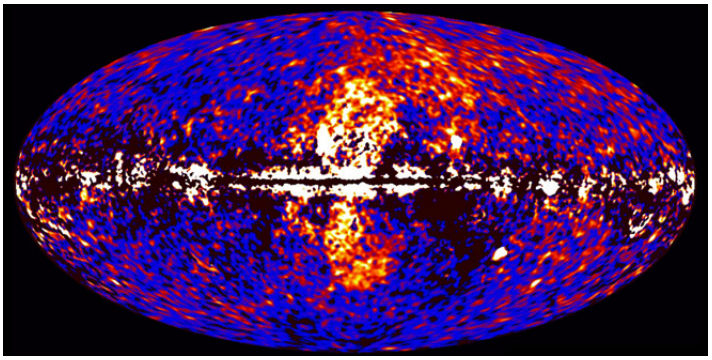

Figure 1.8: Full-sky map of the γ-ray emission observed with the Fermi spatial telescope FERMI. Figure extracted from http://www.nasa.gov

It is also important to remark 2 works ([32] y [35]) where, through the analysis of the γ-ray emission from different extragalactic sources (Andromeda galaxy, Perseus Cluster, Coma Cluster, etc.), they identify an emission line of $E = 3.55-3.57\pm0.03$ KeV that does not match with any known emission line. This new line can be interpreted as the emission produced by the decay of an sterile neutrino with a mass of $m_{sn} = 7.06\pm0.05 KeV$, into photons and neutrinos. This interpretation is consistent

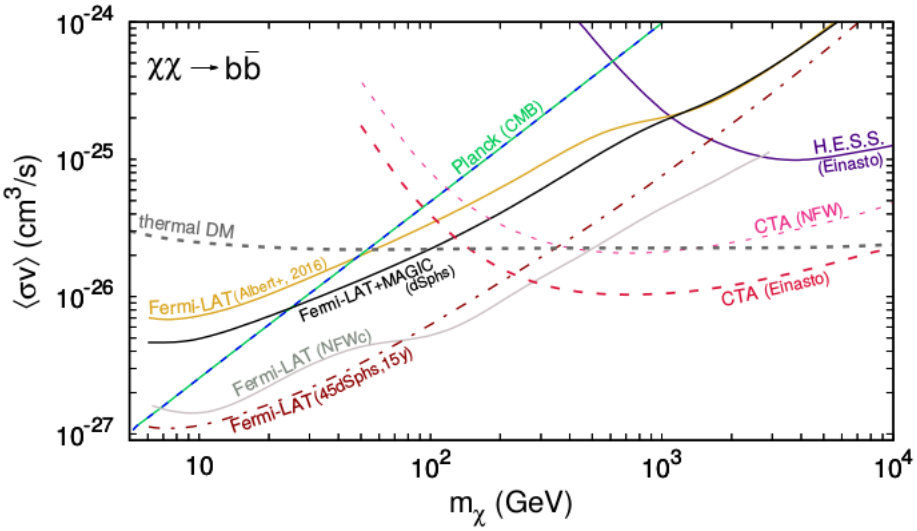

Figure 1.9: Current constrains in the mass vs cross-section space for the annihilation of dark matter particles. In gold line is shown the results obtained with the analysis of dwarf galaxies with Fermi-LAT [11] data. In black line is plotted the analysis combining Fermi-LAT and MAGIC [69] data. The results obtained with HESS observations [3] are shown in violet lines, while the constraints obtained with Fermi-LAT observations of the Milky-way centre [57] are shown in grey lines. The results obtained by Planck collaboration through the analysis of CMB data [79] are plotted in dotted blue and green lines. The constraints obtained with the joined analysis of 45 dwarf galaxies [40] are shown in dotted brown lines, while the limits obtained by CTA collaboration [4] assuming an NFW profile are shown in dotted pink lines, while the constraints obtained assuming an Einasto profile are shown in dotted red line. In grey dotted line is shown the cross section of the annihilation that correspond to the thermal production of WIMPs particles [99]. Figure extracted from [87]

with all the constraints that comes from cosmological observations. Although this results were confirmed with Suzaku telescope observations of the Perseus Cluster [105] and for the observations of XMM-Newton telescope of galactic centre [33], the Suzaku observations of other galaxy clusters [105] and the HITOMI XMM-Newton telescope observation of Perseus cluster [9] do not found such emission line. It is also worth to remark different works that discuss possible astrophysics explanations for this signal, including emission lines of potassium and chlorine atoms transitions [105] or the charge interchange between sulphide ions and neutral hydro-

gen atoms [58, 94]. They also found similar signals in supernovas remnants [64].

- Particle colliders:

  The dark matter can also be produced in the collide of very energetic standard model particles. If a dark matter particle is produced in this way it would escape through the detectors without producing any signal. Nevertheless its presence can be deduced by a decrease of the momentum of the detected particles, this is why particle colliders are more sensitive to lighter dark matter particles, that may be produced with a higher momentum [23].

- Astrophysical probes:

  Taking into account the specific properties of each dark matter candidate, and how this properties may affect some astrophysical phenomena, we can put stringent constraints in certain dark matter parameters. As an example, the coupling between dark matter with standard model particles can affect the cooling of compact objects (stars, neutron stars, white dwarfs, etc.) and the transparency of the extragalactic medium [22, 63]. In a similar manner, the fluctuations of the CMB are sensitive to the annihilation rate of the dark matter during the recombination epoch, that may change significantly the power spectra [73]. It is also work to remark that even very small interactions between dark matter and photons may have important consequences either in the power spectra of the CMB [107] and in the large scale structure formation of the universe [29, 90, 91].

1.2.2 Dark Energy.

The discovery of the acceleration expansion rate of the universe ([82] and [76]), give place to a new era in modern cosmology. This result is one of the most important evidence that the $\approx 70\%$ of the universe is composed of a new form of matter, different either from baryonic matter and dark matter, known as dark energy. This results were also confirmed by other cosmological probes as the CMB fluctuations [78] and the baryonic acoustic oscillations (BAO) [50].

The dark energy is characterised by an state equation $P_{DE} = \omega_{DE}\rho_{DE}$, where P_{DE} is the pressure and ρ_{DE} is the dark energy density. In order to produce an accelerated cosmic expansion the dark energy should have $\omega_{DE} < -1/3$.

One of the principal candidates to dark energy is the cosmological constant Λ that have $\omega_\Lambda = -1$. Although this constant could be interpreted as the vacuum energy, there are several order of magnitude of difference between the value measured in cosmological observations with the one measured through particle physics experiments [106]. Another authors proposed a new type of matter as the quinteseence ([54], [81]), Chaplygin gas [65] or the k-essence ([42], [21]). On the other hand *Buchert et al. 2000* [34] proposed that the inhomogeneities in the large scale matter distribution may produce an accelerated expansion similar to the one observed.

Bibliography

[1] C. E. Aalseth, P. S. Barbeau, J. Colaresi, J. I. Collar, J. Diaz Leon, J. E. Fast, N. E. Fields, T. W. Hossbach, A. Knecht, M. S. Kos, M. G. Marino, H. S. Miley, M. L. Miller, J. L. Orrell, and K. M. Yocum. CoGeNT: A search for low-mass dark matter using p-type point contact germanium detectors,. *PhRvD*, 88(1):012002, July 2013. doi: 10.1103/PhysRevD.88.012002.

[2] C. E. Aalseth, F. Acerbi, P. Agnes, I. F. M. Albuquerque, T. Alexander, A. Alici, A. K. Alton, P. Antonioli, S. Arcelli, R. Ardito, and et al. DarkSide-20k: A 20 Tonne Two-Phase LAr TPC for Direct Dark Matter Detection at LNGS,. *ArXiv e-prints*, July 2017.

[3] H. Abdallah, A. Abramowski, F. Aharonian, F. Ait Benkhali, A. G. Akhperjanian, E. Anguner, M. Arrieta, P. Aubert, M. Backes, A. Balzer, and et al. Search for Dark Matter Annihilations towards the Inner Galactic Halo from 10 Years of Observations with H.E.S.S.,. *Physical Review Letters*, 117(11):111301, September 2016. doi: 10.1103/PhysRevLett.117.111301.

[4] M. Actis, G. Agnetta, F. Aharonian, A. Akhperjanian, J. Aleksic, E. Aliu, D. Allan, I. Allekotte, F. Antico, L. A. Antonelli, and et al. Design concepts for the Cherenkov Telescope Array CTA: an advanced facility for ground-based high-energy gamma-ray astronomy,. *Experimental Astronomy*, 32:193–316, December 2011. doi: 10.1007/s10686-011-9247-0.

[5] R. Agnese, A. J. Anderson, M. Asai, D. Balakishiyeva, R. Basu Thakur, D. A. Bauer, J. Billard, A. Borgland, M. A. Bowles, D. Brandt, P. L. Brink, R. Bunker, B. Cabrera, D. O. Caldwell, D. G. Cerdeno, H. Chagani, J. Cooley, B. Cornell,

C. H. Crewdson, P. Cushman, M. Daal, P. C. F. Di Stefano, T. Doughty, L. Esteban, S. Fallows, E. Figueroa-Feliciano, G. L. Godfrey, S. R. Golwala, J. Hall, H. R. Harris, S. A. Hertel, T. Hofer, D. Holmgren, L. Hsu, M. E. Huber, A. Jastram, O. Kamaev, B. Kara, M. H. Kelsey, A. Kennedy, M. Kiveni, K. Koch, B. Loer, E. Lopez Asamar, R. Mahapatra, V. Mandic, C. Martinez, K. A. McCarthy, N. Mirabolfathi, R. A. Moffatt, D. C. Moore, P. Nadeau, R. H. Nelson, K. Page, R. Partridge, M. Pepin, A. Phipps, K. Prasad, M. Pyle, H. Qiu, W. Rau, P. Redl, A. Reisetter, Y. Ricci, T. Saab, B. Sadoulet, J. Sander, K. Schneck, R. W. Schnee, S. Scorza, B. Serfass, B. Shank, D. Speller, A. N. Villano, B. Welliver, D. H. Wright, S. Yellin, J. J. Yen, B. A. Young, J. Zhang, and SuperCDMS Collaboration. Search for Low-Mass Weakly Interacting Massive Particles Using Voltage-Assisted Calorimetric Ionization Detection in the SuperCDMS Experiment,. *Physical Review Letters*, 112(4):041302, January 2014. doi: 10.1103/PhysRevLett.112.041302.

[6] R. Agnese, A. J. Anderson, D. Balakishiyeva, R. Basu Thakur, D. A. Bauer, J. Billard, A. Borgland, M. A. Bowles, D. Brandt, P. L. Brink, R. Bunker, B. Cabrera, D. O. Caldwell, D. G. Cerdeno, H. Chagani, Y. Chen, J. Cooley, B. Cornell, C. H. Crewdson, P. Cushman, M. Daal, P. C. F. Di Stefano, T. Doughty, L. Esteban, S. Fallows, E. Figueroa-Feliciano, M. Fritts, G. L. Godfrey, S. R. Golwala, M. Graham, J. Hall, H. R. Harris, S. A. Hertel, T. Hofer, D. Holmgren, L. Hsu, M. E. Huber, A. Jastram, O. Kamaev, B. Kara, M. H. Kelsey, A. Kennedy, M. Kiveni, K. Koch, A. Leder, B. Loer, E. Lopez Asamar, R. Mahapatra, V. Mandic, C. Martinez, K. A. McCarthy, N. Mirabolfathi, R. A. Moffatt, D. C. Moore, R. H. Nelson, S. M. Oser, K. Page, W. A. Page, R. Partridge, M. Pepin, A. Phipps, K. Prasad, M. Pyle, H. Qiu, W. Rau, P. Redl, A. Reisetter, Y. Ricci, H. E. Rogers, T. Saab, B. Sadoulet, J. Sander, K. Schneck, R. W. Schnee, S. Scorza, B. Serfass, B. Shank, D. Speller, S. Upadhyayula, A. N. Villano, B. Welliver, D. H. Wright, S. Yellin, J. J. Yen, B. A. Young, J. Zhang, and SuperCDMS Collaboration. Maximum likelihood analysis of low energy CDMS II germanium data,. *PhRvD*, 91(5):052021, March 2015. doi: 10.1103/PhysRevD.91.052021.

[7] R. Agnese, A. J. Anderson, T. Aramaki, I. Arnquist, W. Baker, D. Barker, R. Basu Thakur, D. A. Bauer, A. Borgland, M. A. Bowles, P. L. Brink, R. Bunker, B. Cabrera, D. O. Caldwell, R. Calkins, C. Cartaro, D. G. Cerdeno, H. Chagani, Y. Chen,

J. Cooley, B. Cornell, P. Cushman, M. Daal, P. C. F. Di Stefano, T. Doughty, L. Esteban, S. Fallows, E. Figueroa-Feliciano, M. Fritts, G. Gerbier, M. Ghaith, G. L. Godfrey, S. R. Golwala, J. Hall, H. R. Harris, T. Hofer, D. Holmgren, Z. Hong, E. Hoppe, L. Hsu, M. E. Huber, V. Iyer, D. Jardin, A. Jastram, M. H. Kelsey, A. Kennedy, A. Kubik, N. A. Kurinsky, A. Leder, B. Loer, E. Lopez Asamar, P. Lukens, R. Mahapatra, V. Mandic, N. Mast, N. Mirabolfathi, R. A. Moffatt, J. D. Morales Mendoza, J. L. Orrell, S. M. Oser, K. Page, W. A. Page, R. Partridge, M. Pepin, A. Phipps, S. Poudel, M. Pyle, H. Qiu, W. Rau, P. Redl, A. Reisetter, A. Roberts, A. E. Robinson, H. E. Rogers, T. Saab, B. Sadoulet, J. Sander, K. Schneck, R. W. Schnee, B. Serfass, D. Speller, M. Stein, J. Street, H. A. Tanaka, D. Toback, R. Underwood, A. N. Villano, B. von Krosigk, B. Welliver, J. S. Wilson, D. H. Wright, S. Yellin, J. J. Yen, B. A. Young, X. Zhang, X. Zhao, and SuperCDMS Collaboration. Projected sensitivity of the SuperCDMS SNOLAB experiment,. *PhRvD*, 95(8):082002, April 2017. doi: 10.1103/PhysRevD.95.082002.

[8] F. Aharonian, A. G. Akhperjanian, A. R. Bazer-Bachi, M. Beilicke, W. Benbow, D. Berge, K. Bernlohr, C. Boisson, O. Bolz, V. Borrel, I. Braun, F. Breitling, A. M. Brown, R. Buhler, I. Busching, S. Carrigan, P. M. Chadwick, L.-M. Chounet, R. Cornils, L. Costamante, B. Degrange, H. J. Dickinson, A. Djannati-Atai, L. O'C.Drury, G. Dubus, K. Egberts, D. Emmanoulopoulos, P. Espigat, F. Feinstein, E. Ferrero, A. Fiasson, G. Fontaine, S. Funk, S. Funk, Y. A. Gallant, B. Giebels, J. F. Glicenstein, P. Goret, C. Hadjichristidis, D. Hauser, M. Hauser, G. Heinzelmann, G. Henri, G. Hermann, J. A. Hinton, W. Hofmann, M. Holleran, D. Horns, A. Jacholkowska, O. C. de Jager, B. Khelifi, N. Komin, A. Konopelko, K. Kosack, I. J. Latham, R. Le Gallou, A. Lemiere, M. Lemoine-Goumard, T. Lohse, J. M. Martin, O. Martineau-Huynh, A. Marcowith, C. Masterson, T. J. L. McComb, M. de Naurois, D. Nedbal, S. J. Nolan, A. Noutsos, K. J. Orford, J. L. Osborne, M. Ouchrif, M. Panter, G. Pelletier, S. Pita, G. Puhlhofer, M. Punch, B. C. Raubenheimer, M. Raue, S. M. Rayner, A. Reimer, O. Reimer, J. Ripken, L. Rob, L. Rolland, G. Rowell, V. Sahakian, L. Sauge, S. Schlenker, R. Schlickeiser, U. Schwanke, H. Sol, D. Spangler, F. Spanier, R. Steenkamp, C. Stegmann, G. Superina, J.-P. Tavernet, R. Terrier, C. G. Theoret, M. Tluczykont, C. van Eldik, G. Vasileiadis, C. Venter, P. Vincent, H. J. Volk, S. J. Wagner, and M. Ward. Ob-

servations of the Crab nebula with HESS,. *A&A*, 457:899–915, October 2006. doi: 10.1051/0004-6361:20065351.

[9] F. A. Aharonian, H. Akamatsu, F. Akimoto, S. W. Allen, L. Angelini, K. A. Arnaud, M. Audard, H. Awaki, M. Axelsson, A. Bamba, and et al. Hitomi Constraints on the 3.5 keV Line in the Perseus Galaxy Cluster,. *ApJL*, 837:L15, March 2017. doi: 10.3847/2041-8213/aa61fa.

[10] D. S. Akerib, S. Alsum, H. M. Araú,jo, X. Bai, A. J. Bailey, J. Balajthy, P. Beltrame, E. P. Bernard, A. Bernstein, T. P. Biesiadzinski, E. M. Boulton, R. Bramante, P. Bra,s, D. Byram, S. B. Cahn, M. C. Carmona-Benitez, C. Chan, A. A. Chiller, C. Chiller, A. Currie, J. E. Cutter, T. J. R. Davison, A. Dobi, J. E. Y. Dobson, E. Druszkiewicz, B. N. Edwards, C. H. Faham, S. Fiorucci, R. J. Gaitskell, V. M. Gehman, C. Ghag, K. R. Gibson, M. G. D. Gilchriese, C. R. Hall, M. Hanhardt, S. J. Haselschwardt, S. A. Hertel, D. P. Hogan, M. Horn, D. Q. Huang, C. M. Ignarra, M. Ihm, R. G. Jacobsen, W. Ji, K. Kamdin, K. Kazkaz, D. Khaitan, R. Knoche, N. A. Larsen, C. Lee, B. G. Lenardo, K. T. Lesko, A. Lindote, M. I. Lopes, A. Manalaysay, R. L. Mannino, M. F. Marzioni, D. N. McKinsey, D.-M. Mei, J. Mock, M. Moongweluwan, J. A. Morad, A. S. J. Murphy, C. Nehrkorn, H. N. Nelson, F. Neves, K. O'Sullivan, K. C. Oliver-Mallory, K. J. Palladino, E. K. Pease, P. Phelps, L. Reichhart, C. Rhyne, S. Shaw, T. A. Shutt, C. Silva, M. Solmaz, V. N. Solovov, P. Sorensen, S. Stephenson, T. J. Sumner, M. Szydagis, D. J. Taylor, W. C. Taylor, B. P. Tennyson, P. A. Terman, D. R. Tiedt, W. H. To, M. Tripathi, L. Tvrznikova, S. Uvarov, J. R. Verbus, R. C. Webb, J. T. White, T. J. Whitis, M. S. Witherell, F. L. H. Wolfs, J. Xu, K. Yazdani, S. K. Young, C. Zhang, and LUX Collaboration. Results from a Search for Dark Matter in the Complete LUX Exposure,. *Physical Review Letters*, 118(2):021303, January 2017. doi: 10.1103/PhysRevLett.118.021303.

[11] A. Albert, B. Anderson, K. Bechtol, A. Drlica-Wagner, M. Meyer, M. Sanchez-Conde, L. Strigari, M. Wood, T. M. C. Abbott, F. B. Abdalla, A. Benoit-Levy, G. M. Bernstein, R. A. Bernstein, E. Bertin, D. Brooks, D. L. Burke, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, M. Crocce, C. E. Cunha, C. B. D'Andrea, L. N. da Costa, S. Desai, H. T. Diehl, J. P. Dietrich, P. Doel, T. F. Eifler, A. E. Evrard,

A. Fausti Neto, D. A. Finley, B. Flaugher, P. Fosalba, J. Frieman, D. W. Gerdes, D. A. Goldstein, D. Gruen, R. A. Gruendl, K. Honscheid, D. J. James, S. Kent, K. Kuehn, N. Kuropatkin, O. Lahav, T. S. Li, M. A. G. Maia, M. March, J. L. Marshall, P. Martini, C. J. Miller, R. Miquel, E. Neilsen, B. Nord, R. Ogando, A. A. Plazas, K. Reil, A. K. Romer, E. S. Rykoff, E. Sanchez, B. Santiago, M. Schubnell, I. Sevilla-Noarbe, R. C. Smith, M. Soares-Santos, F. Sobreira, E. Suchyta, M. E. C. Swanson, G. Tarle, V. Vikram, A. R. Walker, R. H. Wechsler, Fermi-LAT Collaboration, and DES Collaboration. Searching for Dark Matter Annihilation in Recently Discovered Milky Way Satellites with Fermi-Lat,. *ApJ*, 834:110, January 2017. doi: 10.3847/1538-4357/834/2/110.

[12] J. Aleksic, E. A. Alvarez, L. A. Antonelli, P. Antoranz, M. Asensio, M. Backes, J. A. Barrio, D. Bastieri, J. Becerra Gonzalez, W. Bednarek, A. Berdyugin, K. Berger, E. Bernardini, A. Biland, O. Blanch, R. K. Bock, A. Boller, G. Bonnoli, D. Borla Tridon, I. Braun, T. Bretz, A. Canellas, E. Carmona, A. Carosi, P. Colin, E. Colombo, J. L. Contreras, J. Cortina, L. Cossio, S. Covino, F. Dazzi, A. de Angelis, G. de Caneva, E. de Cea Del Pozo, B. de Lotto, C. Delgado Mendez, A. Diago Ortega, M. Doert, A. Dominguez, D. Dominis Prester, D. Dorner, M. Doro, D. Elsaesser, D. Ferenc, M. V. Fonseca, L. Font, C. Fruck, R. J. Garcia Lopez, M. Garczarczyk, D. Garrido, G. Giavitto, N. Godinovic, D. Hadasch, D. Hafner, A. Herrero, D. Hildebrand, D. Hohne-Monch, J. Hose, D. Hrupec, B. Huber, T. Jogler, H. Kellermann, S. Klepser, T. Krahenbuhl, J. Krause, A. La Barbera, D. Lelas, E. Leonardo, E. Lindfors, S. Lombardi, M. Lopez, A. Lopez-Oramas, E. Lorenz, M. Makariev, G. Maneva, N. Mankuzhiyil, K. Mannheim, L. Maraschi, M. Mariotti, M. Martinez, D. Mazin, M. Meucci, J. M. Miranda, R. Mirzoyan, H. Miyamoto, J. Moldon, A. Moralejo, P. Munar-Adrover, D. Nieto, K. Nilsson, R. Orito, I. Oya, D. Paneque, R. Paoletti, S. Pardo, J. M. Paredes, S. Partini, M. Pasanen, F. Pauss, M. A. Perez-Torres, M. Persic, L. Peruzzo, M. Pilia, J. Pochon, F. Prada, P. G. Prada Moroni, E. Prandini, I. Puljak, I. Reichardt, R. Reinthal, W. Rhode, M. Ribo, J. Rico, S. Rugamer, A. Saggion, K. Saito, T. Y. Saito, M. Salvati, K. Satalecka, V. Scalzotto, V. Scapin, C. Schultz, T. Schweizer, M. Shayduk, S. N. Shore, A. Sillanpaa, J. Sitarek, I. Snidaric, D. Sobczynska, F. Spanier, S. Spiro, V. Stamatescu, A. Stamerra, B. Steinke, J. Storz, N. Strah,

T. Suric, L. Takalo, H. Takami, F. Tavecchio, P. Temnikov, T. Terzic, D. Tescaro, M. Teshima, O. Tibolla, D. F. Torres, A. Treves, M. Uellenbeck, H. Vankov, P. Vogler, R. M. Wagner, Q. Weitzel, V. Zabalza, F. Zandanel, and R. Zanin. Performance of the MAGIC stereo system obtained with Crab Nebula data,. *Astroparticle Physics*, 35:435–448, February 2012. doi: 10.1016/j.astropartphys.2011.11.007.

[13] J. Angle. Erratum: Search for Light Dark Matter in XENON10 Data [Phys. Rev. Lett. 107, 051301 (2011)],. *Physical Review Letters*, 110(24):249901, June 2013. doi: 10.1103/PhysRevLett.110.249901.

[14] G. Angloher, M. Bauer, I. Bavykina, A. Bento, C. Bucci, C. Ciemniak, G. Deuter, F. von Feilitzsch, D. Hauff, P. Huff, C. Isaila, J. Jochum, M. Kiefer, M. Kimmerle, J.-C. Lanfranchi, F. Petricca, S. Pfister, W. Potzel, F. Probst, F. Reindl, S. Roth, K. Rottler, C. Sailer, K. Schaffner, J. Schmaler, S. Scholl, W. Seidel, M. v. Sivers, L. Stodolsky, C. Strandhagen, R. Strauss, A. Tanzke, I. Usherov, S. Wawoczny, M. Willers, and A. Zoller. Results from 730 kg days of the CRESST-II Dark Matter search,. *European Physical Journal C*, 72:1971, April 2012. doi: 10.1140/epjc/s10052-012-1971-8.

[15] G. Angloher, A. Bento, C. Bucci, L. Canonica, A. Erb, F. von Feilitzsch, N. F. Iachellini, P. Gorla, A. Gutlein, D. Hauff, P. Huff, J. Jochum, M. Kiefer, C. Kister, H. Kluck, H. Kraus, J.-C. Lanfranchi, J. Loebell, A. Munster, F. Petricca, W. Potzel, F. Probst, F. Reindl, S. Roth, K. Rottler, C. Sailer, K. Schaffner, J. Schieck, J. Schmaler, S. Scholl, S. Schonert, W. Seidel, M. von Sivers, L. Stodolsky, C. Strandhagen, R. Strauss, A. Tanzke, M. Uffinger, A. Ulrich, I. Usherov, S. Wawoczny, M. Willers, M. Wustrich, and A. Zoller. Results on low mass WIMPs using an upgraded CRESST-II detector,. *European Physical Journal C*, 74:3184, December 2014. doi: 10.1140/epjc/s10052-014-3184-9.

[16] E. Aprile, M. Alfonsi, K. Arisaka, F. Arneodo, C. Balan, L. Baudis, B. Bauermeister, A. Behrens, P. Beltrame, K. Bokeloh, A. Brown, E. Brown, G. Bruno, R. Budnik, J. M. R. Cardoso, W.-T. Chen, B. Choi, A. P. Colijn, H. Contreras, J. P. Cussonneau, M. P. Decowski, E. Duchovni, S. Fattori, A. D. Ferella, W. Fulgione, F. Gao, M. Garbini, C. Ghag, K.-L. Giboni, L. W. Goetzke, C. Grignon, E. Gross,

W. Hampel, F. Kaether, A. Kish, J. Lamblin, H. Landsman, R. F. Lang, M. Le Calloch, D. Lellouch, C. Levy, K. E. Lim, Q. Lin, S. Lindemann, M. Lindner, J. A. M. Lopes, K. Lung, T. Marrodan Undagoitia, F. V. Massoli, A. J. Melgarejo Fernandez, Y. Meng, M. Messina, A. Molinario, K. Ni, U. Oberlack, S. E. A. Orrigo, E. Pantic, R. Persiani, G. Plante, N. Priel, A. Rizzo, S. Rosendahl, J. M. F. dos Santos, G. Sartorelli, J. Schreiner, M. Schumann, L. Scotto Lavina, P. R. Scovell, M. Selvi, P. Shagin, H. Simgen, A. Teymourian, D. Thers, O. Vitells, H. Wang, M. Weber, and C. Weinheimer. Limits on Spin-Dependent WIMP-Nucleon Cross Sections from 225 Live Days of XENON100 Data,. *Physical Review Letters*, 111 (2):021301, July 2013. doi: 10.1103/PhysRevLett.111.021301.

[17] E. Aprile, J. Aalbers, F. Agostini, M. Alfonsi, F. D. Amaro, M. Anthony, L. Arazi, F. Arneodo, C. Balan, P. Barrow, L. Baudis, B. Bauermeister, T. Berger, P. Breur, A. Breskin, A. Brown, E. Brown, S. Bruenner, G. Bruno, R. Budnik, L. Butikofer, J. M. R. Cardoso, M. Cervantes, D. Cichon, D. Coderre, A. P. Colijn, J. Conrad, H. Contreras, J. P. Cussonneau, M. P. Decowski, P. de Perio, P. Di Gangi, A. Di Giovanni, E. Duchovni, S. Fattori, A. D. Ferella, A. Fieguth, D. Franco, W. Fulgione, M. Galloway, M. Garbini, C. Geis, L. W. Goetzke, Z. Greene, C. Grignon, E. Gross, W. Hampel, C. Hasterok, R. Itay, F. Kaether, B. Kaminsky, G. Kessler, A. Kish, H. Landsman, R. F. Lang, D. Lellouch, L. Levinson, M. Le Calloch, C. Levy, S. Lindemann, M. Lindner, J. A. M. Lopes, A. Lyashenko, S. Macmullin, A. Manfredini, T. Marroda,n Undagoitia, J. Masbou, F. V. Massoli, D. Mayani, A. J. Melgarejo Fernandez, Y. Meng, M. Messina, K. Micheneau, B. Miguez, A. Molinario, M. Murra, J. Naganoma, U. Oberlack, S. E. A. Orrigo, P. Pakarha, B. Pelssers, R. Persiani, F. Piastra, J. Pienaar, G. Plante, N. Priel, L. Rauch, S. Reichard, C. Reuter, A. Rizzo, S. Rosendahl, N. Rupp, J. M. F. dos Santos, G. Sartorelli, M. Scheibelhut, S. Schindler, J. Schreiner, M. Schumann, L. Scotto Lavina, M. Selvi, P. Shagin, H. Simgen, A. Stein, D. Thers, A. Tiseni, G. Trinchero, C. Tunnell, M. von Sivers, R. Wall, H. Wang, M. Weber, Y. Wei, C. Weinheimer, J. Wulf, and Y. Zhang. Physics reach of the XENON1T dark matter experiment.,. *JCAP*, 4: 027, April 2016. doi: 10.1088/1475-7516/2016/04/027.

[18] E. Aprile, J. Aalbers, F. Agostini, M. Alfonsi, F. D. Amaro, M. Anthony, F. Arneodo, P. Barrow, L. Baudis, B. Bauermeister, M. L. Benabderrahmane, T. Berger,

P. A. Breur, A. Brown, E. Brown, S. Bruenner, G. Bruno, R. Budnik, L. Butikofer, J. Calven, J. M. R. Cardoso, M. Cervantes, D. Cichon, D. Coderre, A. P. Colijn, J. Conrad, J. P. Cussonneau, M. P. Decowski, P. de Perio, P. di Gangi, A. di Giovanni, S. Diglio, E. Duchovni, J. Fei, A. D. Ferella, A. Fieguth, D. Franco, W. Fulgione, A. Gallo Rosso, M. Galloway, F. Gao, M. Garbini, C. Geis, L. W. Goetzke, Z. Greene, C. Grignon, C. Hasterok, E. Hogenbirk, R. Itay, B. Kaminsky, G. Kessler, A. Kish, H. Landsman, R. F. Lang, D. Lellouch, L. Levinson, M. Le Calloch, C. Levy, Q. Lin, S. Lindemann, M. Lindner, J. A. M. Lopes, A. Manfredini, T. Marroda,n Undagoitia, J. Masbou, F. V. Massoli, D. Masson, D. Mayani, Y. Meng, M. Messina, K. Micheneau, B. Miguez, A. Molinario, M. Murra, J. Naganoma, K. Ni, U. Oberlack, S. E. A. Orrigo, P. Pakarha, B. Pelssers, R. Persiani, F. Piastra, J. Pienaar, M.-C. Piro, G. Plante, N. Priel, L. Rauch, S. Reichard, C. Reuter, A. Rizzo, S. Rosendahl, N. Rupp, J. M. F. Dos Santos, G. Sartorelli, M. Scheibelhut, S. Schindler, J. Schreiner, M. Schumann, L. Scotto Lavina, M. Selvi, P. Shagin, M. Silva, H. Simgen, M. V. Sivers, A. Stein, D. Thers, A. Tiseni, G. Trinchero, C. D. Tunnell, R. Wall, H. Wang, M. Weber, Y. Wei, C. Weinheimer, J. Wulf, Y. Zhang, and Xenon Collaboration. XENON100 dark matter results from a combination of 477 live days,. *PhRvD*, 94(12):122001, December 2016. doi: 10.1103/PhysRevD.94.122001.

[19] E. Aprile, J. Aalbers, F. Agostini, M. Alfonsi, F. D. Amaro, M. Anthony, F. Arneodo, P. Barrow, L. Baudis, B. Bauermeister, M. L. Benabderrahmane, T. Berger, P. A. Breur, A. Brown, A. Brown, E. Brown, S. Bruenner, G. Bruno, R. Budnik, L. Butikofer, J. Calven, J. M. R. Cardoso, M. Cervantes, D. Cichon, D. Coderre, A. P. Colijn, J. Conrad, J. P. Cussonneau, M. P. Decowski, P. de Perio, P. Di Gangi, A. Di Giovanni, S. Diglio, G. Eurin, J. Fei, A. D. Ferella, A. Fieguth, W. Fulgione, A. Gallo Rosso, M. Galloway, F. Gao, M. Garbini, R. Gardner, C. Geis, L. W. Goetzke, L. Grandi, Z. Greene, C. Grignon, C. Hasterok, E. Hogenbirk, J. Howlett, R. Itay, B. Kaminsky, S. Kazama, G. Kessler, A. Kish, H. Landsman, R. F. Lang, D. Lellouch, L. Levinson, Q. Lin, S. Lindemann, M. Lindner, F. Lombardi, J. A. M. Lopes, A. Manfredini, I. Marics, T. Marroda,n Undagoitia, J. Masbou, F. V. Massoli, D. Masson, D. Mayani, M. Messina, K. Micheneau, A. Molinario, K. Moraa,, M. Murra, J. Naganoma, K. Ni, U. Oberlack, P. Pakarha, B. Pelssers, R. Persiani,

F. Piastra, J. Pienaar, V. Pizzella, M.-C. Piro, G. Plante, N. Priel, L. Rauch, S. Reichard, C. Reuter, B. Riedel, A. Rizzo, S. Rosendahl, N. Rupp, R. Saldanha, J. M. F. dos Santos, G. Sartorelli, M. Scheibelhut, S. Schindler, J. Schreiner, M. Schumann, L. Scotto Lavina, M. Selvi, P. Shagin, E. Shockley, M. Silva, H. Simgen, M. v. Sivers, A. Stein, S. Thapa, D. Thers, A. Tiseni, G. Trinchero, C. Tunnell, M. Vargas, N. Upole, H. Wang, Z. Wang, Y. Wei, C. Weinheimer, J. Wulf, J. Ye, Y. Zhang, and T. Zhu. First Dark Matter Search Results from the XENON1T Experiment,. *ArXiv e-prints*, May 2017.

[20] J. S. Arabadjis, M. W. Bautz, and G. P. Garmire. Chandra Observations of the Lensing Cluster EMSS 1358+6245: Implications for Self-interacting Dark Matter,. *ApJ*, 572:66–78, June 2002. doi: 10.1086/340296.

[21] C. Armendariz-Picon, V. Mukhanov, and P. J. Steinhardt. Dynamical Solution to the Problem of a Small Cosmological Constant and Late-Time Cosmic Acceleration,. *Physical Review Letters*, 85:4438–4441, November 2000. doi: 10.1103/PhysRevLett.85.4438.

[22] R. Barkana. Possible interaction between baryons and dark-matter particles revealed by the first stars. *Nature*, 555:71–74, March 2018. doi: 10.1038/nature25791.

[23] Daniel Bauer, James Buckley, Matthew Cahill-Rowley, Randel Cotta, Alex Drlica-Wagner, Jonathan L. Feng, Stefan Funk, JoAnne Hewett, Dan Hooper, Ahmed Ismail, Manoj Kaplinghat, Alexander Kusenko, Konstantin Matchev, Daniel McKinsey, Tom Rizzo, William Shepherd, Tim M. P. Tait, Alexander M. Wijangco, and Matthew Wood. Dark matter in the coming decade: Complementary paths to discovery and beyond. *Physics of the Dark Universe*, 7:16–23, March 2015. doi: 10.1016/j.dark.2015.04.001.

[24] K. G. Begeman, A. H. Broeils, and R. H. Sanders. Extended rotation curves of spiral galaxies - Dark haloes and modified dynamics,. *MNRAS*, 249:523–537, April 1991.

[25] R. Bernabei, P. Belli, F. Cappella, R. Cerulli, C. J. Dai, A. D'Angelo, H. L. He, A. Incicchitti, H. H. Kuang, J. M. Ma, F. Montecchia, F. Nozzoli, D. Prosperi,

X. D. Sheng, and Z. P. Ye. First results from DAMA/LIBRA and the combined results with DAMA/NaI,. *European Physical Journal C*, 56:333, August 2008. doi: 10.1140/epjc/s10052-008-0662-y.

[26] G. Bertone, D. Hooper, and J. Silk. Particle dark matter: evidence, candidates and constraints.,. *Physics Reports*, 405:279–390, 2004.

[27] G. Bertone, D. Hooper, and J. Silk. Particle dark matter: evidence, candidates and constraints,. *Physics Reports*, 405:279–390, January 2005. doi: 10.1016/j.physrep.2004.08.031.

[28] J. Billard, E. Figueroa-Feliciano, and L. Strigari. Implication of neutrino backgrounds on the reach of next generation dark matter direct detection experiments,. *PhRvD*, 89(2):023524, January 2014. doi: 10.1103/PhysRevD.89.023524.

[29] C. Boehm, J. A. Schewtschenko, R. J. Wilkinson, C. M. Baugh, and S. Pascoli. Using the Milky Way satellites to study interactions between cold dark matter and radiation,. *MNRAS*, 445:L31–L35, November 2014. doi: 10.1093/mnrasl/slu115.

[30] K. Bolejko. Emerging spatial curvature can resolve the tension between high-redshift CMB and low-redshift distance ladder measurements of the Hubble constant. *PhRvD*, 97(10):103529, May 2018. doi: 10.1103/PhysRevD.97.103529.

[31] J. R. Bond and A. S. Szalay. The collisionless damping of density fluctuations in an expanding universe,. *ApJ*, 274:443–468, November 1983. doi: 10.1086/161460.

[32] A. Boyarsky, O. Ruchayskiy, D. Iakubovskyi, and J. Franse. Unidentified Line in X-Ray Spectra of the Andromeda Galaxy and Perseus Galaxy Cluster. *Physical Review Letters*, 113(25):251301, December 2014. doi: 10.1103/PhysRevLett.113.251301.

[33] A. Boyarsky, J. Franse, D. Iakubovskyi, and O. Ruchayskiy. Checking the Dark Matter Origin of a 3.53 keV Line with the Milky Way Center,. *Physical Review Letters*, 115(16):161301, October 2015. doi: 10.1103/PhysRevLett.115.161301.

[34] Thomas Buchert, Martin Kerscher, and Christian Sicka. Back reaction of inhomogeneities on the expansion: The evolution of cosmological parameters. *PhRvD*, 62: 043525, August 2000. doi: 10.1103/PhysRevD.62.043525.

[35] E. Bulbul, M. Markevitch, A. Foster, R. K. Smith, M. Loewenstein, and S. W. Randall. Detection of An Unidentified Emission Line in the Stacked X-ray spectrum of Galaxy Clusters,. *ArXiv e-prints*, February 2014.

[36] J. Carr, C. Balazs, T. Bringmann, T. Buanes, M. Daniel, M. Doro, C. Farnier, M. Fornasa, J. Gaskins, G. Gomez-Vargas, M. Hayashida, K. Kohri, V. Lefranc, A. Morselli, E. Moulin, N. Mirabal, J. Rico, T. Saito, M. Sanchez-Conde, M. Wilkinson, M. Wood, G. Zaharijas, and H. Zechlin. Prospects for Indirect Dark Matter Searches with the Cherenkov Telescope Array (CTA),. In *34th International Cosmic Ray Conference (ICRC2015)*, volume 34 of *International Cosmic Ray Conference*, page 1203, July 2015.

[37] Bradley W. Carrol and Dale A. Ostlie. *Modern Astronomy*. Pearson, 2007.

[38] Sean Carrol. *Spacetime and Geometry: An Introduction to General Relativity*. Cummings, 2003.

[39] CDMS Collaboration, R. Agnese, Z. Ahmed, A. J. Anderson, S. Arrenberg, D. Balakishiyeva, R. Basu Thakur, D. A. Bauer, J. Billard, A. Borgland, D. Brandt, P. L. Brink, T. Bruch, R. Bunker, B. Cabrera, D. O. Caldwell, D. G. Cerdeno, H. Chagani, J. Cooley, B. Cornell, C. H. Crewdson, P. Cushman, M. Daal, F. Dejongh, E. D. C. E Silva, T. Doughty, L. Esteban, S. Fallows, E. Figueroa-Feliciano, J. Filippini, J. Fox, M. Fritts, G. L. Godfrey, S. R. Golwala, J. Hall, R. H. Harris, S. A. Hertel, T. Hofer, D. Holmgren, L. Hsu, M. E. Huber, A. Jastram, O. Kamaev, B. Kara, M. H. Kelsey, A. Kennedy, P. Kim, M. Kiveni, K. Koch, M. Kos, S. W. Leman, B. Loer, E. Lopez Asamar, R. Mahapatra, V. Mandic, C. Martinez, K. A. McCarthy, N. Mirabolfathi, R. A. Moffatt, D. C. Moore, P. Nadeau, R. H. Nelson, K. Page, R. Partridge, M. Pepin, A. Phipps, K. Prasad, M. Pyle, H. Qiu, W. Rau, P. Redl, A. Reisetter, Y. Ricci, T. Saab, B. Sadoulet, J. Sander, K. Schneck, R. W. Schnee, S. Scorza, B. Serfass, B. Shank, D. Speller, K. M. Sundqvist, A. N. Villano, B. Welliver, D. H. Wright, S. Yellin, J. J. Yen, J. Yoo, B. A. Young, and J. Zhang.

Silicon Detector Dark Matter Results from the Final Exposure of CDMS II,. *ArXiv e-prints*, April 2013.

[40] E. Charles, M. Sanchez-Conde, B. Anderson, R. Caputo, A. Cuoco, M. Di Mauro, A. Drlica-Wagner, G. A. Gomez-Vargas, M. Meyer, L. Tibaldo, M. Wood, G. Zaharijas, S. Zimmer, M. Ajello, A. Albert, L. Baldini, K. Bechtol, E. D. Bloom, F. Ceraudo, J. Cohen-Tanugi, S. W. Digel, J. Gaskins, M. Gustafsson, N. Mirabal, and M. Razzano. Sensitivity projections for dark matter searches with the Fermi large area telescope,. *Physics Reports*, 636:1–46, June 2016. doi: 10.1016/j.physrep.2016.05.001.

[41] X. Chen and M. Kamionkowski. Particle decays during the cosmic dark ages,. *PhRvD*, 70(4):043502, August 2004. doi: 10.1103/PhysRevD.70.043502.

[42] T. Chiba, T. Okabe, and M. Yamaguchi. Kinetically driven quintessence,. *PhRvD*, 62(2):023511, July 2000. doi: 10.1103/PhysRevD.62.023511.

[43] D. Clowe, M. Bradac, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky. A Direct Empirical Proof of the Existence of Dark Matter,. *ApJL*, 648:L109–L113, September 2006. doi: 10.1086/508162.

[44] W. A. Dawson, D. Wittman, M. J. Jee, P. Gee, J. P. Hughes, J. A. Tyson, S. Schmidt, P. Thorman, M. Bradac, S. Miyazaki, B. Lemaux, Y. Utsumi, and V. E. Margoniner. Discovery of a Dissociative Galaxy Cluster Merger with Large Physical Separation,. *ApJL*, 747:L42, March 2012. doi: 10.1088/2041-8205/747/2/L42.

[45] A. de Rujula, S. L. Glashow, and U. Sarid. Charged dark matter,. *Nuclear Physics B*, 333:173–194, March 1990. doi: 10.1016/0550-3213(90)90227-5.

[46] DEAP Collaboration, P.-A. Amaudruz, M. Batygov, B. Beltran, J. Bonatt, M. G. Boulay, B. Broerman, J. F. Bueno, A. Butcher, B. Cai, M. Chen, R. Chouinard, B. T. Cleveland, K. Dering, J. DiGioseffo, F. Duncan, T. Flower, R. Ford, P. Giampa, P. Gorel, K. Graham, D. R. Grant, E. Guliyev, A. L. Hallin, M. Hamstra, P. Harvey, C. J. Jillings, M. Kuzniak, I. Lawson, O. Li, P. Liimatainen, P. Majewski, A. B. McDonald, T. McElroy, K. McFarlane, J. Monroe, A. Muir, C. Nantais, C. Ng, A. J. Noble, C. Ouellet, K. Palladino, P. Pasuthip, S. J. M. Peeters, T. Pollmann,

W. Rau, F. Retiere, N. Seeburn, K. Singhrao, P. Skensved, B. Smith, T. Sonley, J. Tang, E. Vazquez-Jauregui, L. Veloce, J. Walding, and M. Ward. DEAP-3600 Dark Matter Search,. *ArXiv e-prints*, October 2014.

[47] E. Di Valentino, C. Boehm, E. Hivon, and F. R. Bouchet. Reducing the H_0 and σ_8 tensions with dark matter-neutrino interactions. *PhRvD*, 97(4):043513, February 2018. doi: 10.1103/PhysRevD.97.043513.

[48] S. Dimopoulos, D. Eichler, R. Esmailzadeh, and G. D. Starkman. Getting a charge out of dark matter,. *PhRvD*, 41:2388–2397, April 1990. doi: 10.1103/PhysRevD.41.2388.

[49] Scott Dodelson. *Modern cosmology*. Academic Press, 2003.

[50] D. J. Eisenstein, I. Zehavi, D. W. Hogg, R. Scoccimarro, M. R. Blanton, R. C. Nichol, R. Scranton, H.-J. Seo, M. Tegmark, Z. Zheng, S. F. Anderson, J. Annis, N. Bahcall, J. Brinkmann, S. Burles, F. J. Castander, A. Connolly, I. Csabai, M. Doi, M. Fukugita, J. A. Frieman, K. Glazebrook, J. E. Gunn, J. S. Hendry, G. Hennessy, Z. Ivezic, S. Kent, G. R. Knapp, H. Lin, Y.-S. Loh, R. H. Lupton, B. Margon, T. A. McKay, A. Meiksin, J. A. Munn, A. Pope, M. W. Richmond, D. Schlegel, D. P. Schneider, K. Shimasaku, C. Stoughton, M. A. Strauss, M. SubbaRao, A. S. Szalay, I. Szapudi, D. L. Tucker, B. Yanny, and D. G. York. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies,. *ApJ*, 633:560–574, November 2005. doi: 10.1086/466512.

[51] John Ellis, J.S. Hagelin, D.V. Nanopoulos, K. Olive, and M. Srednicki. Supersymmetric relics from the big bang. *Nuclear Physics B*, 238(2):453 – 476, 1984. ISSN 0550-3213. doi: https://doi.org/10.1016/0550-3213(84)90461-9. URL http://www.sciencedirect.com/science/article/pii/0550321384904619.

[52] A. Enea Romano. Hubble trouble or Hubble bubble? *ArXiv e-prints*, September 2016.

[53] M. Fairbairn, P. Scott, and J. Edsjo. The zero age main sequence of WIMP burners,. *PhRvD*, 77(4):047301, February 2008. doi: 10.1103/PhysRevD.77.047301.

[54] Y. Fujii. Origin of the gravitational constant and particle masses in a scale-invariant scalar-tensor theory,. *PhRvD*, 26:2580–2588, November 1982. doi: 10.1103/PhysRevD.26.2580.

[55] A. Geringer-Sameth and S. M. Koushiappas. Exclusion of Canonical Weakly Interacting Massive Particles by Joint Analysis of Milky Way Dwarf Galaxies with Data from the Fermi Gamma-Ray Space Telescope,. *Physical Review Letters*, 107(24): 241303, December 2011. doi: 10.1103/PhysRevLett.107.241303.

[56] O. Y. Gnedin and J. P. Ostriker. Limits on Collisional Dark Matter from Elliptical Galaxies in Clusters,. *ApJ*, 561:61–68, November 2001. doi: 10.1086/323211.

[57] G. A. Gomez-Vargas, M. A. Sanchez-Conde, J.-H. Huh, M. Peiro, F. Prada, A. Morselli, A. Klypin, D. G. Cerdeno, Y. Mambrini, and C. Munoz. Constraints on WIMP annihilation for contracted dark matter in the inner Galaxy with the Fermi-LAT,. *JCAP*, 10:029, October 2013. doi: 10.1088/1475-7516/2013/10/029.

[58] L. Gu, J. Kaastra, A. J. J. Raassen, P. D. Mullen, R. S. Cumbee, D. Lyons, and P. C. Stancil. A novel scenario for the possible X-ray line feature at ˜3.5 keV. Charge exchange with bare sulfur ions. *A&A*, 584:L11, December 2015. doi: 10.1051/0004-6361/201527634.

[59] D. Harvey, R. Massey, T. Kitching, A. Taylor, and E. Tittley. The nongravitational interactions of dark matter in colliding galaxy clusters,. *Science*, 347:1462–1465, March 2015. doi: 10.1126/science.1261381.

[60] J. F. Hennawi and J. P. Ostriker. Observational Constraints on the Self-interacting Dark Matter Scenario and the Growth of Supermassive Black Holes,. *ApJ*, 572: 41–54, June 2002. doi: 10.1086/340226.

[61] H. Hoekstra, H. K. C. Yee, and M. D. Gladders. Current status of weak gravitational lensing,. *NewAR*, 46:767–781, November 2002. doi: 10.1016/S1387-6473(02)00245-2.

[62] J. Holder, V. A. Acciari, E. Aliu, T. Arlen, M. Beilicke, W. Benbow, S. M. Bradbury, J. H. Buckley, V. Bugaev, Y. Butt, K. L. Byrum, A. Cannon, O. Celik, A. Cesarini,

L. Ciupik, Y. C. K. Chow, P. Cogan, P. Colin, W. Cui, M. K. Daniel, T. Ergin, A. D. Falcone, S. J. Fegan, J. P. Finley, G. Finnegan, P. Fortin, L. F. Fortson, A. Furniss, G. H. Gillanders, J. Grube, R. Guenette, G. Gyuk, D. Hanna, E. Hays, D. Horan, C. M. Hui, T. B. Humensky, A. Imran, P. Kaaret, N. Karlsson, M. Kertzman, D. B. Kieda, J. Kildea, A. Konopelko, H. Krawczynski, F. Krennrich, M. J. Lang, S. Lebohec, G. Maier, A. McCann, M. McCutcheon, P. Moriarty, R. Mukherjee, T. Nagai, J. Niemiec, R. A. Ong, D. Pandel, J. S. Perkins, M. Pohl, J. Quinn, K. Ragan, L. C. Reyes, P. T. Reynolds, H. J. Rose, M. Schroedter, G. H. Sembroski, A. W. Smith, D. Steele, S. P. Swordy, J. A. Toner, L. Valcarcel, V. V. Vassiliev, R. Wagner, S. P. Wakely, J. E. Ward, T. C. Weekes, A. Weinstein, R. J. White, D. A. Williams, S. A. Wissel, M. Wood, and B. Zitzer. Status of the VERITAS Observatory,. In F. A. Aharonian, W. Hofmann, and F. Rieger, editors, *American Institute of Physics Conference Series*, volume 1085 of *American Institute of Physics Conference Series*, pages 657–660, December 2008. doi: 10.1063/1.3076760.

[63] F. Iocco, M. Taoso, F. Leclercq, and G. Meynet. Main Sequence Stars with Asymmetric Dark Matter. *Physical Review Letters*, 108(6):061301, February 2012. doi: 10.1103/PhysRevLett.108.061301.

[64] T. Jeltema and S. Profumo. Discovery of a 3.5 keV line in the Galactic Centre and a critical look at the origin of the line across astronomical targets. *MNRAS*, 450: 2143–2152, June 2015. doi: 10.1093/mnras/stv768.

[65] A. Kamenshchik, U. Moschella, and V. Pasquier. An alternative to quintessence,. *Physics Letters B*, 511:265–268, July 2001. doi: 10.1016/S0370-2693(01)00571-8.

[66] Edward W. Kolb and Richard Slansky. Dimensional reduction in the early universe: Where have the massive particles gone? *Physics Letters B*, 135(5): 378 – 382, 1984. ISSN 0370-2693. doi: https://doi.org/10.1016/0370-2693(84)90298-3. URL http://www.sciencedirect.com/science/article/pii/0370269384902983.

[67] L. V. E. Koopmans and T. Treu. The Structure and Dynamics of Luminous and Dark Matter in the Early-Type Lens Galaxy of 0047-281 at z = 0.485,. *ApJ*, 583: 606–615, February 2003. doi: 10.1086/345423.

[68] M. Kuzniak, M. G. Boulay, and T. Pollmann. Surface roughness interpretation of 730 kg days CRESST-II results,. *Astroparticle Physics*, 36:77–82, August 2012. doi: 10.1016/j.astropartphys.2012.05.005.

[69] MAGIC Collaboration. Limits to dark matter annihilation cross-section from a combined analysis of MAGIC and Fermi-LAT observations of dwarf satellite galaxies,. *JCAP*, 2:039, February 2016. doi: 10.1088/1475-7516/2016/02/039.

[70] M. Markevitch, A. H. Gonzalez, D. Clowe, A. Vikhlinin, W. Forman, C. Jones, S. Murray, and W. Tucker. Direct Constraints on the Dark Matter Self-Interaction Cross Section from the Merging Galaxy Cluster 1E 0657-56,. *ApJ*, 606:819–824, May 2004. doi: 10.1086/383178.

[71] P. Natarajan, A. Loeb, J.-P. Kneib, and I. Smail. Constraints on the Collisional Nature of the Dark Matter from Gravitational Lensing in the Cluster A2218,. *ApJL*, 580:L17–L20, November 2002. doi: 10.1086/345547.

[72] M. S. Owers, S. W. Randall, P. E. J. Nulsen, W. J. Couch, L. P. David, and J. C. Kempner. The Dissection of Abell 2744: A Rich Cluster Growing Through Major and Minor Mergers,. *ApJ*, 728:27, February 2011. doi: 10.1088/0004-637X/728/1/27.

[73] N. Padmanabhan and D. P. Finkbeiner. Detecting dark matter annihilation with CMB polarization: Signatures and experimental prospects,. *PhRvD*, 72(2):023508, July 2005. doi: 10.1103/PhysRevD.72.023508.

[74] PandaX-II Collaboration, :, A. Tan, M. Xiao, X. Cui, X. Chen, Y. Chen, D. Fang, C. Fu, K. Giboni, F. Giuliani, H. Gong, S. Hu, X. Huang, X. Ji, Y. Ju, S. Lei, S. Li, X. Li, X. Li, H. Liang, Q. Lin, H. Liu, J. Liu, W. Lorenzon, Y. Ma, Y. Mao, K. Ni, X. Ren, M. Schubnell, M. Shen, F. Shi, H. Wang, J. Wang, M. Wang, Q. Wang, S. Wang, X. Wang, Z. Wang, S. Wu, X. Xiao, P. Xie, B. Yan, Y. Yang, J. Yue, X. Zeng, H. Zhang, H. Zhang, H. Zhang, T. Zhang, L. Zhao, J. Zhou, N. Zhou, and X. Zhou. Dark Matter Results from First 98.7-day Data of PandaX-II Experiment,. *ArXiv e-prints*, July 2016.

[75] P. J. E. Peebles. The origin of galaxies and clusters of galaxies,. *Science*, 224: 1385–1391, June 1984. doi: 10.1126/science.224.4656.1385.

[76] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, N. J. Nunes, R. Pain, C. R. Pennypacker, R. Quimby, C. Lidman, R. S. Ellis, M. Irwin, R. G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaefer, B. J. Boyle, A. V. Filippenko, T. Matheson, A. S. Fruchter, N. Panagia, H. J. M. Newberg, W. J. Couch, and Supernova Cosmology Project. Measurements of Omega and Lambda from 42 High-Redshift Supernovae,. *ApJ*, 517:565–586, June 1999. doi: 10.1086/307221.

[77] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2015 results. XIII. Cosmological parameters. *A&A*, 594:A13, September 2016. doi: 10.1051/0004-6361/201525830.

[78] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2015 results. XIII. Cosmological parameters,. *aap*, 594:A13, September 2016. doi: 10.1051/0004-6361/201525830.

[79] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2015 results. XIII. Cosmological parameters,. *A&A*, 594:A13, September 2016. doi: 10.1051/0004-6361/201525830.

[80] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J. P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, R. C. Butler, E. Calabrese, J. F. Cardoso, J. Carron, A. Challinor, H. C. Chiang, J. Chluba, L. P. L. Colombo, C. Combet, D. Contreras, B. P. Crill, F. Cuttaia, P. de Bernardis, G. de Zotti, J. Delabrouille, J. M. Delouis, E. Di Valentino, J. M. Diego, O. Doré, M. Douspis, A. Ducout, X. Dupac, S. Dusini, G. Efstathiou, F. Elsner, T. A.

Enßlin, H. K. Eriksen, Y. Fantaye, M. Farhang, J. Fergusson, R. Fernandez-Cobos, F. Finelli, F. Forastieri, M. Frailis, E. Franceschi, A. Frolov, S. Galeotta, S. Galli, K. Ganga, R. T. Génova-Santos, M. Gerbino, T. Ghosh, J. González-Nuevo, K-M. Górski, S. Gratton, A. Gruppuso, J. E. Gudmundsson, J. Hamann, W. Handley, D. Herranz, E. Hivon, Z. Huang, A. H. Jaffe, W. C. Jones, A. Karakci, E. Keihänen, R. Keskitalo, K. Kiiveri, J. Kim, T. S. Kisner, L. Knox, N. Krachmalnicoff, M. Kunz, H. Kurki-Suonio, G. Lagache, J. M. Lamarre, A. Lasenby, M. Lattanzi, C. R. Lawrence, M. Le Jeune, P. Lemos, J. Lesgourgues, F. Levrier, A. Lewis, M. Liguori, P. B. Lilje, M. Lilley, V. Lindholm, M. López-Caniego, P. M. Lubin, Y. Z. Ma, J. F. Macías-Pérez, G. Maggio, D. Maino, N. Mandolesi, A. Mangilli, A. Marcos-Caballero, M. Maris, P. G. Martin, M. Martinelli, E. Martínez-González, S. Matarrese, N. Mauri, J. D. McEwen, P. R. Meinhold, A. Melchiorri, A. Mennella, M. Migliaccio, M. Millea, S. Mitra, M. A Miville-Deschênes, D. Molinari, L. Montier, G. Morgante, A. Moss, P. Natoli, H. U. Nørgaard-Nielsen, L. Pagano, D. Paoletti, B. Partridge, G. Patanchon, H. V. Peiris, F. Perrotta, V. Pettorino, F. Piacentini, L. Polastri, G. Polenta, J. L. Puget, J. P. Rachen, M. Reinecke, M. Remazeilles, A. Renzi, G. Rocha, C. Rosset, G. Roudier, J. A. Rubiño-Martín, B. Ruiz-Granados, L. Salvati, M. Sandri, M. Savelainen, D. Scott, E. P. S. Shellard, C. Sirignano, G. Sirri, L. D. Spencer, R. Sunyaev, A. S. Suur-Uski, J. A. Tauber, D. Tavagnacco, M. Tenti, L. Toffolatti, M. Tomasi, T. Trombetti, L. Valenziano, J. Valiviita, B. Van Tent, L. Vibert, P. Vielva, F. Villa, N. Vittorio, B. D. Wandelt, I. K. Wehus, M. White, S. D. M. White, A. Zacchei, and A. Zonca. Planck 2018 results. VI. Cosmological parameters. *arXiv e-prints*, art. arXiv:1807.06209, July 2018.

[81] B. Ratra and P. J. E. Peebles. Cosmological consequences of a rolling homogeneous scalar field,. *PhRvD*, 37:3406–3427, June 1988. doi: 10.1103/PhysRevD.37.3406.

[82] A. G. Riess, L. M. Macri, S. L. Hoffmann, D. Scolnic, S. Casertano, A. V. Filippenko, B. E. Tucker, M. J. Reid, D. O. Jones, J. M. Silverman, R. Chornock, P. Challis, W. Yuan, P. J. Brown, and R. J. Foley. A 2.4% Determination of the Local Value of the Hubble Constant. *ArXiv e-prints*, April 2016.

[83] A. G. Riess, L. M. Macri, S. L. Hoffmann, D. Scolnic, S. Casertano, A. V. Fil-

ippenko, B. E. Tucker, M. J. Reid, D. O. Jones, J. M. Silverman, R. Chornock, P. Challis, W. Yuan, P. J. Brown, and R. J. Foley. A 2.4% Determination of the Local Value of the Hubble Constant,. *ApJ*, 826:56, July 2016. doi: 10.3847/0004-637X/826/1/56.

[84] E. Ripamonti, M. Mapelli, and A. Ferrara. The impact of dark matter decays and annihilations on the formation of the first structures,. *MNRAS*, 375:1399–1408, March 2007. doi: 10.1111/j.1365-2966.2006.11402.x.

[85] L. J. Rosenberg and K. A. van Bibber. Searches for invisible axions.,. *Physics Reports*, 325:1–39, February 2000. doi: 10.1016/S0370-1573(99)00045-9.

[86] L. Roszkowski, E. M. Sessolo, and S. Trojanowski. WIMP dark matter candidates and searches - current issues and future prospects,. *ArXiv e-prints*, July 2017.

[87] L. Roszkowski, E. M. Sessolo, and S. Trojanowski. WIMP dark matter candidates and searches - current issues and future prospects,. *ArXiv e-prints*, July 2017.

[88] P. Salati and J. Silk. A stellar probe of dark matter annihilation in galactic nuclei,. *ApJ*, 338:24–31, March 1989. doi: 10.1086/167177.

[89] C. Savage, G. Gelmini, P. Gondolo, and K. Freese. Compatibility of DAMA/LIBRA dark matter detection with other searches,. *JCAP*, 4:010, April 2009. doi: 10.1088/1475-7516/2009/04/010.

[90] J. A. Schewtschenko, R. J. Wilkinson, C. M. Baugh, C. Bœhm, and S. Pascoli. Dark matter-radiation interactions: the impact on dark matter haloes. *MNRAS*, 449: 3587–3596, June 2015. doi: 10.1093/mnras/stv431.

[91] J. A. Schewtschenko, C. M. Baugh, R. J. Wilkinson, C. Boehm, S. Pascoli, and T. Sawala. Dark matter-radiation interactions: the structure of Milky Way satellite galaxies,. *MNRAS*, 461:2282–2287, September 2016. doi: 10.1093/mnras/stw1078.

[92] Peter Schneider. *Extragalactic Astronomy and Cosmology*. Springer-Verlag Berlin Heidelberg, 2015.

[93] P. C. Scott, J. Edsjo, and M. Fairbairn. Low Mass Stellar Evolution with WIMP Capture and Annihilation,. In H. V. Klapdor-Kleingrothaus and G. F. Lewis, editors, *Dark Matter in Astroparticle and Particle Physics*, pages 387–392, April 2008. doi: 10.1142/9789812814357_0038.

[94] C. Shah, S. Dobrodey, S. Bernitt, R. Steinbrügge, J. R. Crespo López-Urrutia, L. Gu, and J. Kaastra. Laboratory Measurements Compellingly Support a Charge-exchange Mechanism for the 'Dark Matter' ~ 3.5 keV X-Ray Line. *ApJ*, 833:52, December 2016. doi: 10.3847/1538-4357/833/1/52.

[95] D. Spolyar, K. Freese, and P. Gondolo. Dark Matter and the First Stars: A New Phase of Stellar Evolution,. *Physical Review Letters*, 100(5):051101, February 2008. doi: 10.1103/PhysRevLett.100.051101.

[96] V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, J. A. Peacock, S. Cole, P. Thomas, H. Couchman, A. Evrard, J. Colberg, and F. Pearce. Simulations of the formation, evolution and clustering of galaxies and quasars,. *Nature*, 435:629–636, June 2005. doi: 10.1038/nature03597.

[97] J. Stadler and C. Boehm. Constraints on γ-CDM interactions matching the Planck data precision. *ArXiv e-prints*, February 2018.

[98] G. D. Starkman, A. Gould, R. Esmailzadeh, and S. Dimopoulos. Opening the window on strongly interacting dark matter,. *PhRvD*, 41:3594–3603, June 1990. doi: 10.1103/PhysRevD.41.3594.

[99] G. Steigman, B. Dasgupta, and J. F. Beacom. Precise relic WIMP abundance and its impact on searches for dark matter annihilation,. *PhRvD*, 86(2):023506, July 2012. doi: 10.1103/PhysRevD.86.023506.

[100] R. Strauss, G. Angloher, A. Bento, C. Bucci, L. Canonica, X. Defay, A. Erb, F. von Feilitzsch, N. Ferreiro Iachellini, P. Gorla, A. Gutlein, D. Hauff, J. Jochum, M. Kiefer, H. Kluck, H. Kraus, J. C. Lanfranchi, J. Loebell, A. Munster, C. Pagliarone, F. Petricca, W. Potzel, F. Probst, F. Reindl, K. Schaffner, J. Schieck, S. Schonert, W. Seidel, L. Stodolsky, C. Strandhagen, A. Tanzke, H. H. Trinh Thi,

C. Turkoglu, M. Ufflnger, A. Ulrich, I. Usherov, S. Wawoczny, M. Willers, M. Wustrich, and A. Zoller. The CRESST-III low-mass WIMP detector,. In *Journal of Physics Conference Series*, volume 718 of *Journal of Physics Conference Series*, page 042048, May 2016. doi: 10.1088/1742-6596/718/4/042048.

[101] SuperCDMS Collaboration, R. Agnese, A. J. Anderson, T. Aramaki, M. Asai, W. Baker, D. Balakishiyeva, D. Barker, R. Basu Thakur, D. A. Bauer, J. Billard, A. Borgland, M. A. Bowles, P. L. Brink, R. Bunker, B. Cabrera, D. O. Caldwell, R. Calkins, D. G. Cerdeno, H. Chagani, Y. Chen, J. Cooley, B. Cornell, P. Cushman, M. Daal, P. C. F. Di Stefano, T. Doughty, L. Esteban, S. Fallows, E. Figueroa-Feliciano, M. Ghaith, G. L. Godfrey, S. R. Golwala, J. Hall, H. R. Harris, T. Hofer, D. Holmgren, L. Hsu, M. E. Huber, D. Jardin, A. Jastram, O. Kamaev, B. Kara, M. H. Kelsey, A. Kennedy, A. Leder, B. Loer, E. Lopez Asamar, P. Lukens, R. Mahapatra, V. Mandic, N. Mast, N. Mirabolfathi, R. A. Moffatt, J. D. Morales Mendoza, S. M. Oser, K. Page, W. A. Page, R. Partridge, M. Pepin, A. Phipps, K. Prasad, M. Pyle, H. Qiu, W. Rau, P. Redl, A. Reisetter, Y. Ricci, A. Roberts, H. E. Rogers, T. Saab, B. Sadoulet, J. Sander, K. Schneck, R. W. Schnee, S. Scorza, B. Serfass, B. Shank, D. Speller, D. Toback, R. Underwood, S. Upadhyayula, A. N. Villano, B. Welliver, J. S. Wilson, D. H. Wright, S. Yellin, J. J. Yen, B. A. Young, and J. Zhang. WIMP-Search Results from the Second CDMSlite Run,. *ArXiv e-prints*, September 2015.

[102] M. Szydagis, LUX Collaboration, and LZ Collaboration. The Present and Future of Searching for Dark Matter with LUX and LZ,. In *Proceedings of the 38th International Conference on High Energy Physics (ICHEP2016). 3-10 August 2016. Chicago, USA. Online at ¡A href=“http://pos.sissa.it/cgi-bin/reader/conf.cgi?confid=282”¿http://pos.sissa.it/cgi-bin/read er/-conf.cgi?confid=282¡/A¿, id.220*, page 220, 2016.

[103] M. Taoso, G. Bertone, and A. Masiero. Dark matter candidates: a ten-point test,. *JCAP*, 3:022, March 2008. doi: 10.1088/1475-7516/2008/03/022.

[104] The Fermi-LAT Collaboration, :, M. Ackermann, A. Albert, B. Anderson, L. Baldini, J. Ballet, G. Barbiellini, D. Bastieri, K. Bechtol, R. Bellazzini, E. Bissaldi, E. D. Bloom, E. Bonamente, A. Bouvier, T. J. Brandt, J. Bregeon, M. Brigida,

P. Bruel, R. Buehler, S. Buson, G. A. Caliandro, R. A. Cameron, M. Caragiulo, P. A. Caraveo, C. Cecchi, E. Charles, A. Chekhtman, J. Chiang, S. Ciprini, R. Claus, J. Cohen-Tanugi, J. Conrad, F. D'Ammando, A. de Angelis, C. D. Dermer, S. W. Digel, E. d. C. e Silva, P. S. Drell, A. Drlica-Wagner, R. Essig, C. Favuzzi, E. C. Ferrara, A. Franckowiak, Y. Fukazawa, S. Funk, P. Fusco, F. Gargano, D. Gasparrini, N. Giglietto, M. Giroletti, G. Godfrey, G. A. Gomez-Vargas, I. A. Grenier, S. Guiriec, M. Gustafsson, M. Hayashida, E. Hays, J. Hewitt, R. E. Hughes, T. Jogler, T. Kamae, J. Knodlseder, D. Kocevski, M. Kuss, . Larsson, L. Latronico, M. Llena Garde, F. Longo, F. Loparco, M. N. Lovellette, P. Lubrano, G. Martinez, M. Mayer, M. N. Mazziotta, P. F. Michelson, W. Mitthumsiri, T. Mizuno, A. A. Moiseev, M. E. Monzani, A. Morselli, I. V. Moskalenko, S. Murgia, R. Nemmen, E. Nuss, T. Ohsugi, E. Orlando, J. F. Ormes, J. S. Perkins, F. Piron, G. Pivato, T. A. Porter, S. Raino, R. Rando, M. Razzano, S. Razzaque, A. Reimer, O. Reimer, S. Ritz, M. Sanchez-Conde, N. Sehgal, C. Sgro, E. J. Siskind, P. Spinelli, L. Strigari, D. J. Suson, H. Tajima, H. Takahashi, J. B. Thayer, L. Tibaldo, M. Tinivella, D. F. Torres, Y. Uchiyama, T. L. Usher, J. Vandenbroucke, G. Vianello, V. Vitale, M. Werner, B. L. Winer, K. S. Wood, M. Wood, G. Zaharijas, and S. Zimmer. Dark Matter Constraints from Observations of 25 Milky Way Satellite Galaxies with the Fermi Large Area Telescope. *ArXiv e-prints*, October 2013.

[105] O. Urban, N. Werner, S. W. Allen, A. Simionescu, J. S. Kaastra, and L. E. Strigari. A Suzaku search for dark matter emission lines in the X-ray brightest galaxy clusters,. *MNRAS*, 451:2447–2461, August 2015. doi: 10.1093/mnras/stv1142.

[106] S. Weinberg. The cosmological constant problem,. *Reviews of Modern Physics*, 61: 1–23, January 1989. doi: 10.1103/RevModPhys.61.1.

[107] R. J. Wilkinson, J. Lesgourgues, and C. Boehm. Using the CMB angular power spectrum to study Dark Matter-photon interactions,. *JCAP*, 4:026, April 2014. doi: 10.1088/1475-7516/2014/04/026.

[108] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln,. *Helvetica Physica Acta*, 6:110–127, 1933.

[109] F. Zwicky. On the Masses of Nebulae and of Clusters of Nebulae. *ApJ*, 86:217, October 1937. doi: 10.1086/143864.

Chapter 2

A brief introduction to Machine Learning Techniques.

Machine Learning techniques are computer algorithms that learn to perform certain tasks (prediction, classification, etc.) through training with known examples. Formally, *Tom Mitchell* [3] defines that a computer program learns to perform a task T, based on a training E and taking into account a measure P of its performance, if this measure P when making T improves with the training E .

While this definition may seem complex at first, the main idea behind these techniques is very natural and will be clear studying the following example.

Ejemplo: Learning to classify astronomical objects.

In this example we will try to classify astronomical objects into stars, variable stars and quasars, using only photometric information. Specifically, we will use their colours $u-g$, $g-r$, $r-i$ and $i-z$. Where u, g, r, i and z are the SDSS photometric bands. Using the aforementioned data set, in which we know both the optical emission of each object and its actual classification, we will train a machine learning algorithm known as *Random Forest* [a]. Then we will measure its performance taking into account the number of objects in which the algorithm have a correct prediction and the amount of objects with a wrong prediction. The result of the training can be seen in figure 2.1, where it is shown that as we increase the number of examples in the training set (i.e we increase the training E), it also increase the performance (P) the method. This can be understood taking into account that as we increase the number of examples in the

training set, the algorithm can learn better the main characteristics of each class. It is worth to mention that this learning process will never be perfect, that is to say that there is always the possibility that our algorithm fails to predict the correct class of a new object. Nowadays most of the work while dealing with machine learning techniques, is dedicated to decrease the errors at predicting new observations.

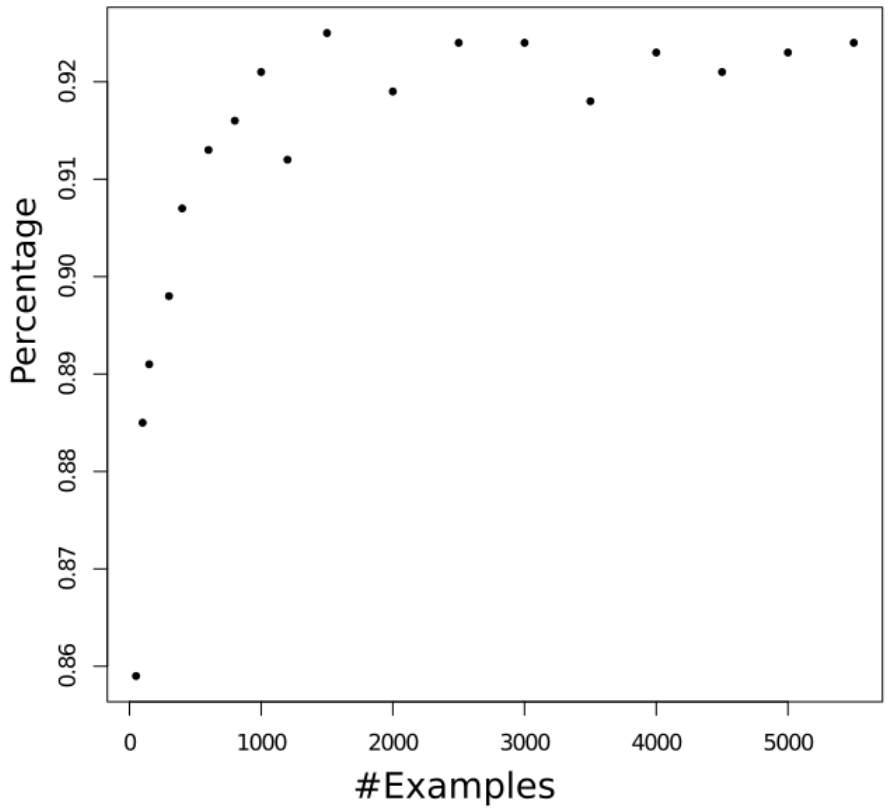

Figure 2.1: Percentage of objects well classified as a function of the number of examples present in the training set

[a] The details of this algorithm will be introduced in the section 2.1.1.

Machine Learning techniques can be classified into 2 main categories: Supervised Learning and Unsupervised Learning. It is important to clarify that there are other categories of machine learning methods, such as Reinforcement Learning, that will not be discussed in this thesis.

Broadly speaking, the goal of supervised learning is to establish a correspondence between input data and the desired output data, using known instances as a training set. An example of this type of techniques are the machine learning algorithms that tries to classify a series of observations among several categories. This problem is known as a classification problem.

On the other hand, the main goal of unsupervised learning is to find structures or patterns in the data set, analysing it without having any a priori information.

2.1 Supervised Learning.

As stated previously, the main objective of supervised learning is to learn a correlation or correspondence between a set of input variables x (also known as features) and an output set y analysing a known data set $D = (x_i, y_i)_{i=1}^{N}$. Where D is the training set and N is the number of examples in this set that are going to be studied. In the simplest scheme, each example x_i is a m-dimensional vector of numbers, however these examples can be much more complex structures such as images, text, time series, graphs, etc. Likewise, each element y_i of the output set can be a complex structure, however in most of the problems, the output set is a finite categorical variable (classes) or a real number. When y_i is a categorical variable, the problem is known as classification, whereas when $y_i \in \Re^n$ the problem is called regression [4]. This type of problem can be formalised if we assume $y = f(x)$ for some unknown function f, then the objective of our algorithm is to estimate this function from a set of examples and then, to make predictions over new observations.

2.1.1 Most important supervised learning models.

- *K-nearest neighbours* (KNN):

  One of the simplest supervised learning algorithms is the *k-nearest neighbour*. This algorithm estimates the classification of a new object taking into account the classes of the k objects, belonging to the training set, that are closer to the object to be classified. For this it is necessary to define a metric in the feature space. Formally we can say that, in a regression problem, the prediction y of a new object with *features* x is:

$$y(x) = \frac{1}{k} \sum_{x_i \in N_{k,x}} y_i \tag{2.1}$$

where $N_{k,x}$ is the subset of the training set consisting of the k observations closer to the object we want to classify, and the values y_i are the outputs of those objects. On the other hand, in a classification problem, the result will be the predominant class in the k examples closest to the object under study.

Is important to remark that, in addition to having to define a suitable metric to the problem that we want to study, we must choose the number of neighbours k that we are going to take into account when classifying a new observation. If we choose a large k, we could lose the local characteristic of our classification but, on the other hand, if we choose a very small k we will over-fit the training set and start modelling the noise instead of the original signal (See section 2.1.2).

In figure 2.2 (left) we can see an example of a KNN algorithm that was trained with $k = 1$, while in the right we can see an example one trained with $k = 15$ in the same parameter space and with the same training set. It can be seen that when we work with $k = 1$ the algorithm tends to over-fit the input data.

In section 2.1.2, we will discuss techniques to choose the best number of neighbours and so, to avoid over-fitting the training data.

- Logistic Regression:

The Logistic regression algorithm consists on modelling the posterior probability $P(k|x)$ of each class k through linear functions in each input parameter x_i. It is worth to mention that each probability should be in the interval $[0, 1]$ and the sum of all the probabilities must be 1. With this in mind, we can write our model as:

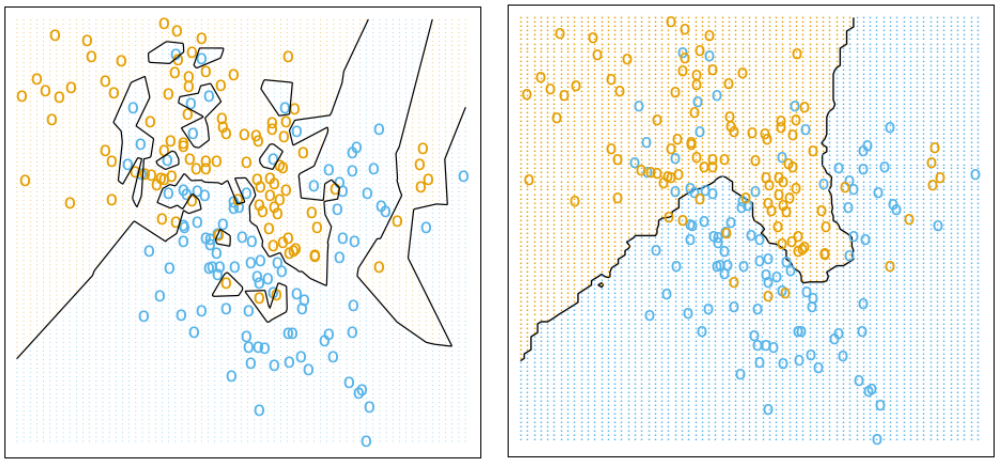

Figure 2.2: *k-nearest neighbour* algorithm. (Left). KNN trained with $k = 1$. (Right.) KNN trained with $k = 15$. Figure extracted from *Hastie et al. 2001* [2].

$$
\begin{aligned}
log\frac{P(k=1|X=x_i)}{P(k=K|X=x_i)} &= \beta_{1,0} + \sum_{j=1}^{n} \beta_{1,j}x^j \\
log\frac{P(k=2|X=x_i)}{P(k=K|X=x_i)} &= \beta_{2,0} + \sum_{j=1}^{n} \beta_{2,j}x^j \\
&\vdots \\
log\frac{P(k=K-1|X=x_i)}{P(k=K|X=x_i)} &= \beta_{K-1,0} + \sum_{j=1}^{n} \beta_{K-1,j}x^j \\
P(k=K|X=x_i) &= \frac{1}{1+\sum_{l=1}^{K-1} exp(\beta_{l,0} + \sum_{j=1}^{n} \beta_{l,j}x^j)}
\end{aligned}
\tag{2.2}
$$

where the last line of the equation 2.2 ensures that the probability that a given object belongs to some class is equal to 1. It is worth clarifying that although in this case we use as a normalising factor the posterior probability that the object belongs to the class $k = K$, you can choose the probability of belonging to any class without affecting the final result.

Finally to uniquely define the model, it is necessary to estimate the values of $\vec{\beta} \equiv \beta_{i,j}$. In order to do this you can use different methods taking into account the number

of classes that we want to study and the dimensionality of our parameter space. One of the most used techniques to adjust the parameters $\vec{\beta}$ taking N observations is maximising the likelihood defined by:

$$l(\vec{\beta}) = \sum_{i=1}^{N} log(p(k = k_i | X = x_i; \vec{\beta})) \tag{2.3}$$

In a similar manner, taking into account the dimensionality of the problem, you can use different methods (Descendent gradient method, Newton's method, etc.) to find the values that maximise the likelihood.

- *Support Vector Machine*: A *Support Vector Machine* (SVM) is a supervised learning algorithm originally created for problems of binary classification. Given a training set $\{(x_1, y_1), ..., (x_n, y_n)\}$ with $y_i = \pm 1$ in a feature space of dimension d, this method look for hyper-planes, i.e. hyper-surfaces of dimension $d - 1$, that separate the classes. Mathematically, we can define an hyper-plane as

$$x : f(x) = x^T \beta + \beta_0 = 0 \tag{2.4}$$

where β is a unit vector, x are the features and x^T is the transposed vector. Then we can define a classification rule taking into account $sgn(f(x) = x^T \beta + \beta_0)$ (it can be shown that $f(x)$ is geometrically the distance from the point x to the hyperplane). Assuming that the classes are perfectly separable (See figure 2.3 (a) and (b)) we can find a β such that $y_i f(x_i) > 0 \forall i$. To deal with problems where classes are not perfectly separable we can introduce cost variables ξ_i that quantify the error in the classification (see figure 2.3 (c)), and then impose a maximum on the sum of the costs $\sum_i \xi_i < C$. This technique is known as *soft margin* and allows to SVM methods to avoid overfitting the training set, and so, to have a better generalisation.

  As can be seen in figure 2.3 (a), there are cases in which there are several hyper-planes that separate the classes and that can be used as a reference for the classification. It is common to keep the hyper-plane that maximises the separation or margin between classes (See figure 2.3 (b)). Mathematically this translates into:

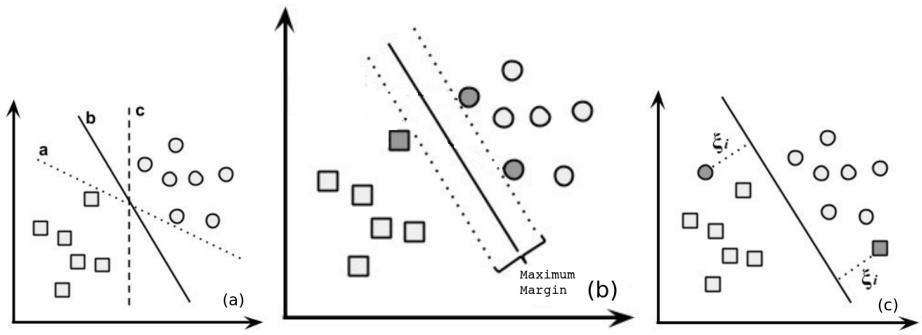

Figure 2.3: Classification schema using *Support Vector Machine*.

$$\max_{\beta,\beta_0} M \tag{2.5}$$

$$\begin{aligned} y_i(x_i^T\beta + \beta_0) &\geq M(1-\xi_i) \\ \sum_i \xi_i &\leq C \end{aligned} \tag{2.6}$$

where M is the margin between classes. It can be shown that the hyperplane will depend only on a subset of observations (called support vectors) that geometrically fall inside or in the margin that separates both classes.

It is important to note that this algorithm only looks for linear hyper-planes in the feature space. A way to generalise this method to more complex hyper-planes is to enlarge the feature space through a set of basis functions $h(x)$. Then the hyper-planes will be searched in a new feature space defined by $h(x_i) = (h_1(x_i), h_2(x_i), ...)$, that translates on non-linear surfaces in the original feature space. Then the classification rule will be defined by $Sgn(f(x) = h(x)^T\beta + \beta_0)$.

As was previously stated, the SVM methods were created to solve binary classification problems, however there are generalisations to solve multiple classification problems and regression problems.

Assuming we want to performed a classification between K different classes, there are 2 types of generalisations. The first option is to generate a binary classifier $f_k(x)$ for each class. Then each classifier f_k will treat as a new class those observations that belong to the k class and as another new class those observations that belong

to all the other original classes. Finally we will have K classifications for each observation, which can lead to ambiguities in the final classification. Another problem that can appear in this generalisation is the imbalance between the new classes. This means that it can be the case that for a given class k there are few observations that belong to that class compared to the number of observations that do not belong to that class.

The second option to generalise the SVM techniques for multiple classification is to generate a classifier $f_{k,k'}$ that studies the limits between each pair of classes. Then we will have $K(K-1)/2$ classifiers, and so, we may have ambiguities when classifying a new observation.

- Decision Trees:

  Decision trees are supervised learning algorithms that can be used for classification and regression problems. The basic idea of these methods is to subdivide the input feature space into sub-regions and then adjust a local model to each of these subsets. This idea can be represented through a tree as can be seen in figure 2.4 (a), where the input space is subdivided into 2 regions in each node of the tree. Finally the feature space is divided into 5 regions (represented in the figure 2.4 (b)) in which a local model was fitted taking into account the examples of the training set that fall into each region.

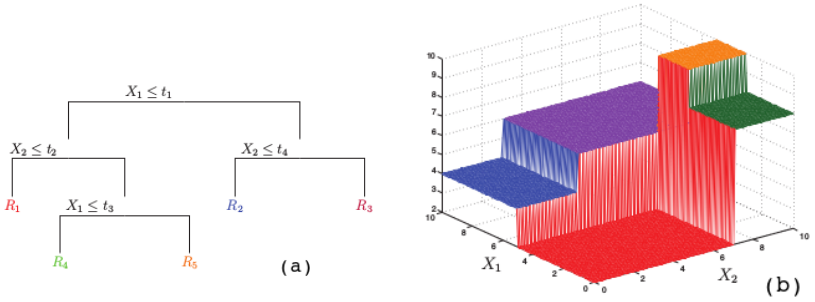

Figure 2.4: Classification scheme using a *Random Forest*. Figures extracted from *Murphy* [4].

Assuming we have N observations (x_i, y_i) with p features $(x_i = (x_{i,1}, ..., x_{i,p}))$ that we are going to use to predict a response y_i, that our input space is divided into M

sub-regions and that we assign a constant c_m to each region R_m, then the prediction for a new observation x is:

$$f(x) = \sum_{m=1}^{M} c_m I_m(x) \tag{2.7}$$

where $I_m(x)$ is a window function that is equal to 1 if $x \in R_m$ and 0 if $x \notin R_m$. If we adopt a minimisation criteria of the sum $\sum_i (y_i - f(x_i)^2$, it can be demonstrated that the constant c_m that best fits the data in each region is the average $\bar{c_m} = prom(y_i|x_i \in R_m)$. To find the best way to partition the input space, we performed a search in a grid of (j, s) in each node, where j is the variable that will be used to subdivide the space and s is the dividing point. Finally each node divides our space in 2 sub-regions

$$R_1(j, s) = \{x|x_j \leq s\} \quad ; \quad R_2(j, s) = \{x|x_j \geq s\} \tag{2.8}$$

Then the problem of finding the best sub-division is reduced to finding the pair (j, s) that minimises the sum of the squares:

$$\min_{j,s} \left[\min_{c1} \sum_{x_i \in R_1(j,s)} (y_i - c_1)^2 + \min_{c2} \sum_{x_i \in R_2(j,s)} (y_i - c_2)^2 \right] \tag{2.9}$$

Once we find the pair (j, s) that gives us the best partition in this node, our input space will be divided into 2 sub-regions. Then we apply the same criteria to subdivide each of these sub-regions in an iterative way [2]. Similarly to the KNN case, if we have very large trees (decision trees with lot of nodes) we are going to over-fit the data (see section 2.1.2) and model the noise, while if we have very small trees we may not capture important details of our problem. In section 2.1.2 we will study different methods to compare supervised learning algorithms (or, as in this case, the same algorithm but with different parameters) and thus, avoid over-fitting without losing important information contained in our training set.

Although the mathematical model presented in this section only serves for regression problems, it can easily be generalised to classification problems taking into

account the distribution of observations of different classes that exist in each sub-region and defining the probability that a new observation belongs to the class K in a given region in a frequentist manner:

$$P(x \in K | x \in R_m) = \frac{N_{K,R_m}}{N_{R_m}} \tag{2.10}$$

where N_{k,R_m} is the number of examples of a given class k in R_m and N_{R_m} is the number of examples in R_m. Then we can classify a new observation taking into account which is the most probable class.

One of the main problems of the decision trees is their instability, this means that small changes in the training set can lead to very different predictions. That is why in general it is advisable to train many decision trees and then average the results of each one in order to obtain the final result. The algorithm known as *Random Forest* is an implementation of this technique [1].

This method consists on training N decision trees selecting randomly $m < p$ features in each node that will be studied to sub-divide the input space as explained above. In this way we can reduce the correlation between different trees. Although the value of m is a parameter that can be adjusted to each problem to obtain a better performance, is common to use $m = \sqrt{p}$ or $m = p/3$ [2] for classification and regression problems respectively.

It is important to note that the technique of training different machine learning algorithms (or the same algorithm but with different parameters) and then averaging the results of each one to obtained the final result can be used to combine very different and independent machine learning algorithms. This method is known as *bagging*. However, if the algorithms are approximately linear the final result will be approximately the same as if bagging had not been used, and so, it only makes sense to combine highly non-linear algorithms, such as the decision trees.

- Artificial Neural Networks (ANN):

  One of the most important machine learning algorithms are the Artificial Neural Networks (ANN). This algorithm was developed as a model of the human brain, where the basic unit is the neuron. In artificial networks each neuron, or node,

receives a signal that will be analysed and then produces another signal that will be sent to the neurons connected to that node. Then an artificial neural network is a set of neurons (see figure 2.5) inter-connected between then that has a first layer of neurons that receives the input data (features), a set of intermediate layers that process the data, called hidden layers and, finally, an output layer that produces the prediction for which the algorithm was trained. The number of neurons in each layer and the number of hidden layers is called the architecture of the neural network and should be varied in each problem to look for a better optimisation.

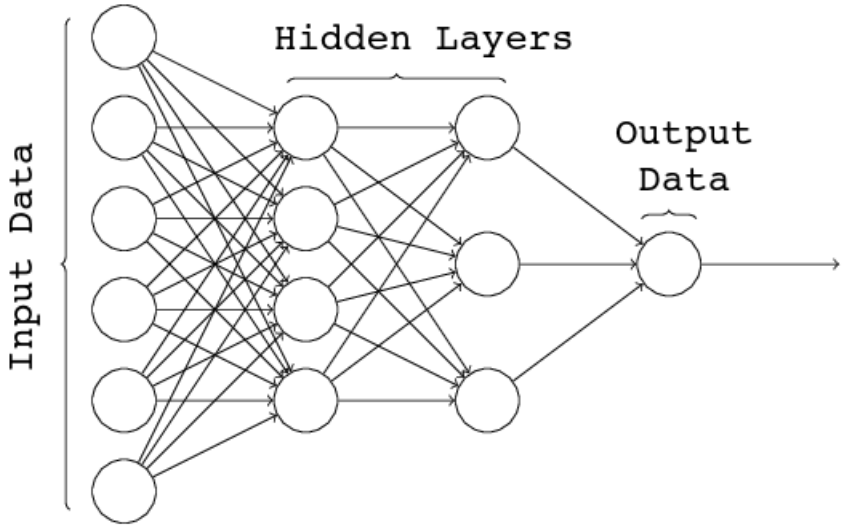

Figure 2.5: Classification scheme of an artificial neural network.

To understand this method in more detail it is necessary to study the behaviour of each neuron individually. Each neuron receives a set of input data x_i and, from these, generates a linear combination $x_i\omega_i$ of them, where ω_i is a vector of weights that depends on each neuron. In addition each neuron consists of an activation function $f(x \cdot \omega)$ that receives the input data weighted by ω and produces a value between 0 and 1. the name that each type of neuron receives depends on the activation function. The simpler Neurons are called perceptrons and are characterised by an activation function where $f(x \cdot \omega) = sgn(x \cdot \omega)$. Another type of neurons widely used are the sigmoidal ones that are defined according to:

$$f(x \cdot \omega) = \frac{1}{1 + e^{x \cdot \omega}} \tag{2.11}$$

Once you have defined the architecture and the type of neurons that will be used in each layer, you need to find for each neuron the weights ω that produce the best result. Although there are lot of methods to do this, the most used algorithm is the one called *backpropagation*. This technique consists on determining the errors made layer by layer, starting with the output layer. Assuming that our training set has n examples of the form (x_i, y_j) with $i = 1 \cdots p$ and $j = 1 \cdots k$, where x_i are the features and y_j are the data that we want to predict, we can define the error of our network as:

$$E = \sum_{i=1}^{n} \sum_{j=1}^{p} (y_{j,i} - \hat{y}_{j,i})^2 \tag{2.12}$$

Then the *backpropagation* algorithm looks for the ω weights that minimise this error following these steps [3]:

1. Initialise all weights ω to random values (between -0.5 and 0.5).
2. For each example of the training set make a prediction, with the network, which we will call $\hat{y_j}$.
3. Estimate the error in each of the predictions $\delta_j = \hat{y_j}(1 - \hat{y_j})(y_j - \hat{y_j})$.
4. For each neuron h in a hidden layer estimate the error $\delta_h = o_h(1-o_h) \sum_{j=1}^{p} \omega_{j,h}\delta_j$, where o_h is the value produced by the activation function of the neuron h and $\omega_{j,h}$ are the weights assigned by the neurons j of the next layer to the values that come out from the h neuron.
5. Update the weights: $\omega_{j,h} = \omega_{j,h} + \eta\delta_j o_{j,h}$, where η is a value what quantifies the rate at which the neurons are going to be updated in each iteration and $o_{j,h}$ is the value produced by the neuron h of a layer which receives the neuron j from the next layer.
6. Go back to item 2 and iterate until you get the required precision.

It is important to note that this algorithm does not ensure convergence at a global level. However the problems that are usually analysed with neural networks are highly dimensional, and so, when we reach a local minimum in a certain parameter usually it is not a local minimum on all the other parameters, and so, the algorithm will not get stuck in there.

It is also important to clarify that there are different algorithms that implement some small variations to what was described here. By example, there are neural networks in which the *backpropagation* algorithm is modifying by adding extra terms in the updating errors or looking for the minimum in a stochastic way. There are also algorithms in which the neurons belonging to hidden layers are interconnected in a cyclic way.

- *Deep learning*:

  As we saw in previous examples, the supervised learning techniques look for correlations between the features of the input data, and the output data that one wants to predict. Historically, due to computational limitations, the algorithms work with, at most, a few tens of features. This is way much of the work was dedicated to find the best features that solve a certain problem. Currently, thanks to the technological advance, there are new methods, called deep learning techniques, which incorporates the process of finding the best features inside the machine learning algorithm. This paradigm shift is clear with an example. At first, when trying to predict whether an image contained a car or not, the first step was to reduced the images of the training set in order to obtain the best features, and then, with these features, to train different supervised learning algorithms that, finally, predicts if there was a car or not in the image. Currently *deep learning* algorithms uses as input data the images and the first layers of these algorithms automatically produced the features that will be used by the following layers to make the prediction (see figure 2.6).

  Although these techniques are being used in lot of applications with very good results, due to the high number of free parameters that they have, they require a great computation power and, in general, of the utilisation of GPU's (Graphical Processing Units) for their training. This is why most of the problems are still studied through traditional methods and, if these method do not reach the required preci-

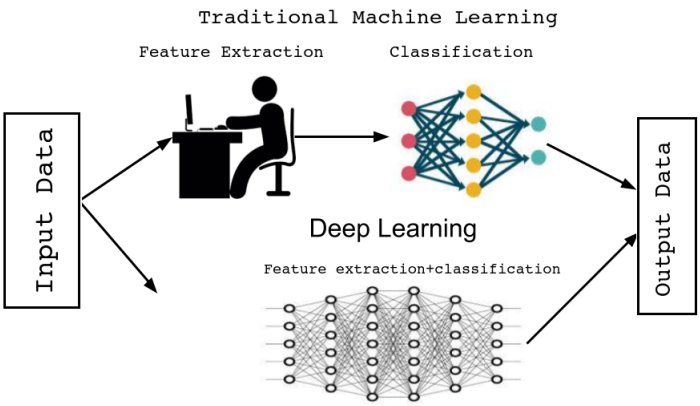

Figure 2.6: Comparison between the classification schemes of traditional machine learning and *Deep Learning* algorithms.

sion, then they are studied with *deep learning* algorithms.

2.1.2 Model evaluation and Over-fitting.

As was stayed in previous section, the variety and quantity of different supervised learning models is very large, and we must also take into account all the different implementations and variations on each algorithm and the different architectures that can be studied on a certain implementation of a model. It is clear that given the amount of different algorithms that can be used to solve the same problem, it is necessary to have a robust and objective measure of the performance of each method in order to be able to compare them and to choose the one with the best performance.

It is important to keep in mind that our ultimate goal is to predict values of y for new observations that are independent of the training set used. Then, in order to accurately measure the performance of our model, it is necessary to evaluate it also in set independent of the used in the training. This is why the original training set should be divided in two mutually exclusive sets, a proper training set and a test set. Then, we must train our model using only the examples of the training set and then the performance should be evaluated making predictions for the examples of the test set. Taking into account that in the test set we also know the real values of those variables that we want to predict, we can compare the predicted values with the real ones and thus have a reliable measure of the performance of our model. In this way we also avoid over-fitting the data, that is, modelling the small

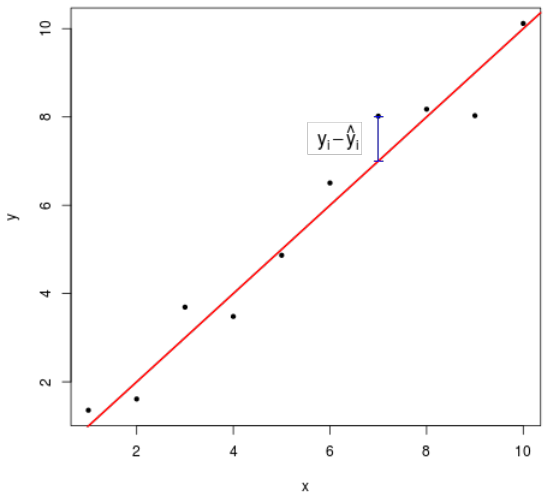

Figure 2.7: Traditional method to evaluate the performance of a supervised learning algorithm.

variations present in our data.

If we are working on a regression problem, the most natural and intuitive way to measure the performance of our algorithm, it is simply through a difference of squares χ^2 defined by:

$$\chi^2 = \frac{\sum_i^N \omega_i (y_i - \hat{y}_i)^2}{\sum_i^N \omega_i} \tag{2.13}$$

where y_i is the real value of the variable that we want to predict, $\hat{y}_i$ is the predicted value by our method and ω_i is the relative weight that we give to each example. The performance of a regression algorithm can also be visualised by plotting the predicted value vs the real one (See figure 2.7), where a perfect prediction will follow a straight line 1-1, and where the value of χ gives a measure of the dispersion of the predictions around the perfect one.

On the other hand, if we are working on a classification problem, the best way to evaluate the performance of our model is counting the amount of correct and incorrect predictions that we made with our algorithm. In turn, this information can be presented in various ways. One of the most common forms is through a confusion matrix (In figure

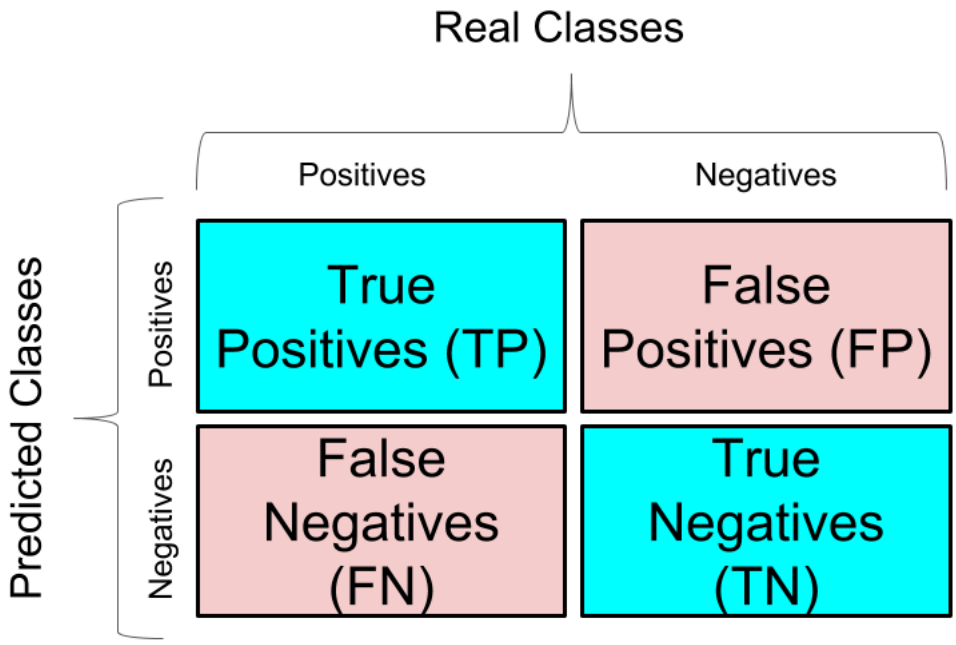

Figure 2.8: Confusion matrix used to evaluate the performance of supervised machine learning algorithms.

2.8 is shown a confusion matrix for a binary classification problem [1]), in which each row represent a predicted class and each column a real class (or viceversa). In this way it can be seen that the diagonal represent the wright predictions, while the off-diagonal elements represent the wrong predictions.

Using the information contained in this matrix it is convenient to define the following statistics that help to understand the performance of our algorithm and are useful when deciding between different techniques.

- True Positive Rate (TPR): Also known as recall or sensitivity, this statistic gives us an idea of the probability of classifying correctly a positive example.

$$TPR = \frac{TP}{TP + FN} \tag{2.14}$$

- False positive rate (FPR): This statistic gives us a measure of the probability of

[1] It is common to called, in binary classification problems, one class as 'positive' and the other 'negative'.

classifying incorrectly a negative example.

$$FPR = \frac{FP}{FP + TN} \tag{2.15}$$

- Precision: This value gives us an idea of the probability that an example predicted as positive is actually positive. In addition, you can define the false discovery rate (FDR) as the probability that an example predicted as positive is actually negative.

$$P = \frac{TP}{TP + FP} = 1 - FDR \tag{2.16}$$

- Accuracy (ACC): Gives us a measure of the probability that the prediction of our model is the correct one.

$$ACC = \frac{TP + TN}{TOTAL} \tag{2.17}$$

Considering these definitions, it is very common to graphically analyse the performance of an algorithm through the so-called ROC curve in which the rate of true positives (TPR) versus the false positive rate (FPR) is plotted. In figure 2.9 you can see an example of a ROC curve in which we compared 3 classifiers. From this graph you can quickly conclude that the classifier with the best performance is the one shown in red since it has a high true positives rate and a low false positive rate. It is worth to remark that a random classifier will have a curve 1-1 on this plane, since the probability that the random method correctly classifies an observation is 50%.

It is important to keep in mind that the choice of the best algorithm depends strongly on both the problem that we want to solve as well as the data set used, the different computational or temporal limitations that may arise and the degree of precision that we want to reach.

2.2 Non-Supervised Learning.

The second major category of automatic learning methods are the so-called unsupervised learning algorithms. This class of algorithms only works with input data and its main goal is to look for correlations or patterns between them, this is why this techniques are also called Knowledge Discovery methods. The main difference with supervised learning is

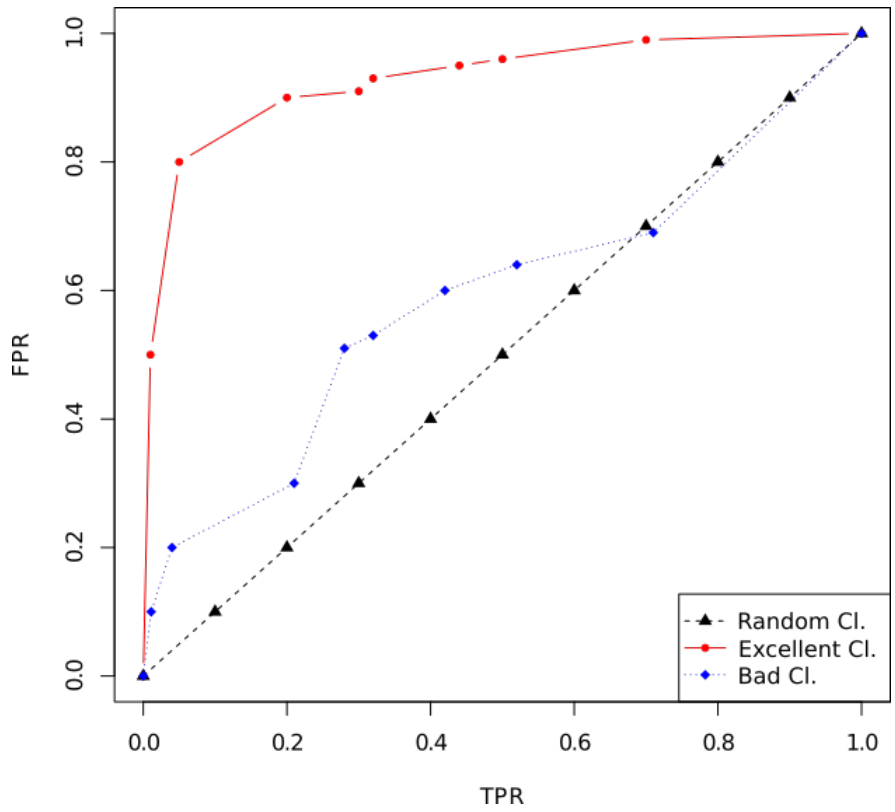

Figure 2.9: Example of ROC curve used to compare different supervised learning methods.

that, by not specifying what kind of correlations we want to find or predict, there is no obvious metric to quantify the error committed by our algorithm.

2.2.1 Most important non-supervised learning models.

- *K-means*:

  The *k-means* algorithm is one of the simplest grouping techniques. Given a set of m input data $\vec{x_i}$ belonging to the feature space $\mathcal{F}$ with $i = 1, .., m$ the algorithm consists on:

  1. Randomly choose k centres $c \in \mathcal{F}$.
  2. Given a distance measure [2] in $\mathcal{F}$, assign each observation to the nearer group.

[2] The Euclidean distance is usually used if $\mathcal{F} = \mathcal{R}^n$

3. Update the value of the centres, averaging the observations that belong to each group.

4. Iterate between step 2 and 3 until the assignments do not change or a convergence criterion is reached.

If worth to note that if the number of groups that best adapt to the problem it is not known a priori, it is necessary to iterate in the number of centres to be fitted.

- Gaussian Mixture:

  The Gaussian mixture is an algorithm similar to *k-means* but it assign each observation to a multi-dimensional Gaussian. Given a set of m input data $\vec{x_i} \in R^n$ with $i = 1, .., m$ the algorithm consists on:

  1. Randomly choose parameters of the n-dimensional k Gaussians (i.e the centres and the covariance matrices).

  2. For each observation estimate the probability to belong to each Gaussian and assign that observation to the most probable one.

  3. Update the parameters of each Gaussian taking into account the observations assigned to it.

  4. Iterate between step 2 and 3 until the assignments do not change or a convergence criterion is reached.

  As in the case of the *k-means* algorithm, if one does not know a priori the number of Gaussians, one should explore different values of this parameter to find the best fit to our data.

- Principal Components Analysis (PCA): The principal component analysis method is an statistical technique that allows to reduce the dimensionality of a data set through the analysis of the correlations between the different variables. This technique allows us to find the best way to represent n observations $x \in \Re^p$ as a linear combination of $q \leq p$ vectors in $\Re^p$.

To study the correlations between the different variables, we construct the covariance matrix defined by:

$$C_{i,j} = \sum_{k=1}^{n} \frac{(x_{i,k} - \bar{x_i})(x_{j,k} - \bar{x_j})}{n} \tag{2.18}$$

where $\bar{x_i} = \sum_{k=1}^{n} \frac{x_{i,k}}{n}$. From the definition it follows that the covariance matrix is a symmetric matrix, which diagonal elements are the variances of each variable and the off-diagonal elements are the covariance between the different variables.

By studying the eigenvalues of this matrix, we can find the directions of greatest correlation. To find such eigenvalues it is necessary to determine the roots of the characteristic polynomial of degree $\leq p$ defined by $det(C_{i,j} - \lambda.I)$. Assuming that we have p real eigenvalues, we can find p eigenvectors that define a new basis in which the covariance matrix is diagonal and whose diagonal elements are the eigenvalues corresponding to each eigenvector. Mathematically this can be expressed as follows:

$$D = P - 1CP \tag{2.19}$$

where $P = (\alpha_1 \alpha_2 \dots \alpha_n)$, $\alpha\ i$ is the eigenvector corresponding to the eigenvalue λ_i in column and D is the diagonal covariance matrix in the base of eigenvectors.

In figure 2.10 you can see an example of a two-dimensional PCA analysis, where you can see that the direction of the first component is the one with the highest correlation in the data. It is important to note that the basis change defined by the P matrix corresponds to a rotation.

Intuitively it can be seen in the example shown in the figure 2.10 that, the data set can be represented only by the coordinate corresponding to the first principal component without losing much information because most of of the correlation of the data is just the correlation in that direction. This is why one of the most important applications of this technique is to reduce the dimensionality of multi-dimensional problems.

In figure 2.11 we can see an example in which the reconstruction of the image of a face is shown as function of the components used for its representation. It can be

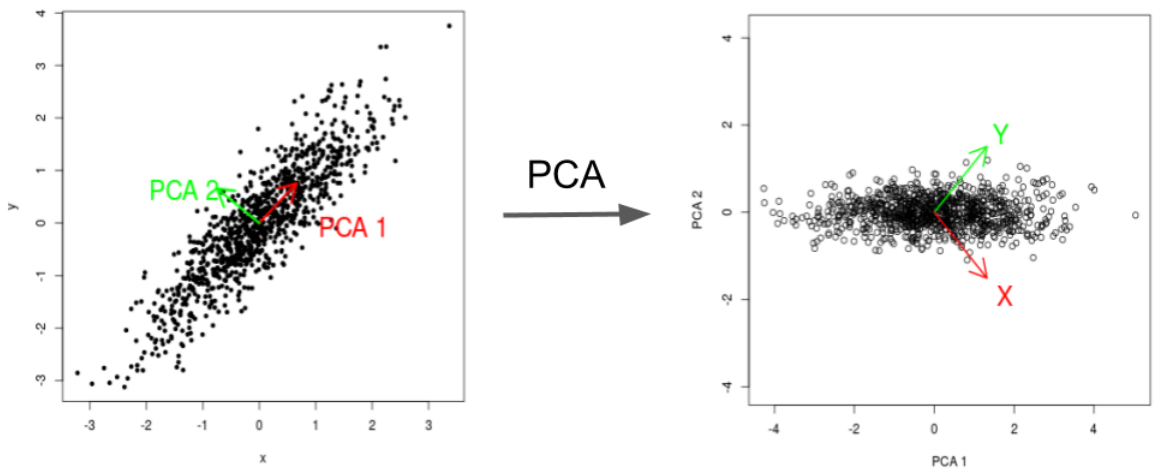

Figure 2.10: Schema of a two-dimensional principal component analysis.

seen that the fewer components we use the more information we loss. However, the information that is gained by incorporating components of lower correlation (see the reconstructions of the last 2 rows in the figure 2.11) it is not much, and so, ignoring these components in the representation of the image will not lead to large errors.

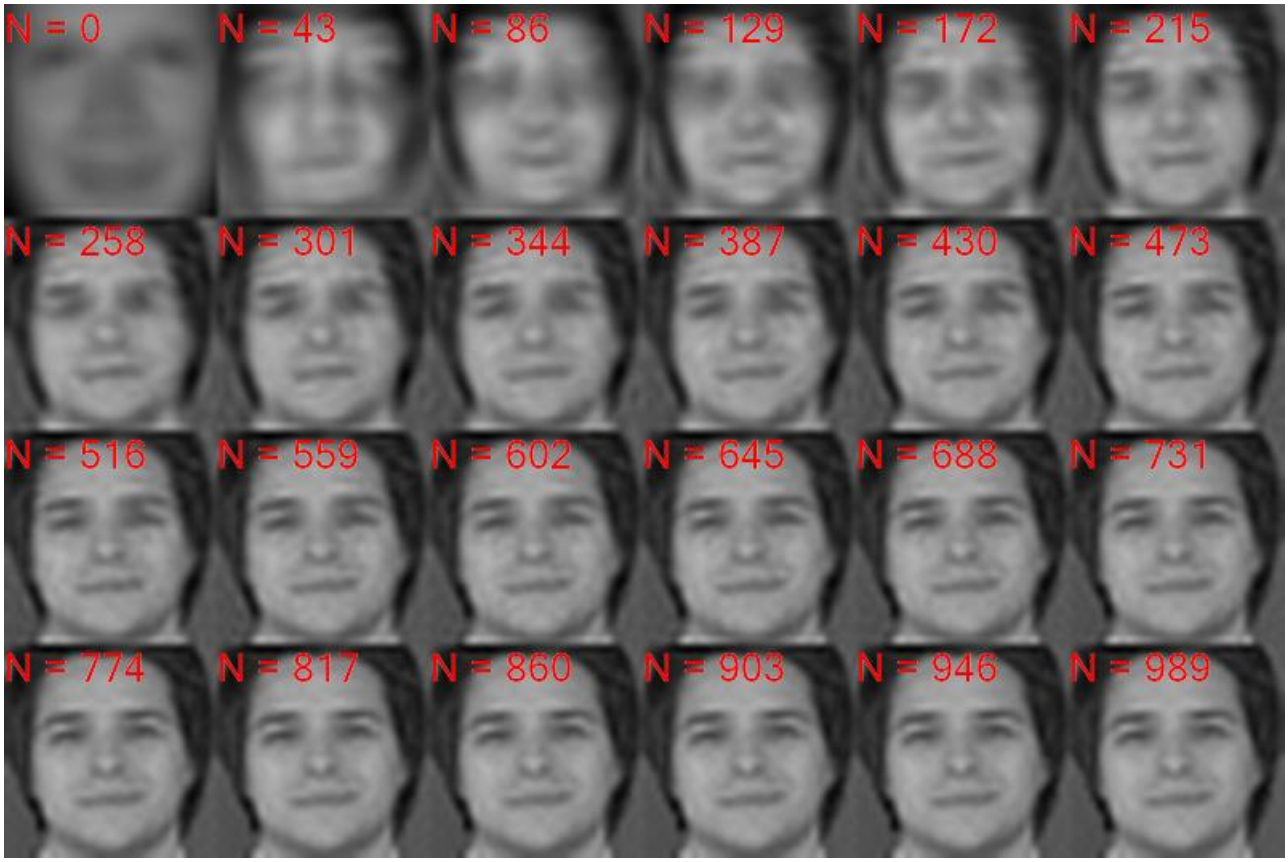

Figure 2.11: Example of a facial reconstruction algorithm by varying the number of the principal components used for its reconstruction. Figure extracted from http://www.declanoller.com/tag/pca/

Bibliography

[1] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, Oct 2001. ISSN 1573-0565. doi: 10.1023/A:1010933404324. URL https://doi.org/10.1023/A:1010933404324.

[2] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York Inc., New York, NY, USA, 2001.

[3] Thomas M. Mitchell. *Machine Learning*. McGraw-Hill, Inc., New York, NY, USA, 1 edition, 1997. ISBN 0070428077, 9780070428072.

[4] Kevin P. Murphy. *Machine Learning: A Probabilistic Perspective*. The MIT Press, 2012. ISBN 0262018020, 9780262018029.

Chapter 3

Construction of a catalogue of merging clusters with Machine Learning techniques.

In the first part of this chapter we will present a brief summary of the main works about merging galaxy clusters and their importance in modern cosmology. In the second part of this chapter, we will present a software, developed during the curse of this doctorate, that builds merging clusters catalogues using photometric and spectroscopic information. The results presented in this chapter were published in the 57^{th} Argentinian Astronomical Association bulletin in the year 2015 [22] and in the journal *Monthly Notices of the Royal Astronomical Society* in the year 2016 [23].

It worth to remark that this work was done in collaboration with Dr. Mariano Dominguez, Dr. Dante Paz and Dr. Manuel Merchán.

3.1 Merging clusters and their importance in cosmology.

The standard cosmological model predicts a hierarchical structure formation, in which the larger structures are formed by the fusion of smaller structures. In this context, the study of merging galaxy clusters, its main properties, merger rates, etc. are fundamental for the development of modern cosmology.

The *bullet cluster* (1E0657-558) is the most studied merging cluster. *Clowe et al. 2006*

[17] performed a joint study of the X-rays emission, optical emission and gravitational lens effect. This system is composed by 2 sub-structures clearly distinguishable in all images. In this study it was found that there is an offset (See figure 3.1 *left*) between the mass distribution, traced by the maps of gravitational lenses (black contours) and the X-rays emission (black points) that comes from the hot gas. On the other hand the distribution of galaxies coincides with the mass distribution of both sub-structures (See figure 3.1 *right*). As is known, most of the baryonic mass of galaxy clusters are in the form of hot gas and not in stars (there are approximately 10 times more mass in hot gas than in stars [67]), This implies that the offset between the distribution of mass and the distribution of gas can not be produced by galaxies.

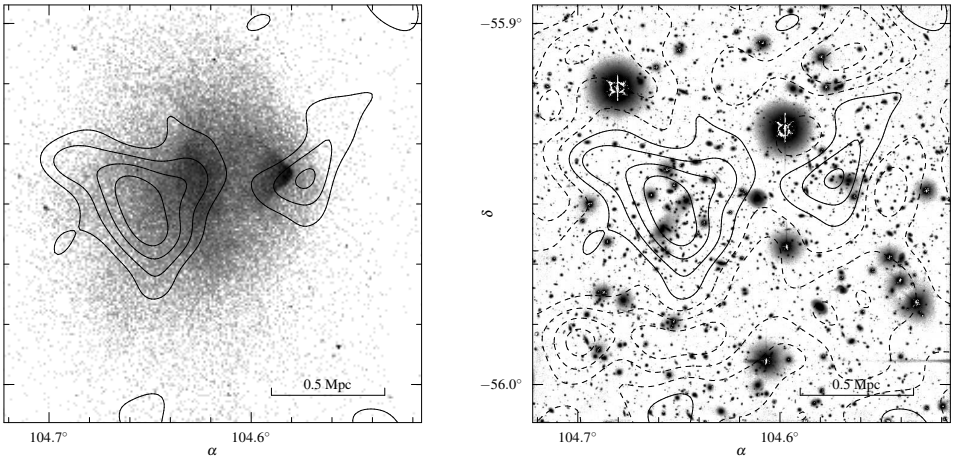

Figure 3.1: *Left* X-ray image superimposed to mass contours estimated with weak gravitational lenses studies. *Right* Optical image superimposed to mass contours estimated with weak gravitational lenses studies. Figures extracted from *Clowe et al.* [17]

In order to explain the properties observed in this system they proposed that the mass distribution of the cluster is dominated by the dark matter. Then this system will actually be the product of 2 sub-clusters that collided. Because galaxies interact only gravitationally, the probability that 2 galaxies collide in the interaction is very low, which translates into their non-collisional behaviour. This causes that galaxies pass by without being substantially affected by the collision. On the other hand, the gas, in addition to the gravitational interactions, also interact through collisions (viscosity, pressure, etc.),

which implies that in the merger of the sub-groups its distribution is affected, producing the observed offset.

Although this was not the first reference to dark matter is considered, by a large part of the scientific community, as one of the most important evidences of it. It is important to clarify that there are also modify gravity theories (MOND, TeVeS, etc.) that can explain the observed phase shift without the need to add a new form of matter [5, 47].

3.1.1 A challenge for the standard cosmological model.

In this section we present a brief summary of the main works that addressed the study of the main properties of the *bullet cluster* and the recent discussion of whether these properties may present a problem for the standard cosmological model.

In order to perform a detailed study of the gravitational lenses effect *Markevitch et al.* [60] modelled the radial mass density profile of the main substructure using a *King* profile $\rho = \rho_0(1 + \frac{r^2}{r_c^2})^{-3/2}$. Fitting this profile they found that $\rho_0 = 2.6 * 10^{-25} g/cm^3$ and $r_c = 210 Kpc$, which gives an estimate of the mass of $M \approx 7 * 10^{13} M_\odot$. For the secondary group they obtained $\rho_0 = 1.3 * 10^{-24} g/cm^3$ and $r_c = 70$ Kpc with a truncation radius of $r_{tr} = 150$ Kpc. Through X-rays emission they found that the relative velocity between both structures is $v_{rel} \approx 4500 km/s$, while the speed of the secondary cluster along the line of sight is $600 km/s$ ([6]), implying that the collision is approximately in the plane of the sky.

The discovery of this cluster, gives place to a discussion about what is the probability of finding an object with similar characteristics in cosmological simulations. In particular, the most striking property of this system is the relative high speed of the secondary cluster with respect to the main substructure ($v_s \approx 4500 km/s$ [60]).

One of the first works to perform a statistical study on the probabilities of finding an object similar to the *bullet cluster* in cosmological simulations was the work done by *Hayashi & White* [44] in 2006. In this work, they look for dark matter haloes, with velocities and masses similar to those observed in the *bullet cluster*, in the *Millenium* simulation [88]. They found that 1 out of 100 secondary haloes of massive clusters ($> 10^{14} M_\odot$) has a speed comparable to that of the bullet cluster. They arrive at the final conclusion that the *bullet cluster* is rare but not impossible in a universe with a ΛCDM cosmology. However, it is important to note that this probability varies significantly with

the speed of the cluster.

In 2007 *Farrar & Rosen* [30], using a velocity of $v_s = 4740km/s$ (measured by *Markevitch* with new data [59]), calculated that the probability of finding a system similar to the *bullet cluster* was $0.8 * 10^{-7}$, which could imply a problem for the standard cosmological model.

Following a different line of work, (*Milosavljevič et al.* [66], *Springel & Farrar* [87], *Mastropietro & Burkert* [61]) studied the properties of the *bullet cluster* through hydrodynamic simulations of the merging process. These works found that the velocity of the gas shock is considerably greater than the velocity of the centre of mass of the secondary cluster. For example, when the shock velocity is $\sim 4500km/s$, the velocity of the centre of mass of the secondary group it is only $\sim 2600km/s$ [87]. Although, taking into account this new speed for the secondary substructure, the tension with the standard cosmological model can be eased, some caution must be had with these results since according to *Mastropietro & Burkert* the properties of gas shock are not reproduced faithfully by these simulations.

In a statistical study, *Lee & Komatsu* [56] quantified the rarity of systems similar to the bullet cluster looking for clusters with a similar speed to the one measured in the secondary substructure of the bullet cluster in a simulation of $3Gpc$ on the side. They studied the velocity distributions of subgroups at a distance between $1 - 3R_{200}$, where R_{200} is the radius in which the density of the dark matter halo is 200 times greater than the density of the universe. Then they compared this distributions with the initial velocities used in the hydrodynamic simulations cited above, finding that no cluster has, at that distance, the speed necessary to explain the X-ray emission ($3000km/s$ [61]), again giving rise to a conflict with the ΛCDM model.

In a similar analysis *Forero-Romero et al.* [33], studied the distribution of offset between dark matter and gas in galaxy clusters of the *MareNostrum* cosmological simulation [40]. They found that simulated clusters show similar and even greater mismatches than those observed in the *bullet cluster*, finally concluding that this property does not represent a problem for the standard cosmological model.

Following this line of study, *Watson et al.* [92] looked for systems similar to the *bullet cluster* in a simulation of $6GPc$ of side and with 6000^3 dark matter particles. They considered secondary systems with masses larger than $10^{13}M_{\odot}$ which were colliding with

clusters of masses greater than $7 * 10^14 M_\odot$ at $z = 0.32$. Taking into account the relative speed between the substructures and the distance between them, they found that the bullet cluster is an extreme system, however, there are other objects that have similar and even higher speeds but with a greater distance between their substructures. Finally, based on these results, they conclude that there will be no problem between the properties of the *bullet cluster* and the standard cosmological model.

Recently *Lage & Farrar* [54] made a detailed study of the fundamental properties of the *bullet cluster*. For this study they performed hydrodynamic simulations varying the initial conditions and then compared them pixel by pixel with the real observations of the *bullet cluster*. In the table 3.1 it can be seen a comparison between the real observations and the results obtained by different authors using hydrodynamic simulations. Using χ^2 as a figure of merit they found that the simulation performed by *Lage & Farrar* [54] is the one that best fit the actual observations.

Authors	M_{Main} ($10^{14}M_\odot$)	M_{Sub} ($10^{14}M_\odot$)	$R_{Initial}$ (kpc)	$V_{Initial}$ (km/sec)	V_{2500} (km/sec)	χ^2
Milosavljevic et al. [66]	12.7	2.54	4600	0	1546	–
Randall et al. [74]	8.56	5.25	4000	4100	4225	–
Springel & Farrar [87]	15.0	1.50	3370	2057	2386	13.67
Mastropietro & Burkert [61]	7.13	1.14	5000	3000	3228	19.93
Lage & Farrar [54]	19.1	2.59	2800	2799	2943	3.92

Table 3.1: Comparison of initial velocities for different simulations. The velocity V_{2500} is estimated assuming that the cluster moves as a point mass in a free fall path from its initial position to a separation of $2.500kPc$. Table extracted from *Lage & Farrar* [54].

To estimate the probability of occurrence of an object with these characteristics, *Lage & Farrar* used the *Horizon* cosmological simulation [51]. They look for dark matter haloes with masses and distances similar to those found in both substructures of the *bullet cluster*. After that they estimated the speeds between them. They found that the speed estimated through their hydrodynamic simulations ($\approx 2943km/s$) is 1.4σ of the average for haloes with similar masses. Finally they conclude that the speed of *bullet cluster* is consistent with a ΛCDM universe.

Taking into account the same sample of halos, *Lage & Farrar* also studied the distribution of concentration index and the distribution of forms, quantified by the ratio of the principal axes of an ellipse fitted to the dark matter particles of each substructure. All

estimated values for the *bullet cluster* are compatible with a standard cosmology, with the possible exception of the concentration index of the main substructure, which is lower than expected for a halo of similar mass.

Finally, they come to the conclusion that the *bullet cluster* does not present a problem for the ΛCDM model, although they emphasise the importance of using larger cosmological simulations in order to have more haloes with masses similar to those found in the *bullet cluster*.

3.1.2 Constraint in the cross-section of the dark matter particle.

Since the discovery of the bullet cluster, many projects were created in order to perform a search for objects with similar features. Thanks to this, we were able to find new interacting clusters (Baby Bullet [11], Abell 520 [18, 48], Abell 2744 [63], etc.).

Using images of the space telescopes *CHANDRA* and *Hubble Harvey et al. 2015* [43] built a sample of 72 colliding clusters. Taking into account that each system has 3 main components (gas, dark matter and galaxies) and that when colliding the distributions of these components can be affected differently according to their physical properties, they developed an scheme (see figure 3.2) to study this 3 components. In magenta is shown the distribution of the diffuse hot gas observed through its X-ray emission. In blue is plotted the distribution of dark matter inferred through the measurement of their gravitational lens effect. Finally, the distribution of the observed galaxies (stars) in the optical band is plotted in green.

Taking this scheme into account, they define the separation between the galaxy distribution and the gas distribution δ_{SG} and the separation between the galaxy distribution and the dark matter distribution in the parallel and perpendicular direction to $S-G$ as δ_{SI} and δ_{DI} respectively. Using these quantities, they define $\beta \equiv \delta_{SI}/\delta_{SG}$ and $\beta_\perp \equiv \delta_{DI}/\delta_{SG}$. Assuming that the galaxies are non-collisional, β is related to the dark matter particle self-interaction cross-section. On the other hand, taking into account the geometry of the problem and, adding the measurement of all the clusters, it is expected that $< \beta_\perp >= 0$.

Applying this method to the 72 clusters, they found $< \beta_\perp >= 0.06 \pm 0.07(68\%CL)$ and $< \beta >= -0.04 \pm 0.07\ (68\%CL)$, which results in a limit for the self-interaction cross-section of the dark matter particle of $\sigma_{dm}/m < 0.47cm^2/g\ (98\%CL)$.

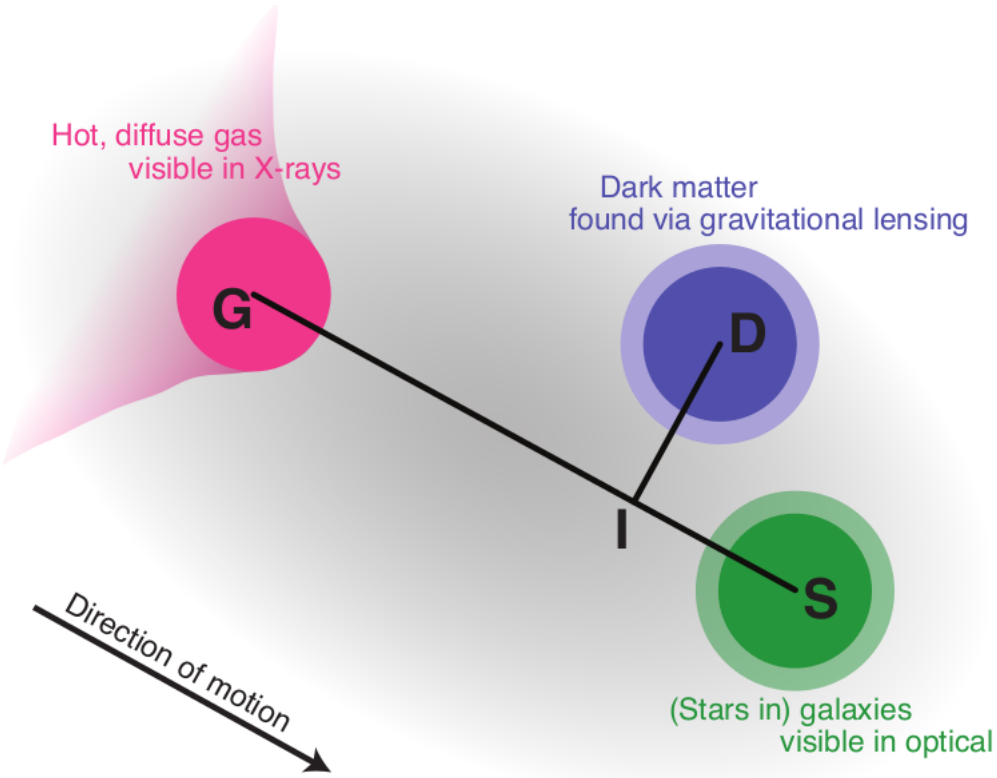

Figure 3.2: Schematic geometry of a merging cluster. Figure extracted from *Harvey et al. 2015* [43].

3.2 The MeSsI (*Merging Systems Identification*) algorithm.

Taking into account the importance of merging clusters studies in modern cosmology, a lot of efforts were focused on the search of such objects. To this end different identification methods, that take advantage of the particular characteristics of the clusters in different electromagnetic bands, were developed [38, 93]. In particular, there is a great collaboration dedicated exclusively to the search, identification and analysis of the colliding systems called *Merging Cluster Collaboration*. This collaboration centres its search using radio and X-ray information [38]. After this search they continued analysing in an individual way each one of the identified systems [7, 37, 39].

In this context we develop an automatic method for the identification of merging galaxy clusters. In order to do this we use supervised learning techniques that, based on properties estimated from photometric and spectroscopic data, classify a given cluster

between merging system or not.

3.2.1 Preparation of the training set using cosmological simulations.

As was specified in section 2 the first step when working with supervised machine learning algorithms is to build a training set representative of the objects that will be classified later. In our case, the ultimate goal is to classify a cluster of galaxies using photometric data (magnitudes and colours), astrometric images (angular positions) and spectroscopic information (redshift).

For the construction of the training sample we used data from the *Millenium* cosmological simulation [88] and from the semi-analytical model of *Guo et al.* [42].

The *Millenium* simulation [88] uses 2160^3 particles that evolve in a cube with $500MPc$ sideways, where each particle has a mass of $m_p = 8.61 * 10^8 M_\odot$. The initial conditions were generated using the `CMBFAST` code [82] to compute a linear power spectra with a ΛCDM model with cosmological parameters consistent with the WMAP 2003 data ($\Omega_{b,0} = 0.045$; $\Omega_{m,0} = 0.25$; $\Omega_{\Lambda,0} = 0.75$; $H_0 = 73$) [86].

The simulation data is organised in different outputs (snapshots) that represent the temporal evolution of the particles, where the snapshot 0 is the beginning of the simulation and the snapshot 63 the current time ($z = 0$).

In order to identify a set of gravitationally bound dark matter particles (haloes of dark matter) they used the `Friends-of-Friends (FOF)` [45] algorithm. This algorithm identify the dark matter haloes present in a given snapshot by grouping the particles according to a spatial proximity criterion. Then, applying an adaptation of the algorithm `SUBFIND` [35] they identified the subhaloes that are inside of other halos. Once you have identified all the halos and subhaloes, you can calculate different haloes properties such as angular momentum, velocity dispersion, mass, radii, etc.

Taking into account that in a ΛCDM universe the formation of structures occurs in a hierarchical way, that is to say that first the smaller structures are formed and then, through mergers of these, the larger structures are formed, it is necessary to study the temporal evolution of dark matter haloes. Once the halos and subhaloes have been identified in the different outputs of a given simulation, the next step is to obtain the merger tree of them. Building a merger tree consists of finding out when and how the (sub) halos merged until they form the (sub) halos present in the *snapshot* 63 corresponding to $z = 0$.

Each particle has an identifier number or ID that allows us to find it in the different snapshots of the simulation. Thus, by taking consecutive pairs of temporary snapshots, one can analyse which particles belong to each (sub)halo in each snapshot and so, study which (sub)halos were merged to give place to the (sub)halos present in the following snapshot [89].

Once the merger trees of each dark matter subhalo is constructed, we know the evolution of this object throughout the simulation. As this simulation is only of dark matter, to be able to build a galaxy catalogue comparable to the real ones, it is necessary to relate dark matter halos with real objects (galaxies, galaxy clusters, etc.). One of the most efficient methods for populating haloes with galaxies are the so-called semi-analytical models. These methods populate the halos through different recipes that model the physical processes that can not be simulated due to the resolution of the simulation [42].

In order to carry out this work, we used the data of the semi-analytical model of *Guo et al.* [42] applied to the *Millenium* cosmological simulation. It is worth mentioning that this model provides us with photometric information (absolute magnitudes in the SDSS bands), $3D$ position and $3D$ speeds for each galaxy. Finally, to build a catalogue similar to the SDSS, it is necessary to choose a position for the observer and, from there, estimate the angular position, redshift and apparent magnitudes of each galaxy. In order to take into advantage all the simulation volume, we built 8 catalogues choosing as observer position each one of the 8 nodes of the simulation cube. Then, we apply the same photometric selection criteria ($14.5 < z < 17.77$) and the same angular mask as the spectroscopic SDSS to each of these catalogues.

To reproduce the same procedures that are performed in real surveys, we use a FOF algorithm to identify galaxy clusters in each of our 8 mock catalogues. Finally, we relate each identified cluster with a dark matter halo of the simulation taking into account which is the halo that shares the largest number of galaxies with the identified group.

By studying the merger tree of the haloes that were associated with the identified groups, we define as a merging cluster a cluster whose associated halo has undergone a major merger (define as a merger in which the ratio among the interacting halos is greater than 0.25) and whose substructures have survived until the snapshot 63 as different subhaloes within the same halo. In figure 3.3 you can see a representative scheme of a halo that had a major merger in the snapshot 61 and whose haloes survived as subhaloes until

the snapshot 63. The final sample has 1447 clusters with masses greater than $10^{13} M_{\odot}$ and more than 30 galaxies [1] at $z = 0$. From these, 132 are merging clusters. The figure 3.4 shows the distribution of the times in which mergers occurred for the groups that are in an interaction process. You can see that this distribution has a maximum in $t \approx 3Gyr$ consistent with what was found by other authors who studied the time it takes to the substructures produced by a merger to disappear according to the *Dressler-Shectman* test [69].

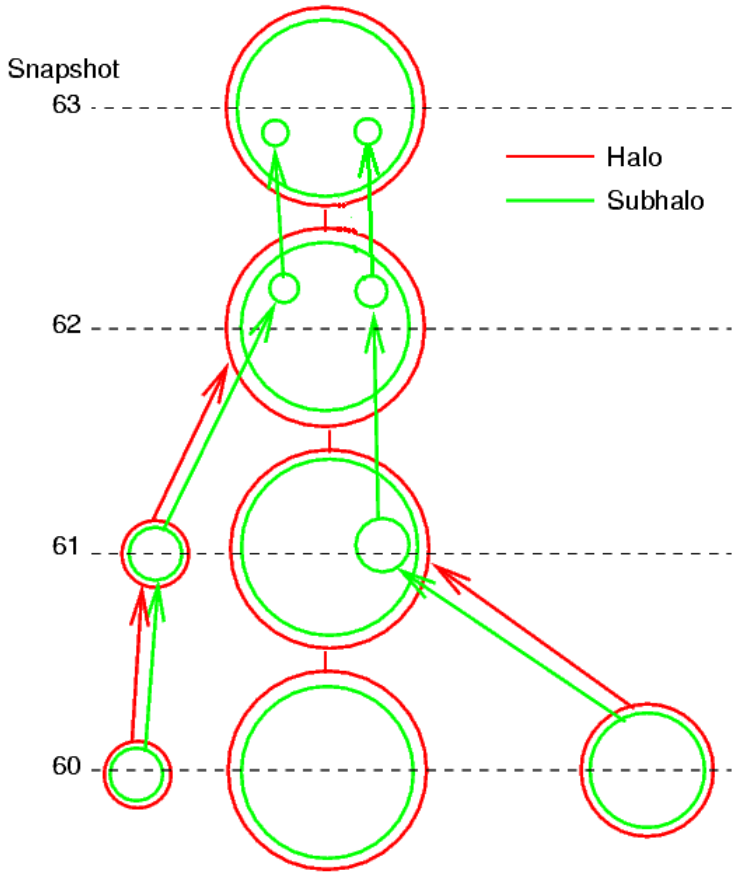

Figure 3.3: Scheme of a halo that had a major merger in the snapshot 61 and whose haloes survived as subhaloes until the snapshot 63.

[1] As explained in section 3.2.2 we set a minimum in the number of member galaxies because the *Dressler & Shectman* test can only be applied to groups with more than 30 galaxies.

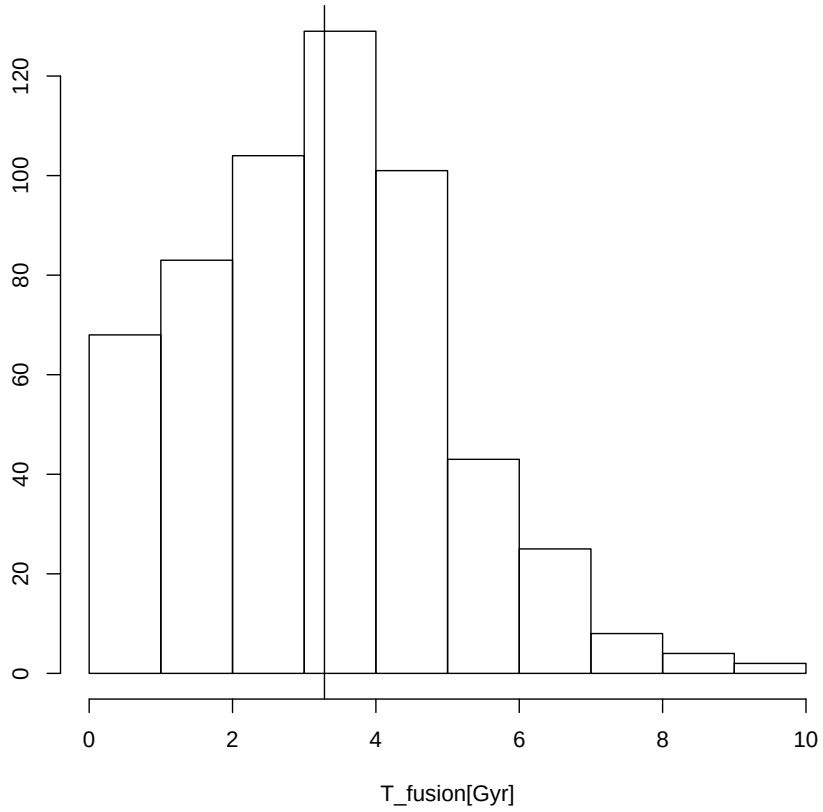

Figure 3.4: Distribution of merger times of the merging clusters sample. In black line is shown the result found by *Pinkney et al. 1996*[69].

3.2.2 Estimation of features for the merging clusters sample.

As explained in the chapter 2, one of the most important processes when working with machine learning algorithms is to find a set of features that serve as predictors when making the classification.

Taking into account the nature of the problem, we can intuit certain properties that should correlate with the interactions between systems. For example, in a merger process like the one we are studying, it is expected that the clusters will have a higher degree of substructures, a non-Gaussian velocity distribution, an asymmetric angular distribution, etc. Considering these properties we have selected the following features:

- The *Dressler & Shectman* test [25]: This test was designed to identify the substructures present in galaxy clusters. To this end, they define the following statistics for each galaxy of the cluster:

$$\delta^2 = \left(\frac{11}{\sigma^2}\right)\left[(\bar{v}_{local} - \bar{v})^2 + (\sigma_{local} - \sigma)^2\right]$$

where σ is the standard deviation of the velocity distribution of the entire clusters, σ_{local} is the standard deviation of the velocity distribution of the $n_{vec} = 10$ galaxies closer to the galaxy we are studying, $\bar{v}$ is the the average radial velocity of the entire cluster and $\bar{v}_{local}$ is the average velocity of the $n_{vec} = 10$ galaxies closer to the galaxy that we are analysing.

Then you can define the Δ statistic for the cluster by

$$\Delta = \sum_{i=1}^{ngal} \delta_i \tag{3.1}$$

where $ngal$ is the number of galaxies in the cluster. If the distribution of velocities is Gaussian and the variations in local velocities are random then $\Delta \approx ngal$, however, Δ can be very different from $ngal$ even without real substructures. This is why it is necessary to perform Monte Carlo simulations in order to calibrate the Δ statistic for each system individually. In each Monte Carlo simulation we randomly distribute the velocities of the galaxies, in order to obtain the same velocity distribution but eliminating the possible substructures.

It is convenient to define the p value of a cluster using

$$p = \frac{N(\Delta_{MC} > \Delta)}{N_{MC}} \tag{3.2}$$

where $N(\Delta_{MC} > \Delta)$ is the number of Monte Carlo realisations in which a value of Δ greater than the real is measured and N_{MC} is the number of Monte Carlo simulations. And so, the smaller is the value of p, the more likely it is that the cluster present substructures.

Finally, we are going to use the Δ value and the p value of each cluster as features for our training.

- Iterative *Dressler-Shectman* test: Taking into account that the statistics δ of each galaxy correlated with the fact that the galaxy belongs or not to a substructure within the cluster, we develop an iterative algorithm. In a first step this algorithm obtains the value of δ for each galaxy applying the *Dressler-Shectman* test with $n_{vec} = 0.2 * ngal$. Then, those galaxies with $\delta_i < \bar{\delta} * 0.7$ are eliminated, leaving only the galaxies with the highest probability of being in a substructure. In the next step we apply again the *Dressler-Shectman* test but with the new sample of galaxies. Finally we say that the algorithm converges if the number of galaxies between 2 consecutive steps is equal or, in other words, that when applying the test all galaxies have $\delta_i > 0.7 * \bar{\delta}$.

  In figure 3.5 is shown, as an example of the application of the iterative test, the angular positions of the galaxies of a cluster that had a convergence in 4 steps. The size of each circle is proportional to the δ statistic of each galaxy.

  Finally we use as feature for our training the number of iterations that the algorithm made until converging.

- Bearing in mind that it is expected that in a merging cluster the radial velocities distribution will deviate from a normal distribution, we use as features different methods that measure the deviation of a given distribution with respect to a Gaussian one. These methods are the *Anderson-Darling* test, the *Cramer-von Mises* test, the *Kolmogorov-Smirnov* test and the *Shapiro-Wilk* test. The 3 first tests are part of the nortest package [41] of R, while the *Shapiro-Wilk* test is part of the R-package stats.

- Finally, we also added as features different properties of the clusters, such as the magnitude in the R-band of the SDSS survey, the colour index G-R and the number of member galaxies of each cluster.

3.2.3 Training different machine learning models.

Once we have built a representative training set for the problem we want to work on, the next step is to look for correlations between the features and the cluster dynamical classifi-

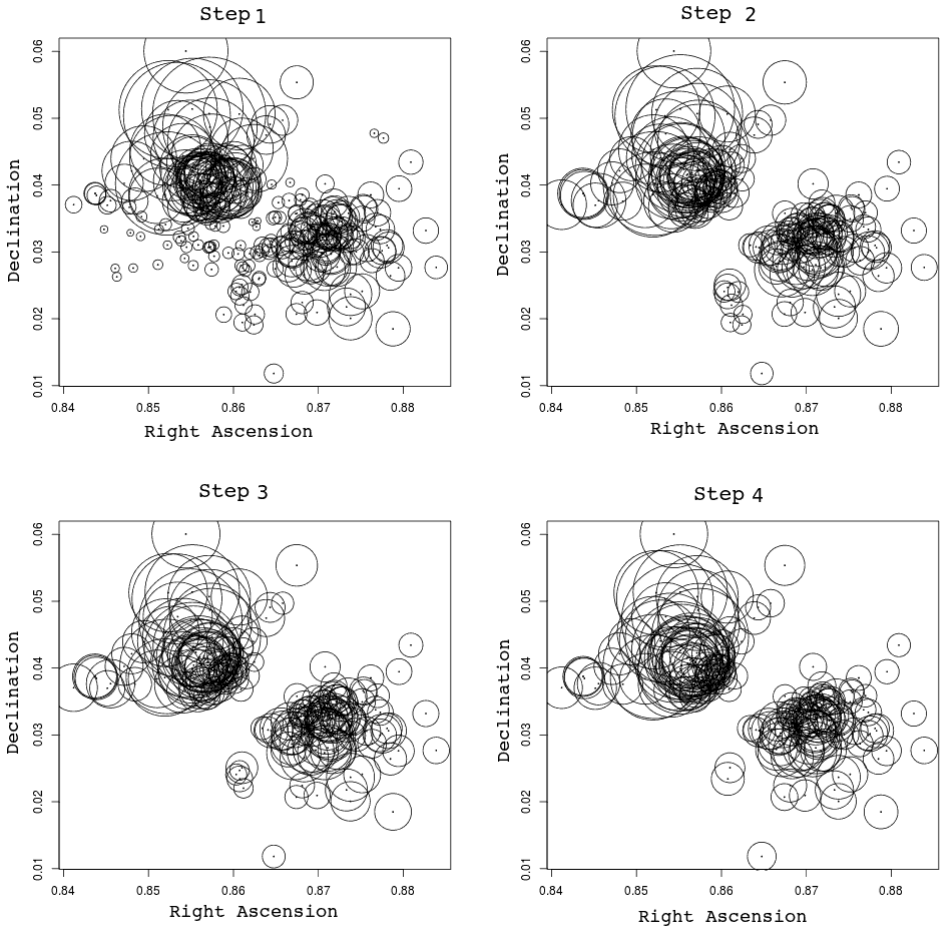

Figure 3.5: From left to right and from top to bottom, is shown the result of the application of the iterative Dressler-Shectman test.

cation. It is important to note that at this stage of the problem, the most efficient procedure would be to investigate if there is any property that separates the merging clusters in a natural way without needing to trained any machine learning algorithm [22]. In figures 3.6 and 3.7 we show the distributions of the features. In red are shown the merging clusters, while the non-merging clusters are plotted in black. As can be seen, none of these parameters clearly separates both classes, showing the need to apply machine learning algorithms that seek for deeper correlations.

As explained in chapter 2 there are different machine learning algorithms, so when it comes to attacking a problem it is necessary to try several models and compare their

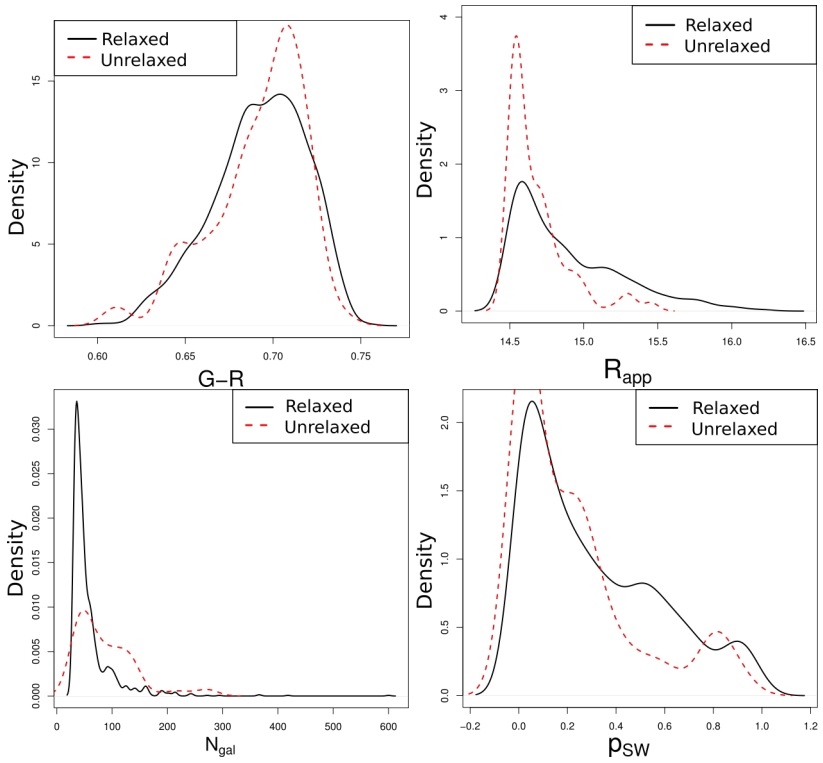

Figure 3.6: Distribution of features for interacting (red) and non-interacting (black) systems.

results. In this work we train 3 different algorithms: Logistic Regression [14, 21], *Support Vector Machine* [20, 64] and *Random Forest* [12, 57].

In order to compare the performance of each algorithm we constructed, from the total training set, 8 independent sub-sets of clusters. Then we train each algorithm using 7 of these sub-sets. Once the algorithm is trained, we use it to predict the dynamic state of the clusters that belong to the remaining sub-set. Since we know the real dynamic state of these systems, we can evaluate the performance of each one of the methods.

In figure 3.8 we shown the ROC curve, which is used to compare the performance of the different methods. It can be observed that the algorithm with the best performance is the *Random Forest*, reaching a true positives rate (TPR) of ≈ 0.85 maintaining the false positive rate (FPR) below 0.1.

For each cluster the *Random Forest* returns a statistic ν_{RF} between 0 and 1 related to the probability of the system of being in an interaction process. Then we say that a group

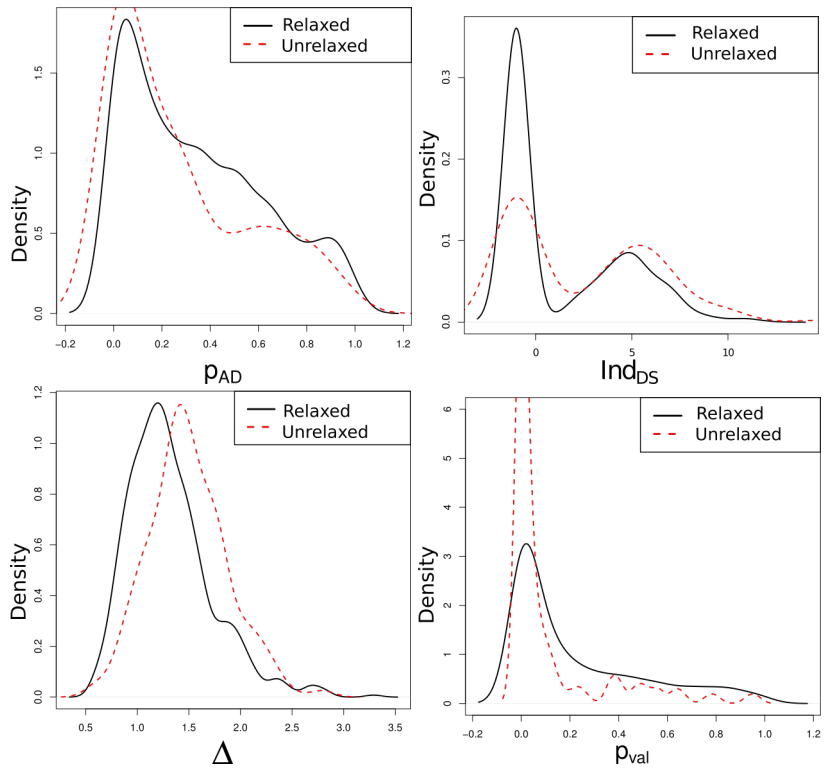

Figure 3.7: Distribution of features for interacting (red) and non-interacting (black) systems.

is identified as merging one by our algorithm if $\nu_{RF} > \nu_{crit}$, where ν_{crit} is a critical value to be defined. It is important to note that we can vary this limit to achieve a final sample as pure as possible (i.e. without non-interacting groups) but without losing completeness. In order to find the best value for ν_{crit} we evaluate the true positives rate, the false positives rate, the effectiveness (number of identified true–mergers divided by the total number of identified mergers) and the normalized length (number of identified mergers divided by the maximum number of identified mergers for the different thresholds) in function of ν_{crit} (See figure 3.9).

Finally we selected a value of $\nu_{crit} = 0.3$, assuring a final sample with a low rate of false detections. It is important to note that this ν_{crit} has a serious impact on the number of cluster that we recover, however we prefer to lose completeness but maintain a high purity in the final sample.

The figure 3.10 shows the true positive rate (TPR) and the effectiveness of the method

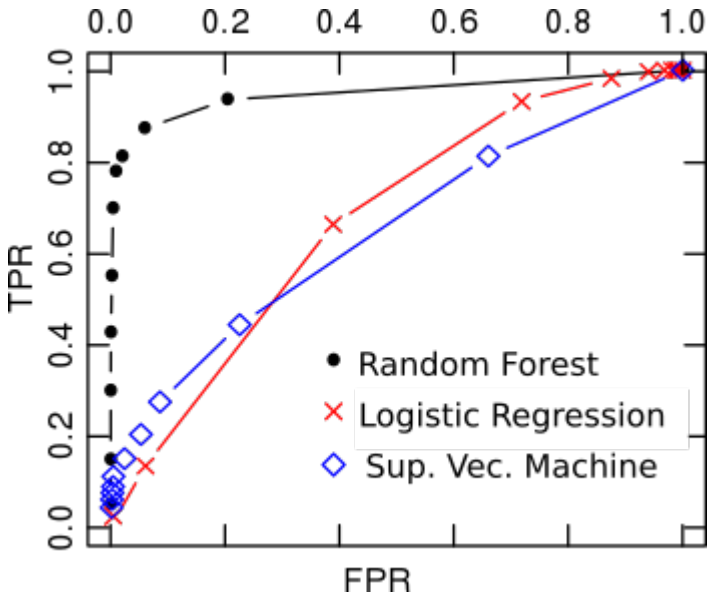

Figure 3.8: Comparative ROC curve for the different machine learning methods used for the classification of the clusters.

as a function of the number of classified merging clusters. It can be seen that as the number of merging clusters found by our method is increased (because the ν_{crit} is reduced) the true positives rate increases, that is, the number of well-classified merging systems increase. On the other hand, by increasing the size of the final sample, the effectiveness of our algorithm decrease, which means that although we increase the number of well-classified merging systems, the number of objects classified as interacting clusters but that in reality are not in a merger process, also increase. We also show in black line, the number of clusters and the expected rates for a $\nu_{crit} = 0.3$.

It is important to clarify that this algorithm classifies a cluster, but does not identify the substructures that are in interaction nor the galaxies that belong to these substructures.

To this end, and taking into account the same set of clusters used previously, we set up a new training set and a new machine learning algorithm with the goal of estimating the probability of a galaxy to belong or not to a colliding substructure. In order to do this we discriminate those galaxies that belong to merging substructures from those that do not belong to a colliding group. For this new algorithm we will use as features the following properties of the galaxies:

- Apparent magnitude of the galaxy in the R-band

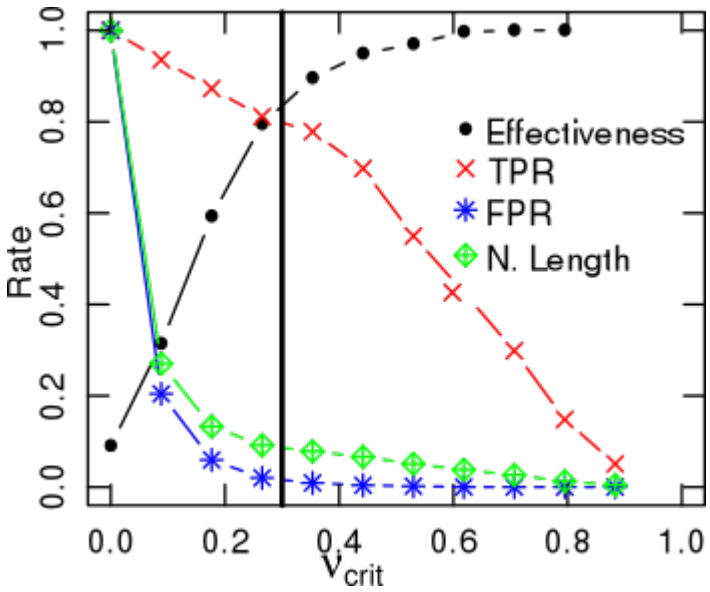

Figure 3.9: True positives rate (red crosses), false positive rate (blue asterisks), effectiveness (black dots) and normalized length of the final sample (green diamonds) as a function of ν_{crit}.

- G-R colour index of the galaxy.
- G-R colour index of the cluster to which the galaxy belongs.
- δ statistic of each galaxy, estimated by the *Dressler-Shectman* [25] test.
- Δ statistic of the cluster to which the galaxy belongs. Calculated by the *Dressler-Shectman* test [25].
- *Shapiro-Wilk* normality test applied to the redial velocity distribution of the 10 galaxies closer to the galaxy under study.

Using these features we train different machine learning algorithms. In figure 3.11 it can be seen the ROC curve for a *Random Forest*. This algorithm was the one that obtained the best performance when classifying the galaxies. This technique returns an statistic that is proportional to the probability of the galaxy to belong to the interacting substructures, As in the case of the clusters, we must choose a cut $\nu_{crit,gal}$ in that probability so that those galaxies with $\nu_{gal} > \nu_{crit,gal}$ are classified as galaxies in interacting substructures, while those galaxies with $\nu_{gal} < \nu_{crit,gal}$ no.

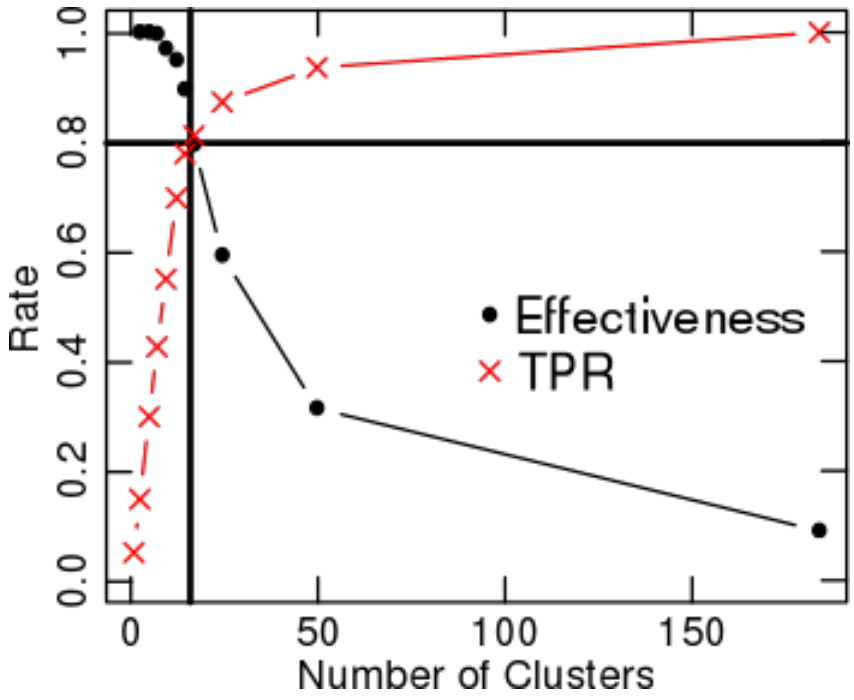

Figure 3.10: True positives rate (red crosses) and effectiveness (black points) as a function of the final size of the sample. In black line is shown the number of clusters and the expected rates for a $\nu_{crit} = 0.3$.

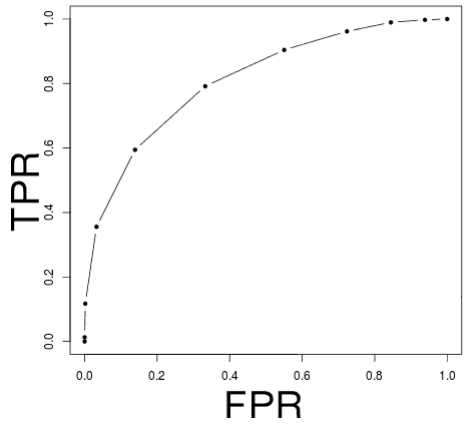

Figure 3.11: ROC curve of *Random Forest* for the galaxy classification.

As can be seen in the ROC curve, this method is not good enough to guarantee a final sample with low contamination and high completeness, that is, regardless of the cut in

$\nu_{crit,gal}$ that we choose, the final sample will have galaxies that are poorly classified. This is why we decided to use this method only to estimate the probability of each galaxy but we did not introduce any cut $\nu_{crit,gal}$. Instead of that, we make a mixture of 2 Gaussian, weighted by the probability estimated with the *Random Forest*, using the package `mclust` [81]. In this way we group those galaxies that are close to each other and that have a high probability of belong to a merged substructure.

After finding the galaxies that belong to each of the substructures that are in interaction, we can estimate different properties of these subsystems. Specifically, we estimate the angular position and the redshift of the centre of each substructure by averaging the properties of the member galaxies.

To estimate the velocity dispersion and the virial radius we use the following formulas:

$$R_{vir} = \frac{\pi n_{gal}(n_{gal}-1)}{2\sum_{i>j}^{n_{gal}} R_{ij}^{-1}} \tag{3.3}$$

$$\sigma = \frac{\sqrt{\pi}}{n_{gal}(n_{gal}-1)} \sum_{i=1}^{n_{gal}-1} \omega_i g_i \tag{3.4}$$

$$\omega_i = i(n_{gal}-1) \tag{3.5}$$

$$g_i = v_{i+1} - v_i \tag{3.6}$$

where R_{vir} is the virial radius, R_{ij} is the projected distance between the galaxy i and the galaxy j and σ is the velocity dispersion.

Taking into account these quantities and, assuming that each substructure is in dynamic equilibrium, it is also possible to estimate a dynamic mass. It is worth mentioning that, although our method seeks to identify substructures that are in interaction and that, therefore, are not in dynamic equilibrium, this is the only estimator of the mass that we can use considering only the redshift and the photometric information of the galaxies. On the other hand, as it can be seen in figures 3.12 and 3.13, comparing with the values estimated in the simulations, we were able to estimate all the properties of the interacting substructures with an acceptable margin of error.

In figure 3.14 we present the geometries of the identified merging clusters. On the x-axis we show the projected angular distance between the substructures normalized to the sum of the virial radii. Mathematically this is: $r_{\rm norm} = d_{1,2}/(r_{\rm vir\,1} + r_{\rm vir\,2})$, where

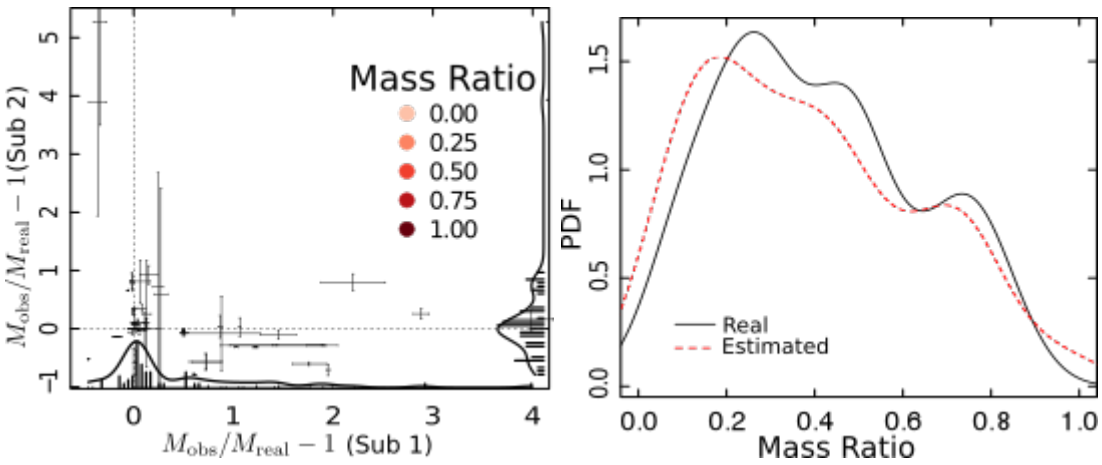

Figure 3.12: *Left*: Comparison between the estimated masses with the real ones for the interacting substructures. The mass ratio between the main and the secondary substructure can be seen in colour scale. *Right*: Comparison between the distribution of estimated and the real mass ratios of the merging substructures.

$d_{1,2}$ is the angular separation between the centres of the substructures and $r_{\mathrm{vir}\,1,2}$ are the corresponding virial radii. On the y-axis is plotted the velocity difference between the substructures, normalized to the sum of the velocities dispersion. Mathematically this is: $v_{\mathrm{norm}} = |v_{1,2}|/(\sigma_1 + \sigma_2)$, where $v_{1,2}$ is the radial velocity difference between the centres and $\sigma_{1,2}$ are the corresponding velocity dispersions.

Taking these parameters into account, 3 cases can be distinguished. Relaxed systems misclassified as merging clusters, indicated with magenta squares in figure 3.14. Well-classified mergers but in which our method fails to identify correctly the interacting substructures, marked with red crosses. Well-classified interacting groups in which our technique correctly reconstructed the substructures, plotted with black dots.

As can be seen, the false positive rate $FPR \sim 15\%$ does not vary as a function of the parameter r_{norm}. However, the well-classified clusters but in which the substructures could not be correctly identified are concentrated in values of $r_{norm} < 0.22$. It is important to clarify that, although for these systems it is not possible to find the substructures that are in the fusion process, the classification of the cluster as a merging one is not altered.

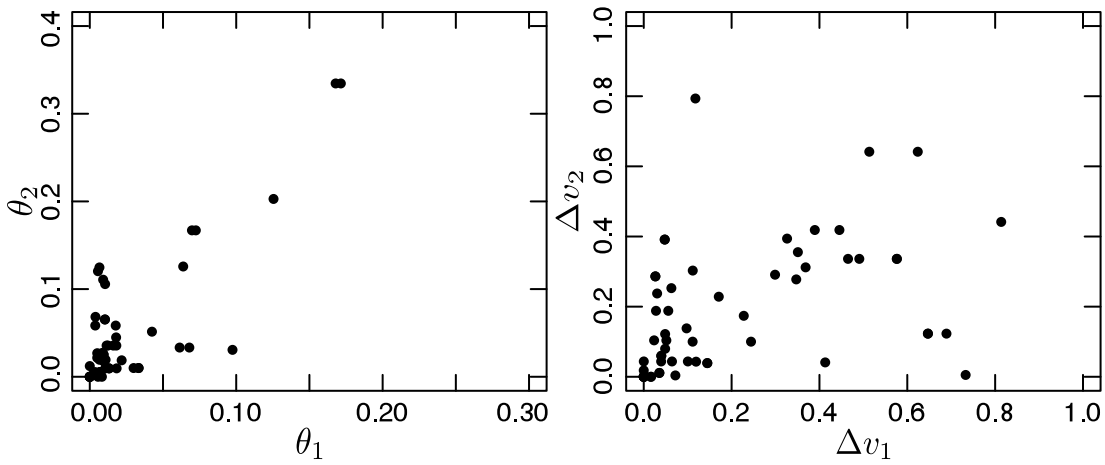

Figure 3.13: *Left*: Real and observed separation between the centres of the interacting substructures, normalized to the virial radius real. *Right*: Absolute value of the radial velocity separation between the merging substructures, normalized to the real velocity dispersion.

3.2.4 Application of the `MeSsI` algorithm to real catalogues.

In order to identify new candidates for merging systems, we applied our algorithm to 3 real galaxy clusters surveys. The WINGS catalogue [15], the HeCS survey [77] and to a cluster catalogue built by applying a `Friends-of-Friends` algorithm to the the SDSS-DR7 [1] spectroscopic data. It is worth mentioning that this technique had been used previously by *Merchán et al.* [62] applied to the galaxies of the SDSS-DR5.

Taking into account the limitations of the *Dressler & Shectman* test and that, in order to estimate the properties of the clusters we need many galaxies, we only study those systems that have more than 30 galaxies. It is also important to clarify that since in the training set the colour G-R of the SDSS was used, for the WINGS catalogue we had to estimate it using the colour b-v and applying the formulas presented by the collaboration 2df based on the results of *Fukugita et al.* [34].

In tables 3.2, 3.3 and 3.4 we present the main properties of the identified merging clusters that have $r_{norm} > 0.22$. In the first column we present the name of the system under study. From the 2^{nd} to the 5^{th} column the estimated mass and position of the main substructure are presented. From the sixth to the ninth column we present the estimated mass and position of the second substructure. Finally in the last column we present previous works in which each object was studied. Those systems that have a previous classification

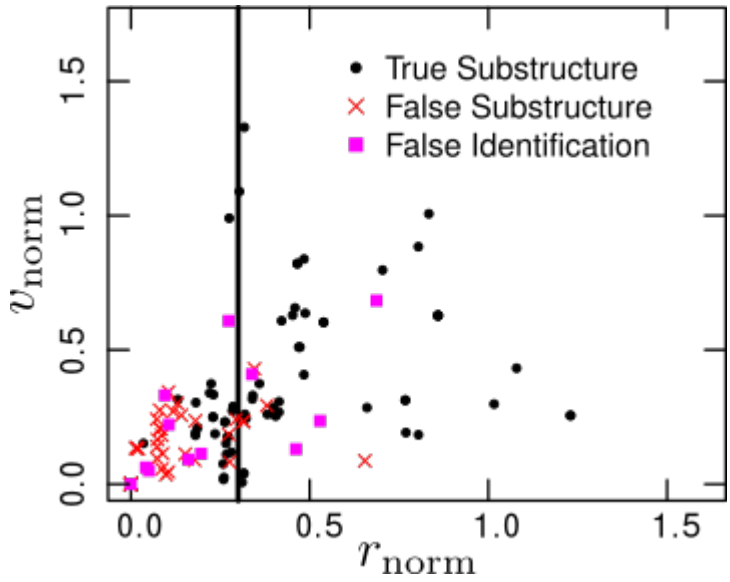

Figure 3.14: Geometry of the classified merging systems.

as a merging cluster are identified with ✓.

The errors reported for each property correspond to the standard deviation computed using 100 realisations of our algorithm for each cluster.

In addition we report the following candidates to interacting groups that have $r_{norm} < 0.22$, and so their substructure reconstruction is not reliable: A2593, A2199 ✓, A2048 ✓, A3266 ✓, A3497, A667, A1201 ✓, A267 ✓, Zw8197, A697 ✓, A750, Zw2701, Zw3146, A1246, A1302, A1413, A1682 ✓, A1763 ✓, A1902, A1918, A1930, A2009, A2034 ✓, A2069 ✓, A2111 ✓, A2219 ✓, A2050, A2259 and RXC 1504.

Name	$M_1[10^{14}M_\odot]$	RA_1 [°]	DEC_1 [°]	z_1	$M_2[10^{14}M_\odot]$	RA_2[°]	DEC_2[°]	z_2	Referencias
A1991	5.7	223.58	18.53	0.0583	2.6	223.67	18.67	0.0586	1,2,3,4,5
	$\pm$1.2	$\pm$0.05	$\pm$0.09	$\pm$0.0003	$\pm$1.02	$\pm$0.6	$\pm$0.1	$\pm$0.0003	1,2,3,4,5
Abell 1424	4.9	179.38	5.08	0.0760	5.1	179.19	5.01	0.0746	3,6.7,8
	$\pm$2.3	$\pm$0.09	$\pm$0.02	$\pm$0.0004	$\pm$1.4	$\pm$0.1	$\pm$0.04	$\pm$0.0005	3,6.7,8
A1589 ✓	5.5	190.25	18.53	0.0721	1.1	190.34	18.22	0.0716	2,22
	$\pm$0.4	$\pm$0.01	$\pm$0.02	$\pm$0.0001	$\pm$0.5	$\pm$0.01	$\pm$0.005	$\pm$0.0002	2,22
31170	25.4	255.63	34.06	0.0993	13.3	255.76	33.90	0.0989	
	$\pm$0.8	$\pm$0.05	$\pm$0.05	$\pm$0.0001	$\pm$1.3	$\pm$0.005	$\pm$0.005	$\pm$0.0002	
A2029/33	24.3	227.73	5.68	0.0796	13.4	227.81	6.13	0.0805	9,10,11,12,23
	$\pm$4.6	$\pm$0.05	$\pm$0.1	$\pm$0.0004	$\pm$1.8	$\pm$0.1	$\pm$0.2	$\pm$0.0009	9,10,11,12,23
A2069 ✓	22.6	230.99	29.94	0.1146	32	231.07	29.86	0.1146	2,3,13
	pm6.8	$\pm$0.05	$\pm$0.04	$\pm$0.0002	$\pm$10	$\pm$0.05	$\pm$0.09	$\pm$0.0004	2,3,13
Abell 2142 ✓	18.3	239.61	27.23	0.0901	11.3	239.33	27.5	0.0893	2
	$\pm$0.6	$\pm$0.005	$\pm$0.005	$\pm$0.0004	$\pm$1.8	$\pm$0.005	$\pm$0.005	$\pm$0.0001	2
A1913	5.5	216.73	16.75	0.0530	2.1	216.84	16.62	0.0533 ± 0.0013	2,3,14
	$\pm$1.1	$\pm$0.02	$\pm$0.06	$\pm$0.0004	$\pm$1.4	$\pm$0.04	$\pm$0.1	$\pm$0.0013	2,3,14
A2399	5.1	329.29	-7.81	0.0576	2.4	329.49	-7.79	0.0581	2,7,14,15,16
	$\pm$0.3	$\pm$0.02	$\pm$0.02	$\pm$0.0001	$\pm$0.3	$\pm$0.04	$\pm$0.02	$\pm$0.0002	2,7,14,15,16
A85 ✓	7.4	10.425	-9.25	0.0559	1.8	10.47	-9.51	0.0573	2,17,18,19,21
	$\pm$0.3	$\pm$0.005	$\pm$0.01	$\pm$0.0001	$\pm$1.1	$\pm$0.01	$\pm$0.05	$\pm$0.002	2,17,18,19,21
55731	1.9	244.72	24.21	0.0661	2.05	244.63	24.32	0.0656 ± 0.0005	×
	$\pm$0.4	$\pm$0.06	$\pm$0.08	$\pm$0.0004	$\pm$0.3	$\pm$0.07	$\pm$0.08	$\pm$0.0005	×
A1750 ✓	8.7	202.80	-1.89	0.0868	7.5	202.82	-1.73	0.0848	2,3,4,7,20
	$\pm$1.1	$\pm$0.02	$\pm$0.02	$\pm$0.0009	$\pm$1.6	$\pm$0.04	$\pm$0.1	$\pm$0.0016	2,3,4,7,20

Table 3.2: This table shows the candidates for interaction groups identified in the SDSS catalogue DR7 [1]. In the first column the name of the group under study is presented. From the 2^{nd} to the 5^{th} column we present the estimated mass and positions of the main substructure of each cluster. From the sixth to the ninth column we present the estimated mass and position of the second substructure. Finally in the last column we present the previous works in which each object was studied. Those clusters that have a previous classification as merging clusters are identified with ✓. References: 1 [90],2 [93], 3 [29], 4 [19], 5 [83], 6 [2], 7 [76], 8 [75], 9 [68], 10 [16], 11 [26], 12 [65], 13 [24], 14 [32], 15 [8], 16 [73], 17 [10], 18 [50], 19 [3], 20 [46], 21 [27], 22 [13], 23 [70]

Name	$M_1[10^{14}M_\odot]$	RA_1 [°]	DEC_1 [°]	z_1	$M_2[10^{14}M_\odot]$	RA_2[°]	DEC_2[°]	z_2	Referencias
A3158 ✓	57.8	55.76	-53.64	0.0644	6.9	55.37	-53.50	0.0624	2,1,3,5,6
	±7.3	±0.02	±0.02	±0.0002	±4.3	±0.01	±0.04	±0.0003	
A3809	8.6	326.68	-43.91	0.0630	6.8	327.07	-43.95	0.0639	3,7
	±0.5	±0.01	±0.005	±0.0001	±1.5	±0.05	±0.01	±0.0003	
Abell 3128 ✓	24.5	52.51	-52.58	0.0629	2.3	52.89	-52.74	0.0627	3,8,9,10
	±4	±0.01	±0.01	±0.0003	±1.1	±0.009	±0.01	±0.0003	
A3266 ✓	24.4	67.66	-61.44	0.0631	9.1	68.11	-61.32	0.0615	9,11,13
	±0.4	±0.01	±0.005	±0.0001	±1.1	±0.08	±0.04	±0.0002	
Abell 500	10.4	69.7	-22.09	0.0679	3.5	69.48	-22.27	0.0687	4,3,8
	±0.3	±0.005	±0.005	±0.0001	±4.6	±0.05	±0.05	±0.0001	
A376	18.6	41.51	36.93	0.0480	5.8	41.39	36.73	0.0474	3,12
	±2.1	±0.04	±0.02	±0.0003	±2.5	±0.1	±0.09	±0.0008	
A3490	42.6	176.43	-34.37	0.0690	63.7	176.19	-34.46	0.0698	3
	±4.2	±0.04	±0.02	±0.0002	±36.1	±0.1	±0.08	±0.0012	

Table 3.3: This table shows the candidates for interaction groups identified in the WINGS catalogue [15]. In the first column the name of the group under study is presented. From the 2^{nd} to the 5^{th} column we present the estimated mass and positions of the main substructure of each cluster. From the sixth to the ninth column we present the estimated mass and position of the second substructure. Finally in the last column we present the previous works in which each object was studied. Those clusters that have a previous classification as merging clusters are identified with ✓. References: 1 [19], 2 [68], 4 [32], 3 [73], 5 [91], 6 [49], 7 [85], 8 [52], 9 [36], 10 [79], 11 [78], 12 [55], 13 [31]

Name	$M_1[10^{14}M_\odot]$	RA_1 [°]	DEC_1 [°]	z_1	$M_2[10^{14}M_\odot]$	RA_2[°]	DEC_2[°]	z_2	Referencias
A689	32.3	129.39	14.99	0.2788	9.9	129.29	15.06	0.2791	×
	±7.9	±0.02	±0.04	±0.0002	±8.5	±0.1	±0.1	±0.0007	
Abell 1758N	59.3	203.07	50.59	0.2768	29.1	203.25	50.57	0.2783	1,4,5,6
✓	±9	±0.02	±0.01	±0.0002	±15.8	±0.02	±0.03	±0.0007	
A1758S	43	203.21	50.42	0.2742	15.8	202.96	50.39	0.2739	1,4,5,6
✓	±4	±0.005	±0.001	±0.0002	±1.6	±0.001	±0.001	±0.0001	
A2631	84.9	354.36	0.26	0.2762	77.1	354.60	0.269	0.2766	1,12
✓	±2.1	±0.001	±0.001	±0.0001	±4.3	±0.005	±0.01	±0.0002	
A2055	20.02	229.68	6.23	0.1013	3.9	229.71	6.19	0.0997	2,3
	±7.6	±0.04	±0.06	±0.0007	±2.2	±0.08	±0.1	±0.0018	
A2261	65.8	260.60	32.06	0.2253	13.4	261.10	32.27	0.2232	2
	±8.1	±0.01	±0.01	±0.0002	±3.8	±0.02	±0.02	±0.0004	
A2187	22.4	246.01	41.23	0.1829	8.9	246.27	41.22	0.1824	×
	±1.9	±0.01	±0.001	±0.0001	±0.9	±0.06	±0.005	±0.0001	
A646	16.9	126.20	47.12	0.1269	16.5	125.89	47.09	0.1263	3
	±3.9	±0.05	±0.01	±0.0002	±4.3	±0.1	±0.02	±0.0004	
A1033	26.02	157.90	35.022	0.1223	14.1	158.06	35.14	0.1222	7
	±1.9	±0.01	±0.005	±0.0001	±5.5	±0.06	±0.03	±0.0004	
Abell 1835	105	210.25	2.87	0.2516	17	210.29	2.75	0.2479	1,8,9
	±14	±0.01	±0.01	±0.0006	±25	±0.09	±0.04	±0.002	
A990	41.8	155.87	49.17	0.1410	3.9	155.84	49.74	0.1416	×
	±3.8	±0.005	±0.02	±0.0001	±9.2	±0.04	±0.2	±0.0005	
CLG 2130.4-0000	37.2	322.50	0.02	0.1355	21.4	322.34	0.03	0.1368	1
	±18.4	±0.05	±0.01	±0.0002	±11.02	±0.02	±0.04	±0.0007	
A1204	12.1	168.33	17.61	0.1703	3.4	168.54	17.52	0.1705	10
	±0.8	±0.01	±0.001	±0.0001	±1.1	±0.03	±0.01	±0.0001	
A1437	59.3	180.10	3.36	0.1324	19.7	180.36	3.30	0.1326	3,11
	±3.4	±0.01	±0.005	±0.0001	±1.8	±0.05	±0.02	±0.0004	
A795	33.5	141.00	14.17	0.1378	19.1	140.95	14.32	0.1384	
	±5.01	±0.02	±0.04	±0.0002	±11.1	±0.05	±0.1	±0.0005	
Zw3179	26.4	156.53	12.62	0.1425	6.2	156.48	12.53	0.1424	
	±3.4	±0.02	±0.01	±0.0001	±7.2	±0.08	±0.1	±0.0003	

Table 3.4: This table shows the candidates for interaction groups identified in the HeCS catalogue [77]. In the first column the name of the group under study is presented. From the 2^{nd} to the 5^{th} column we present the estimated mass and positions of the main substructure of each cluster. From the sixth to the ninth column we present the estimated mass and position of the second substructure. Finally in the last column we present the previous works in which each object was studied. Those clusters that have a previous classification as merging clusters are identified with ✓. References: 1 [93], 2 [29], 3 [32], 4 [9], 5 [71],6 [28], 7 [94], 8 [84], 9 [53], 10 [4], 11 [80], 12 [58]

It is important to note that our method was able to identify clusters that were previously classified as interacting systems by other authors, which demonstrates the effectiveness of our method. But we were also was able to identify new candidates that would expand the list of merging clusters allowing for new statistical studies.

3.2.5 The case of multiple mergers.

It is interesting to note that there are clusters whose properties suggest that are the product of the interaction between more than 2 substructures [63, 72]. In order to study these objects, we look for similar objects in the mock simulated constructed from the *Millenium* simulation. Taking into account the 8 simulated catalogues we find 27 multiple merger cases on a total of 132 merging clusters. To each one of the multiple cases we apply our algorithm allowing the mixture of Gaussians to identify more than 2 groups. We find that this technique is only able to reliably identify and estimate the properties of the main 2 substructures of each cluster, while minor substructures may appear as contamination in the major substructures or not appear linked to any group.

Continuing with the analysis of this particular type of objects, we conducted an individual study of the Abell 1758 cluster [72] that presents 4 substructures in interaction. In a first iteration, our algorithm was able to correctly classify the cluster as a merging system, but failed to correctly reconstruct the known substructures. Taking into account the previous information relative to this system, we separated the cluster into 2 components, one towards the north and one towards the south and we made an individual analysis with MeSsI to each of them. Our algorithm was able to classify each component as an interacting system and to correctly identify the 2 substructures that are interacting in each component. This example shows that any previous information about a cluster could be very important while studying the dynamical properties of a system.

3.2.6 Conclusions.

In this chapter we present the results obtained in the first stage of the doctorate, which were presented in various national and international conferences and published in the international journal *Monthly Notices of the Royal Academy* [23] and in the 57^{th} Argentinian Astronomical Association bulletin in the year 2015 [22].

In this work we introduce a new method to detect interacting galaxy systems using spectroscopic and photometric information This method is based on automatic machine learning techniques trained using a simulated catalogue from the data of the *Millenium* simulation [87] and the semi-analytical model of *Guo et al.* [42]. After comparing the performance of different methods, we found that the algorithm with the best performance is the *Random Forest* with which we obtain a final sample with high purity and with an acceptable completeness. It is important to remind that we used as *features* different properties obtained from the spectro-photometric data of each catalogue.

In addition, using a mixture of Gaussians techniques, we were able to reconstruct the merging substructures and to estimate its main properties, such as the position, the virial radius, the mass and the velocity dispersion.

Finally, we apply our method to 3 real galaxy cluster catalogues: SDSS-DR7, WINGS and HeCS. We found 12, 4 and 16 candidates for merging systems in which we were able to reconstruct the interacting substructures. In addition, we reported another 29 candidates in which, due to its geometry, we were not able to identify its substructures.

It is important to note that 19 of these candidates were previously reported as merging systems, which reinforces the validity of our method. We also reported for the first time 40 new candidates for interacting groups that must be confirmed through individual studies.

Finally we develop a web interface (http://200.16.29.98/martin/merclust) in which any astronomer can upload the data of their galaxy cluster and analyse them using our algorithm for free or download the source code to use it locally.

Bibliography

[1] K. N. Abazajian, J. K. Adelman-McCarthy, M. A. Agueros, S. S. Allam, C. Allende Prieto, D. An, K. S. J. Anderson, S. F. Anderson, J. Annis, N. A. Bahcall, and et al. The Seventh Data Release of the Sloan Digital Sky Survey,. *Astrophys. J. Suppl. Ser.*, 182:543, June 2009. doi: 10.1088/0067-0049/182/2/543.

[2] M. H. Abdullah, G. B. Ali, H. A. Ismail, and M. A. Rassem. Studying the dynamical properties of 20 nearby galaxy clusters. *MNRAS*, 416:2027–2040, September 2011. doi: 10.1111/j.1365-2966.2011.19178.x.

[3] J. A. L. Aguerri and R. Sánchez-Janssen. A study of catalogued nearby galaxy clusters in the SDSS-DR4. II. Cluster substructure. *A&A*, 521:A28, October 2010. doi: 10.1051/0004-6361/200913901.

[4] F. Andrade-Santos, G. B. Lima Neto, and T. F. Laganá. A New Method to Quantify X-Ray Substructures in Clusters of Galaxies. *ApJ*, 746:139, February 2012. doi: 10.1088/0004-637X/746/2/139.

[5] G. W. Angus, B. Famaey, and H. S. Zhao. Can MOND take a bullet? Analytical comparisons of three versions of MOND beyond spherical symmetry,. *MNRAS*, 371: 138–146, September 2006. doi: 10.1111/j.1365-2966.2006.10668.x.

[6] R. Barrena, A. Biviano, M. Ramella, E. E. Falco, and S. Seitz. The dynamical status of the cluster of galaxies 1E0657-56,. *aap*, 386:816–828, May 2002. doi: 10.1051/0004-6361:20020244.

[7] B. Benson, D. M. Wittman, N. Golovich, M. James Jee, R. J. van Weeren, and W. A. Dawson. MC^2: A Deeper Look at ZwCl 2341.1+0000 with Bayesian

Galaxy Clustering and Weak Lensing Analyses. *ApJ*, 841:7, May 2017. doi: 10.3847/1538-4357/aa6d66.

[8] H. Böhringer et al. Substructure of the galaxy clusters in the REXCESS sample: observed statistics and comparison to numerical simulations. *A&A*, 514:A32, May 2010. doi: 10.1051/0004-6361/200913911.

[9] W. Boschin, M. Girardi, R. Barrena, and M. Nonino. Abell 1758N from an optical point of view: new insights on a merging cluster with diffuse radio emission. *A&A*, 540:A43, April 2012. doi: 10.1051/0004-6361/201118076.

[10] G. Boué, F. Durret, C. Adami, G. A. Mamon, O. Ilbert, and V. Cayatte. An optical view of the filament region of Abell 85. *A&A*, 489:11–22, October 2008. doi: 10.1051/0004-6361:20078972.

[11] M. Bradac, S. W. Allen, T. Treu, H. Ebeling, R. Massey, R. G. Morris, A. von der Linden, and D. Applegate. Revealing the Properties of Dark Matter in the Merging Cluster MACS J0025.4-1222,. *ApJ*, 687:959–967, November 2008. doi: 10.1086/591246.

[12] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, Oct 2001. ISSN 1573-0565. doi: 10.1023/A:1010933404324. URL https://doi.org/10.1023/A:1010933404324.

[13] J. O. Burns, G. Rhee, F. N. Owen, and J. Pinkney. Clumped X-ray emission around radio galaxies in Abell clusters. *ApJ*, 423:94–115, March 1994. doi: 10.1086/173792.

[14] Angelo Canty and B. D. Ripley. *boot: Bootstrap R Functions*, 2015. R package version 1.3-15.

[15] A. Cava et al. WINGS-SPE Spectroscopy in the WIde-field Nearby Galaxy-cluster Survey. *A&A*, 495:707–719, March 2009. doi: 10.1051/0004-6361:200810997.

[16] T. E. Clarke, E. L. Blanton, and C. L. Sarazin. The Complex Cooling Core of A2029: Radio and X-Ray Interactions. *ApJ*, 616:178–191, November 2004. doi: 10.1086/424911.

[17] D. Clowe, M. Bradac, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky. A Direct Empirical Proof of the Existence of Dark Matter,. *ApJL*, 648: L109–L113, September 2006. doi: 10.1086/508162.

[18] D. Clowe, M. Markevitch, M. Bradac, A. H. Gonzalez, S. M. Chung, R. Massey, and D. Zaritsky. On Dark Peaks and Missing Mass: A Weak-lensing Mass Reconstruction of the Merging Cluster System A520,. *ApJ*, 758:128, October 2012. doi: 10.1088/0004-637X/758/2/128.

[19] S. A. Cohen, R. C. Hickox, G. A. Wegner, M. Einasto, and J. Vennik. Star Formation and Substructure in Galaxy Clusters. *ApJ*, 783:136, March 2014. doi: 10.1088/0004-637X/783/2/136.

[20] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995. ISSN 0885-6125. doi: 10.1023/A:1022627411411. URL http://dx.doi.org/10.1023/A%3A1022627411411.

[21] A. C. Davison and D. V. Hinkley. *Bootstrap Methods and Their Applications*. CUP, Cambridge, 1997. URL http://statwww.epfl.ch/davison/BMA/. ISBN 0-521-57391-2.

[22] M. de los Rios, M. J. Domínguez, and D. Paz. Construcción de un catálogo de cúmulos de galaxias en proceso de colisión. *Boletin de la Asociacion Argentina de Astronomia La Plata Argentina*, 57:22–24, August 2015.

[23] M. de los Rios, M. J. Domínguez R., D. Paz, and M. Merchán. The MeSsI (merging systems identification) algorithm and catalogue. *MNRAS*, 458:226–232, May 2016. doi: 10.1093/mnras/stw215.

[24] A. Drabent, M. Hoeft, R. F. Pizzo, A. Bonafede, R. J. van Weeren, and U. Klein. Diffuse radio emission in the complex merging galaxy cluster Abell2069. *A&A*, 575:A8, March 2015. doi: 10.1051/0004-6361/201424828.

[25] A. Dressler and S. A. Shectman. Evidence for substructure in rich clusters of galaxies from radial-velocity measurements,. *AJ*, 95:985–995, April 1988. doi: 10.1086/114694.

[26] R. J. H. Dunn, A. C. Fabian, and G. B. Taylor. Radio bubbles in clusters of galaxies. *MNRAS*, 364:1343–1353, December 2005. doi: 10.1111/j.1365-2966.2005.09673.x.

[27] F. Durret, G. Lima Neto, and W. Forman. An XMM-Newton view of the cluster of galaxies Abell 85. *A&A*, 432:809–821, March 2005. doi: 10.1051/0004-6361:20041666.

[28] F. Durret, T. F. Laganá, and M. Haider. The merging cluster Abell 1758 revisited: multi-wavelength observations and numerical simulations. *A&A*, 529:A38, May 2011. doi: 10.1051/0004-6361/201015978.

[29] M. Einasto et al. Multimodality in galaxy clusters from SDSS DR8: substructure and velocity distribution. *A&A*, 540:A123, April 2012. doi: 10.1051/0004-6361/201118697.

[30] G. R. Farrar and R. A. Rosen. A New Force in the Dark Sector?,. *Physical Review Letters*, 98(17):171302, April 2007. doi: 10.1103/PhysRevLett.98.171302.

[31] A. Finoguenov, M. J. Henriksen, F. Miniati, U. G. Briel, and C. Jones. A Puzzling Merger in A3266: The Hydrodynamic Picture from XMM-Newton. *ApJ*, 643:790–796, June 2006. doi: 10.1086/503285.

[32] P. Flin and J. Krywult. Substructures in Abell clusters of galaxies. *A&A*, 450:9–14, April 2006. doi: 10.1051/0004-6361:20041635.

[33] J. E. Forero-Romero, S. Gottlober, and G. Yepes. Bullet Clusters in the MARENOSTRUM Universe,. *ApJ*, 725:598–604, December 2010. doi: 10.1088/0004-637X/725/1/598.

[34] M. Fukugita, K. Shimasaku, and T. Ichikawa. Galaxy Colors in Various Photometric Band Systems. *pasp*, 107:945, October 1995. doi: 10.1086/133643.

[35] S. Ghigna, B. Moore, F. Governato, G. Lake, T. Quinn, and J. Stadel. Dark matter haloes within clusters,. *MNRAS*, 300:146–162, October 1998. doi: 10.1046/j.1365-8711.1998.01918.x.

[36] M. Girardi, E. Escalera, D. Fadda, G. Giuricin, F. Mardirossian, and M. Mezzetti. Optical Substructures in 48 Galaxy Clusters: New Insights from a Multiscale Analysis. *ApJ*, 482:41–62, June 1997.

[37] N. Golovich, W. A. Dawson, D. Wittman, G. Ogrean, R. van Weeren, and A. Bonafede. MC^2: Dynamical Analysis of the Merging Galaxy Cluster MACS J1149.5+2223. *ApJ*, 831:110, November 2016. doi: 10.3847/0004-637X/831/1/110.

[38] N. Golovich, W. A. Dawson, D. M. Wittman, M. J. Jee, B. Benson, B. Lemaux, R. J. van Weeren, F. Andrade-Santos, D. Sobral, F. de Gasperin, M. Bruggen, M. Bradac, K. Finner, and A. Peter. Merging Cluster Collaboration: Optical and Spectroscopic Survey of a Radio-Selected Sample of Twenty Nine Merging Galaxy Clusters. *ArXiv e-prints*, November 2017.

[39] Nathan Golovich, Reinout J. van Weeren, William A. Dawson, M. James Jee, and David Wittman. MC^2: Multiwavelength and Dynamical Analysis of the Merging Galaxy Cluster ZwCl 0008.8+5215: An Older and Less Massive Bullet Cluster. *ApJ*, 838:110, April 2017. doi: 10.3847/1538-4357/aa667f.

[40] S. Gottlober and G. Yepes. Shape, Spin, and Baryon Fraction of Clusters in the MareNostrum Universe,. *ApJ*, 664:117–122, July 2007. doi: 10.1086/517907.

[41] Juergen Gross and Uwe Ligges. *nortest: Tests for Normality*, 2015. URL `http://CRAN.R-project.org/package=nortest`. R package version 1.0-3.

[42] Q. Guo et al. From dwarf spheroidals to cD galaxies: simulating the galaxy population in a ΛCDM cosmology. *MNRAS*, 413:101–131, May 2011. doi: 10.1111/j.1365-2966.2010.18114.x.

[43] D. Harvey, R. Massey, T. Kitching, A. Taylor, and E. Tittley. The nongravitational interactions of dark matter in colliding galaxy clusters,. *Science*, 347:1462–1465, March 2015. doi: 10.1126/science.1261381.

[44] E. Hayashi and S. D. M. White. How rare is the bullet cluster?,. *MNRAS*, 370: L38–L41, July 2006. doi: 10.1111/j.1745-3933.2006.00184.x.

[45] J. P. Huchra and M. J. Geller. Groups of galaxies. I - Nearby groups,. *ApJ*, 257: 423–437, June 1982. doi: 10.1086/160000.

[46] H. S. Hwang and M. G. Lee. Galaxy activity in merging binary galaxy clusters. *MNRAS*, 397:2111–2122, August 2009. doi: 10.1111/j.1365-2966.2009.15100.x.

[47] N. S Israel and J. W Moffat. The Train Wreck Cluster Abell 520 and the Bullet Cluster 1E0657-558 in a Generalized Theory of Gravitation,. *ArXiv e-prints*, April 2016.

[48] M. J. Jee, H. Hoekstra, A. Mahdavi, and A. Babul. Hubble Space Telescope/Advanced Camera for Surveys Confirmation of the Dark Substructure in A520,. *ArXiv e-prints*, January 2014.

[49] M. Johnston-Hollitt, M. Sato, J. A. Gill, M. C. Fleenor, and A.-M. Brick. Radio observations of the Horologium-Reticulum supercluster - I. A3158: excess star-forming galaxies in a merging cluster? *MNRAS*, 390:289–303, October 2008. doi: 10.1111/j.1365-2966.2008.13730.x.

[50] J. C. Kempner, C. L. Sarazin, and P. M. Ricker. Chandra Observations of A85: Merger of the South Subcluster. *ApJ*, 579:236–246, November 2002. doi: 10.1086/342748.

[51] J. Kim, C. Park, J. R. Gott, III, and J. Dubinski. The Horizon Run N-Body Simulation: Baryon Acoustic Oscillations and Topology of Large-scale Structure of the Universe,. *ApJ*, 701:1547-1559, August 2009. doi: 10.1088/0004-637X/701/2/1547.

[52] V. Kolokotronis, S. Basilakos, M. Plionis, and I. Georgantopoulos. Searching for cluster substructure using APM and ROSAT data. *MNRAS*, 320:49–60, January 2001. doi: 10.1046/j.1365-8711.2001.03924.x.

[53] P. M. Korngut et al. MUSTANG High Angular Resolution Sunyaev-Zel'dovich Effect Imaging of Substructure in Four Galaxy Clusters. *ApJ*, 734:10, June 2011. doi: 10.1088/0004-637X/734/1/10.

[54] C. Lage and G. R. Farrar. The bullet cluster is not a cosmological anomaly,. *JCAP*, 2:038, February 2015. doi: 10.1088/1475-7516/2015/02/038.

[55] K. Lakhchaura and K. P. Singh. Dynamics of 10 Clusters of Galaxies with Substructures. *AJ*, 147:156, June 2014. doi: 10.1088/0004-6256/147/6/156.

[56] J. Lee and E. Komatsu. Bullet Cluster: A Challenge to Λ,CDM Cosmology,. *ApJ*, 718:60–65, July 2010. doi: 10.1088/0004-637X/718/1/60.

[57] Andy Liaw and Matthew Wiener. Classification and regression by randomforest. *R News*, 2(3):18–22, 2002. URL http://CRAN.R-project.org/doc/Rnews/.

[58] A. W. Mann and H. Ebeling. X-ray-optical classification of cluster mergers and the evolution of the cluster merger fraction. *MNRAS*, 420:2120–2138, March 2012. doi: 10.1111/j.1365-2966.2011.20170.x.

[59] M. Markevitch. Chandra Observation of the Most Interesting Cluster in the Universe,. In A. Wilson, editor, *The X-ray Universe 2005*, volume 604 of *ESA Special Publication*, page 723, January 2006.

[60] M. Markevitch, A. H. Gonzalez, D. Clowe, A. Vikhlinin, W. Forman, C. Jones, S. Murray, and W. Tucker. Direct Constraints on the Dark Matter Self-Interaction Cross Section from the Merging Galaxy Cluster 1E 0657-56,. *ApJ*, 606:819–824, May 2004. doi: 10.1086/383178.

[61] C. Mastropietro and A. Burkert. Simulating the Bullet Cluster,. *MNRAS*, 389:967–988, September 2008. doi: 10.1111/j.1365-2966.2008.13626.x.

[62] M. Merchán and A. Zandivarez. Galaxy groups in the 2dF Galaxy Redshift Survey: the catalogue. *MNRAS*, 335:216–222, September 2002. doi: 10.1046/j.1365-8711.2002.05623.x.

[63] J. Merten, D. Coe, R. Dupke, R. Massey, A. Zitrin, E. S. Cypriano, N. Okabe, B. Frye, F. G. Braglia, Y. Jimenez-Teja, N. Benitez, T. Broadhurst, J. Rhodes, M. Meneghetti, L. A. Moustakas, L. Sodre, Jr., J. Krick, and J. N. Bregman. Creation of cosmic structure in the complex galaxy cluster merger Abell 2744,. *MNRAS*, 417:333–347, October 2011. doi: 10.1111/j.1365-2966.2011.19266.x.

[64] David Meyer, Evgenia Dimitriadou, Kurt Hornik, Andreas Weingessel, and Friedrich Leisch. *e1071, TU Wien*, 2014. URL http://CRAN.R-project.org/package=e1071. R package version 1.6-4.

[65] E. T. Million and S. W. Allen. Chandra measurements of non-thermal-like X-ray emission from massive, merging, radio halo clusters. *MNRAS*, 399:1307–1327, November 2009. doi: 10.1111/j.1365-2966.2009.15359.x.

[66] M. Milosavljevic, J. Koda, D. Nagai, E. Nakar, and P. R. Shapiro. The Cluster-Merger Shock in 1E 0657-56: Faster than a Speeding Bullet?,. *ApJL*, 661:L131–L134, June 2007. doi: 10.1086/518960.

[67] Houjun Mo, Frank van den Bosch, and Simon White. *Galaxy Formation and Evolution*. Cambridge University Press, 2010. doi: 10.1017/CBO9780511807244.

[68] V. Parekh, K. van der Heyden, C. Ferrari, G. Angus, and B. Holwerda. Morphology parameters: substructure identification in X-ray galaxy clusters. *A&A*, 575:A127, March 2015. doi: 10.1051/0004-6361/201424123.

[69] J. Pinkney, K. Roettiger, J. O. Burns, and C. M. Bird. Evaluation of Statistical Tests for Substructure in Clusters of Galaxies,. *Astrophys. J. Suppl. Ser.*, 104:1, May 1996. doi: 10.1086/192290.

[70] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. Balbi, A. J. Banday, and et al. Planck intermediate results. VIII. Filaments between interacting clusters. *A&A*, 550:A134, February 2013. doi: 10.1051/0004-6361/201220194.

[71] B. Ragozzine, D. Clowe, M. Markevitch, A. H. Gonzalez, and M. Bradač. Weak-lensing Results for the Merging Cluster A1758. *ApJ*, 744:94, January 2012. doi: 10.1088/0004-637X/744/2/94.

[72] B. Ragozzine, D. Clowe, M. Markevitch, A. H. Gonzalez, and M. Bradač. Weak-lensing Results for the Merging Cluster A1758. *ApJ*, 744:94, January 2012. doi: 10.1088/0004-637X/744/2/94.

[73] M. Ramella et al. Substructures in WINGS clusters. *A&A*, 470:39–51, July 2007. doi: 10.1051/0004-6361:20077245.

[74] S. W. Randall, M. Markevitch, D. Clowe, A. H. Gonzalez, and M. Bradac. Constraints on the Self-Interaction Cross Section of Dark Matter from Numerical Simulations of the Merging Galaxy Cluster 1E 0657-56,. *ApJ*, 679:1173–1180, June 2008. doi: 10.1086/587859.

[75] G. F. R. N. Rhee, M. P. van Haarlem, and P. Katgert. Substructure in Abell clusters. *A&A*, 246:301–312, June 1991.

[76] K. Rines and A. Diaferio. CIRS: Cluster Infall Regions in the Sloan Digital Sky Survey. I. Infall Patterns and Mass Profiles. *AJ*, 132:1275–1297, September 2006. doi: 10.1086/506017.

[77] K. Rines, M. J. Geller, A. Diaferio, and M. J. Kurtz. Measuring the Ultimate Halo Mass of Galaxy Clusters: Redshifts and Mass Profiles from the Hectospec Cluster Survey (HeCS). *ApJ*, 767:15, April 2013. doi: 10.1088/0004-637X/767/1/15.

[78] K. Roettiger and R. Flores. A Prediction of Observable Rotation in the Intracluster Medium of Abell 3266. *ApJ*, 538:92–97, July 2000. doi: 10.1086/309132.

[79] J. A. Rose, A. E. Gaba, W. A. Christiansen, D. S. Davis, N. Caldwell, R. W. Hunstead, and M. Johnston-Hollitt. Multiple Merging Events in the Double Cluster A3128/A3125. *AJ*, 123:1216–1246, March 2002. doi: 10.1086/339183.

[80] S. Schindler. X-ray and optical observations of three clusters of galaxies: Abell 901, Abell 1437, and Abell 3570. *A&As*, 142:433–441, March 2000. doi: 10.1051/aas:2000161.

[81] Luca Scrucca, Michael Fop, Thomas Brendan Murphy, and Adrian E. Raftery. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal*, 8(1):205–233, 2016.

[82] U. Seljak and M. Zaldarriaga. A Line-of-Sight Integration Approach to Cosmic Microwave Background Anisotropies,. *ApJ*, 469:437, October 1996. doi: 10.1086/177793.

[83] M. Sharma et al. A Chandra X-Ray Observation of A1991: The Late Stages of Infall? *ApJ*, 613:180–188, September 2004. doi: 10.1086/422866.

[84] G. P. Smith and J. E. Taylor. Connecting Substructure in Galaxy Cluster Cores at z = 0.2 with Cluster Assembly Histories. *ApJL*, 682:L73–L76, August 2008. doi: 10.1086/591271.

[85] J. M. Solanes, E. Salvador-Solé, and G. González-Casado. Substructure in the ENACS clusters. *A&A*, 343:733–739, March 1999.

[86] D. N. Spergel, L. Verde, H. V. Peiris, E. Komatsu, M. R. Nolta, C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters,. *Astrophys. J. Suppl. Ser.*, 148:175–194, September 2003. doi: 10.1086/377226.

[87] V. Springel and G. R. Farrar. The speed of the 'bullet' in the merging galaxy cluster 1E0657-56,. *MNRAS*, 380:911–925, September 2007. doi: 10.1111/j.1365-2966.2007.12159.x.

[88] V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, J. A. Peacock, S. Cole, P. Thomas, H. Couchman, A. Evrard, J. Colberg, and F. Pearce. Simulations of the formation, evolution and clustering of galaxies and quasars,. *Nature*, 435:629–636, June 2005. doi: 10.1038/nature03597.

[89] D. Tweed, J. Devriendt, J. Blaizot, S. Colombi, and A. Slyz. Building merger trees from cosmological N-body simulations. Towards improving galaxy formation models using subhaloes,. *aap*, 506:647–660, November 2009. doi: 10.1051/0004-6361/200911787.

[90] A. Vikhlinin, A. Kravtsov, W. Forman, C. Jones, M. Markevitch, S. S. Murray, and L. Van Speybroeck. Chandra Sample of Nearby Relaxed Galaxy Clusters: Mass, Gas Fraction, and Mass-Temperature Relation. *ApJ*, 640:691–709, April 2006. doi: 10.1086/500288.

[91] Y. Wang, H. Xu, L. Gu, J. Gu, Z. Qin, J. Wang, Z. Zhang, and X.-P. Wu. A joint Chandra and XMM-Newton view of Abell 3158: a massive off-centre cool gas clump as a robust diagnostic of a merger stage. *MNRAS*, 403:1909–1918, April 2010. doi: 10.1111/j.1365-2966.2010.16264.x.

[92] W. A. Watson, I. T. Iliev, J. M. Diego, S. Gottlober, A. Knebe, E. Martinez-Gonzalez, and G. Yepes. Statistics of extreme objects in the Juropa Hubble Volume simulation,. *MNRAS*, 437:3776–3786, February 2014. doi: 10.1093/mnras/stt2173.

[93] Z. L. Wen and J. L. Han. Substructure and dynamical state of 2092 rich clusters of galaxies derived from photometric data. *MNRAS*, 436:275–293, November 2013. doi: 10.1093/mnras/stt1581.

[94] J. D. Wing and E. L. Blanton. An Examination of the Optical Substructure of Galaxy Clusters Hosting Radio Sources. *ApJ*, 767:102, April 2013. doi: 10.1088/0004-637X/767/2/102.

Chapter 4

Statistical studies of galaxy clusters according to their dynamical status.

4.1 Introduction.

Galaxy clusters are one of the main astronomical objects where it has been demonstrated the presence of magnetic fields [8, 10] and studied their properties [4, 13], however the role that they play in the formation of structures is still under debate. Taking into account the electronic densities inside galaxy clusters, the magnetic field should be able to produce a signal of the Faraday Rotation (RF) effect that can give us information about the physics of the intra-cluster gas [15, 19, 23].

Although the presence of magnetic fields in galaxy cluster was demonstrated, the role of these fields on the evolution of the universe and in the formation of structures it is not understood in its totality. Taking into account the hierarchical structure formation and the physics of the electromagnetic processes, it is expected that the merger of 2 or more systems can affect the magnetic fields in the intra-cluster medium.

Numerous previous works [3, 14, 16, 17, 20] have studied the cosmological properties of magnetic fields and their influence on the evolution of the universe. In particular *Stasyszyn et al.* [18] studied the relationship between the large-scale structure of the universe (LSS) and the RF signal using hydrodynamic simulations. Although they found that LSS should produce a signal, this will be below the sensitivity of the current instruments. However, with the new generation of instruments and, by means of statistical techniques,

the effects produced by the fields of the LSS [7, 12, 22] could be detected.

One of the first works to statistically study the intra-cluster medium (ICM) of 16 galaxy groups was *Clarke et al.* [9], where they prove that the magnetic fields intensity increases towards the center of the systems. Modelling the inferred electron density from X-rays observations, they calculate an average magnetic field of $< |B| >= 5 - 10\ (l/10\mathrm{kpc})^{1/2}\ h_{75}^{1/2} \mu G$, where l is the characteristic length of the ICM cell. On another study *Boehringer et al.* [5] analyse the correlation between the RF signal and different properties of the galaxy clusters. Using 1722 extragalactic clusters from CLASSIX they find that the RF signal increases with the ICM density and infer a magnetic field of $< |B| >= 2 - 6\ (l/10\mathrm{kpc})^{1/2} \mu G$.

In this chapter we present the results of the study of the correlation between the dynamic state of galaxy clusters and their magnetic fields through the analysis of the RF signal in the line of sight of each cluster. In order to do this we use a sample of relaxed clusters and a sample of unrelaxed clusters and, taking into account the RF signal in the line of sight of each one, we infer properties on its magnetic fields.

We found an increase in the RF signal in the unrelaxed clusters that suggest the presence of physical processes that increase the magnetic fields in this systems.

This work was done in collaboration with Dr. Federico Stasyszyn and the results were recently submitted for its publication on the journal *Monthly Notices of the Royal Astronomy*.

4.2 Observational data.

4.2.1 Galaxy clusters survey.

Taking into account that in the construction of the merging cluster catalogue through machine learning techniques we decide to maintain low contamination rather than have a high completeness, we do not have enough number of clusters to make an statistical analysis. This is why we decided to study the correlation between galaxy clusters and RF measurements using the catalogue built by *Wen & Han* [25]. In this catalogue the authors classify the groups in relaxed and not relaxed systems. This catalogue was constructed based on photometric data from the SDSS/DR8, and has 589 and 1503 relaxed and not relaxed systems respectively. It is important to have in mind that relaxed groups are those that are not

interacting and that, therefore, should present an smooth and symmetric mass and luminosity distribution. While, the unrelaxed systems are those that have large substructures and that may be the product of interactions with other similar clusters.

Interactions between clusters introduce significant modifications on all the components of the systems (galaxies, ICM, dark matter distribution, etc.), which produces particular characteristics in different bands of the electromagnetic spectrum (presence of radio haloes, symmetrical distribution of X-ray emission, etc.). Keeping in mind this *Wen & Han* calibrated their method using a set of clusters with multiband information and a known dynamic state. Then they estimate 3 parameters that quantify different properties of the light distribution of these systems, namely [1]:

- α: Taking into account the smoothed luminosity map I, they define the parameter $S^2 = \sum_{ij} I^2(x_i, y_j)$ where (x, y) is the position of a given pixel of the map in a coordinate system centred on each cluster. Then, taking into account that the relaxed groups have a more symmetrical luminosity distribution than the unrelaxed ones, they define the parameter $\Delta^2 = \sum_{ij}[I(x_i, y_j) - I(-x_i, -y_j)]^2/2$. Finally they define the asymmetry parameter as $\alpha = \frac{\Delta^2}{S^2}$, where $\alpha = 0$ implies a symmetric distribution and $\alpha = 1$ an asymmetrical one.

- β: Bearing in mind that it is expected that the smoothed luminosity surface brightness maps of relaxed systems are similar in all directions and that, on the other hand, if a system has significant substructures, it is expected that their maps will have an abrupt change in a certain direction, the β parameter quantifies the difference in the brightness profile in different directions.

- δ: Taking into account the spherical symmetry of the luminosity distribution in relaxed groups, the luminosity maps can be correctly fitted by a $2 - dimensional$ *King* profile. On the other hand, clusters with a lot of degree of substructures should not be correctly adjusted with this model. The parameter δ quantifies the deviation of the luminosity map of a given group from a $2 - dimensional$ *King* profile.

Finally using these parameters and, calibrating the method taking into account the dynamic state of the known systems, they define the relaxation parameter Γ as:

[1] For a detailed definitions of these parameters see *Wen & Han* [25]

Table 4.1: Correlation between galaxy and AGN surveys in different photometric bands.

Survey	Radio	Unrelax	Relax
Veron et al. 2010 [24]	$30'$	72%	74%
Assef et al. 2018 [2]	$30'$	97%	96%
Abdo et al. 2010 [1]	$30'$	1.9%	1.8%
Horiuchi et al. 2004 [11]	$30'$	0.8%	0.3%

$$\Gamma = \beta - 1.90\alpha + 3.58\delta + 0.10 \tag{4.1}$$

They found that relaxed clusters have $\Gamma > 0$, while unrelaxed systems have $\Gamma < 0$.

After that, they estimate the Γ parameter for 2092 clusters of the SDSS/DR8 survey. In the final sample, they found 589 relaxed systems and 1503 non-relaxed systems. In figure 4.1 you can see the distributions of the main properties of both samples. It can be seen that all the distributions are very similar, with the exception of the the R_{mag} magnitude distribution, where the relaxed groups are slightly brighter. Excluding this magnitude, the 2 samples have statistically the same properties and only differ in their dynamical state.

Taking into account the possibility of having AGNs inside the clusters that may affect the ICM in a considerable way, we study the probability of having an AGN inside a cluster. Specifically we study the correlation between the relaxed and unrelaxed systems with the AGNs present in the catalogues *Veron et al. 2010* [24], *Assef et al. 2018* [2], *Abdo et al. 2010* [1] (γ-Rays) and *Horiuchi et al. 2004* [11] (Radio). With this goal in mind we count how many clusters have at least one AGN in radials bins. We found that both samples (relaxed and unrelaxed) have the same proportions of clusters with and without AGN in all catalogues. This reinforces the idea that the only difference between both samples it their dynamical state. These results are summarised in the table 4.1.

4.2.2 Observations of the Faraday Rotation effect.

For the RF information we use the catalogue of *Taylor et al.* [21] that contains RF signal measurements of 37543 polarised sources from the NRAO VLA Sky Survey (NVSS). It is important to note that this catalogue only has observations in 2 frequencies, which may cause ambiguities for high RF values, however it has the advantage that it has a large

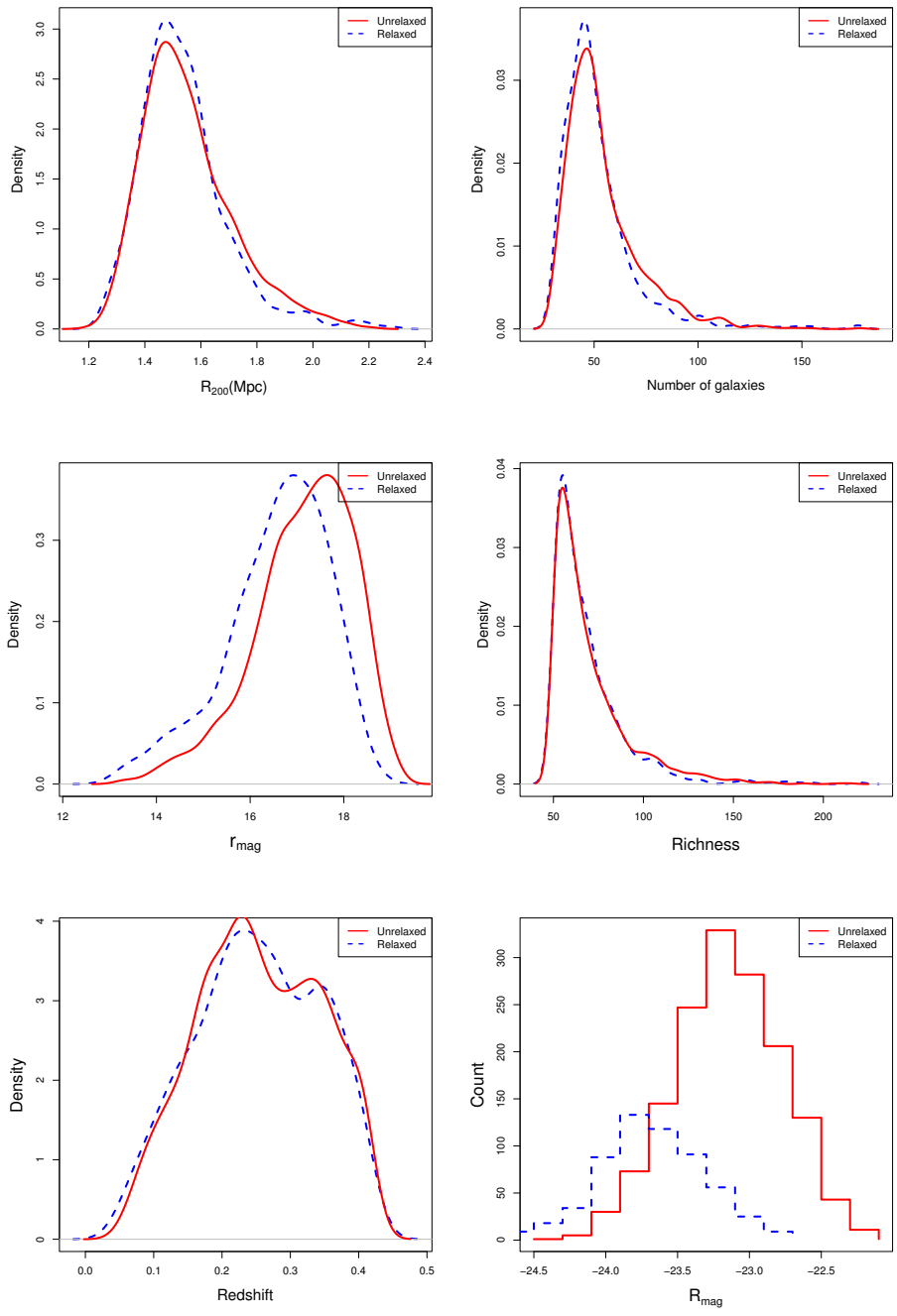

Figure 4.1: Comparison of the main properties of the relaxed and unrelaxed clusters. *Top Left*: Distribution of the cluster radius (R_{200}). *Top Right*: Distribution of the number of member galaxies. *Centre Left*: Distribution of apparent magnitude r_{mag}. *Centre Right*: Distribution of the cluster richness. *Bottom Left*: Distribution of redshifts. *Bottom Right*: Histogram of absolute magnitude R_{mag}.

coverage area that allows us to obtain a better correlation with the distribution of galaxy clusters. In addition, we only find 2 extreme RF values, so these ambiguities will not statistically contribute and, therefore, will not have an influence on our analysis.

In order to remove the contamination introduced by our galaxy, we estimate the average of the RF values in a circle of 6°around each cluster and then we subtract this value as described in *Boehringer et al.* [5] and *Stasyszyn et al.* [18].

In addition to the *Taylor et al.* catalogue [21] we also analysed the correlation between the RF measurements of the catalogue of *Xu et al.* [26] and the galaxy clusters. Although this is a smaller catalogue, it has the advantage that it have estimations of the redshift of the RF source, which allows us to study only those sources that are behind each one of the analysed clusters.

4.3 Statistical analysis.

In order to study the correlation between the RF measurements and the dynamical state of the galaxy clusters, we estimate the RF standard deviation (σ_{RF}) in 5 radial bins around each cluster. In figure 4.2 you can see in solid lines the standard deviation as a function of the angular distance for the relaxed (black) and unrelaxed samples (red). It can be seen that the unrelaxed clusters present a greater standard deviation in the central regions than relaxed systems. With the goal of studying if the differences in the standard deviation have a correlation with the number of galaxy members of each cluster, we made the same analysis but only taking into account those systems with more than 40 galaxies. These results are shown in dotted lines. As you can see the difference between relaxed and unrelaxed clusters increases when we use only groups with more than 40 members. This implies that this effect is dominated by larger systems. To estimate the errors in the measurement of the standard deviation we perform a bootstrapping technique that consists on estimating σ_{RF} by removing one galaxy cluster each time, and then estimating the dispersion in the measurements of σ_{RF}.

In order to avoid a bias that may be produced by the difference in the luminosity distributions of the relaxed and unrelaxed samples (see figure 4.1), we decided to build a random sub-sample of the unrelaxed groups that follows the R_{mag} distribution and the numerical density of the relaxed systems. Taking into account this sub-sample, we estimate

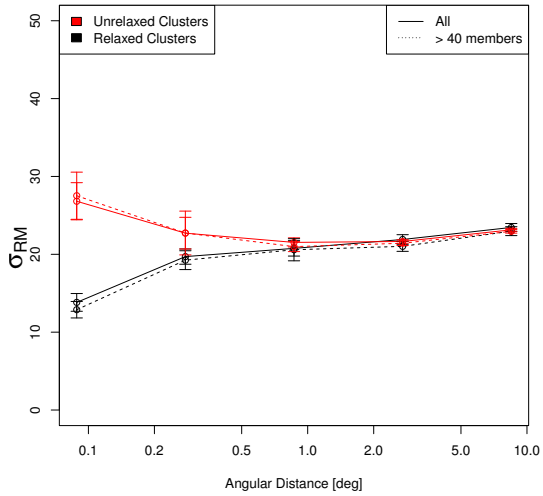

Figure 4.2: Standard deviation of the Faraday Rotation measurements as a function of the angular distance to the centre of the cluster. In red are shown the results for the non-relaxed systems, while in black are shown the results for the relaxed systems. In solid lines are shown the results for the complete samples, while the results for the subsamples with more than 40 member galaxies are shown in dotted lines.

the deviation σ_{RF} in radial bins as previously explained. Again we find that those systems that are not relaxed present a greater dispersion in the central areas.

In figure 4.3 we show σ_{RF} as a function of the projected distance to the centre of the cluster. Again can be observed the tendency that those unrelaxed systems have a greater dispersion in the inner zones for the sample with more than 40 galaxies (dotted lines) and for the total sample (solid lines).

Continuing with the analysis, we divide the sample galaxy clusters into bright ($R_{mag} < -23.5$) and faint ($R_{mag} > -23.5$) systems and estimate the standard deviation σ_{RF} in radial bins for both samples without taking into account the dynamic state of the systems. In figure 4.4 we present the dispersion σ_{RF} as a function of the angular distance. As can be seen there is not a significant difference between these samples, which indicates that the difference previously found, is produced by the dynamical state of the clusters.

In figure 4.5 you can see the standard deviation as a function of the projected distance normalized to the radius R_{200} of each system. It can be seen that the difference between the relaxed and unrelaxed samples persists until approximately 2 times the radius R_{200},

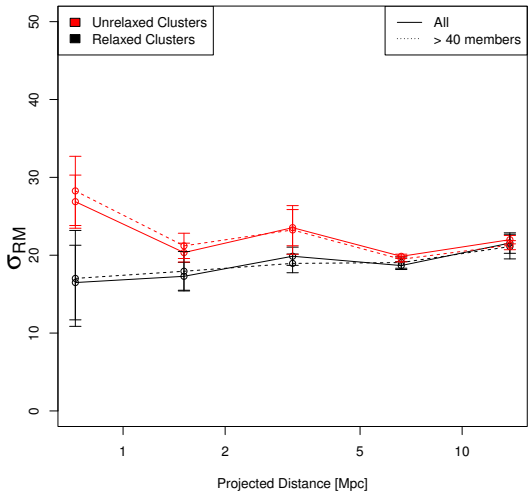

Figure 4.3: Standard deviation of the Faraday Rotation measurements as a function of the projected distance to the centre of the cluster. In red are shown the results for the unrelaxed sample, while in black are plotted the results for relaxed systems. In solid lines are shown the results for the complete samples, while in dotted lines are plotted the results for the subsample with more than 40 member galaxies.

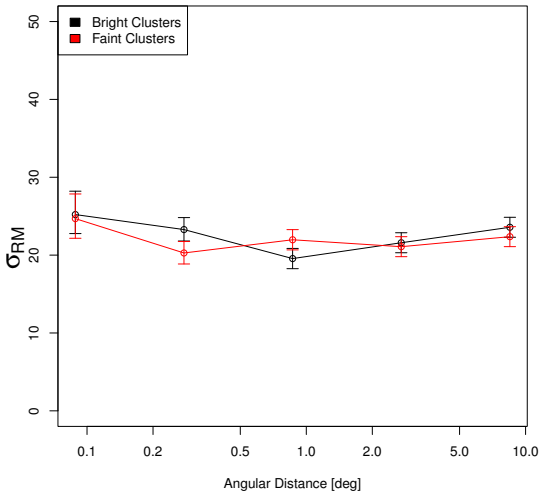

Figure 4.4: Standard deviation of the Faraday Rotation measurements as a function of the angular distance to the centre of the cluster. In black are shown the results for the bright systems sample, while in red are shown the results for the faint systems.

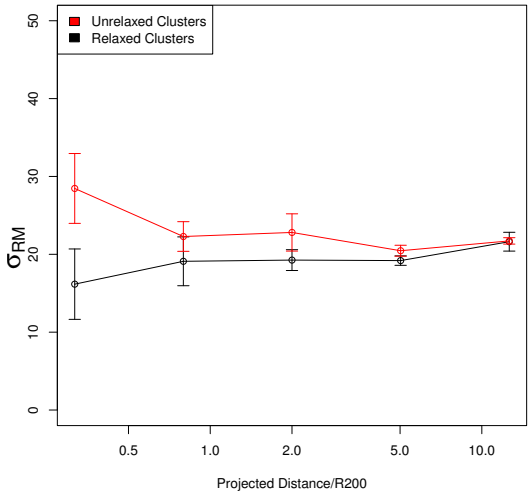

Figure 4.5: Standard deviation of the Faraday Rotation measurements as a function of the projected distance to the centre of the cluster, normalized to the radius R_{200}. In red are shown the results for the unrelaxed sample, while in black are shown the results for the relaxed ones. In solid lines are shown the results for the complete samples, while in dotted lines are shown the results for the subsample with systems with more than 40 member galaxies.

showing that this difference is inherent to the clusters.

As can be seen in all the figures, the difference between the relaxed and unrelaxed groups increases when we only study those systems with more than 40 galaxies. This can be understood because the smaller systems do not really contribute to this effect, due to the fact that they have relatively smaller masses and sizes and so, they do not produce a significant amount of Faraday rotation.

Bearing in mind that Faraday's rotation is an integrated effect along the line of sight, a natural source of uncertainty in our analysis is the integration without knowing the distance towards the polarised sources. To reduce this source of error we use the catalogue of *Xu et al.* [26], which has approximately 3600 confirmed extragalactic sources on radio with an estimation if their redshift. This allows us to estimate the standard deviation σ_{RF} using only those sources that are behind each group. In figure 4.6 you can see the results of this analysis, where we observe again the tendency for unrelaxed systems to have a greater dispersion in the central zones. A problem of this catalogue is the small number of sources

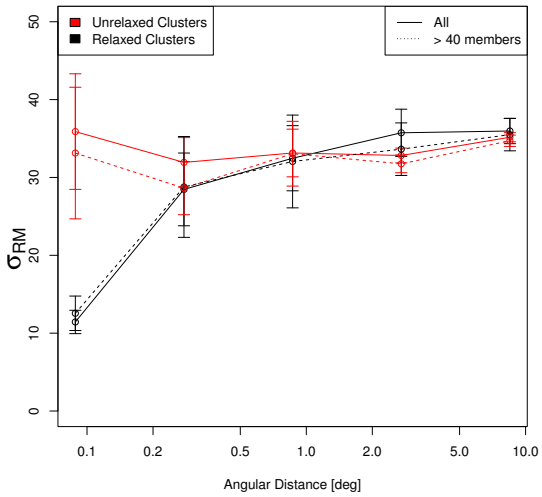

Figure 4.6: Standard deviation of the Faraday Rotation measurements as a function of the angular distance to the centre of the cluster, for the catalogue of *Xu et al.* [26] and taking into account only the polarised sources behind each of the studied groups. In red are shown the results for the unrelaxed systems, while in black are shown the results for the relaxed systems. In solid lines are shown the results for the complete samples, while in dotted lines are shown the results for the subsamples with systems with more than 40 member galaxies.

(approximately the 6% of the number of sources that have the catalogue of *Taylor et al.* [21]) and low angular coverage, which translates into greater errors in the measurement.

4.3.1 Density distributions.

Continuing with the analysis of the correlations between RF measurements and the dynamic state of the galaxy clusters we studied the probability of a random distribution of galaxy groups to overlap RF measurements.

In order to do this we generated 1000 catalogues randomly distributing the angular positions of the galaxy clusters and calculating the number of systems with at least one RF measurement at less than one radius R_{200}. The figure 4.7 shows the probability distribution of the random catalogues (relaxed systems in black and unrelaxed systems in red) to have at least one RF measurement at less than one radius R_{200}. In a comparative way, it is shown in solid black vertical line and in vertical red dotted line the amount of relaxed and

unrelaxed groups in the real catalogue respectively that have at least one RF measurement within a radius R_{200}. It can be seen that both relaxed and unrelaxed systems present values that are more than 3σ of the average found for the random catalogues and with opposite tendencies. For a better comparison, the smallest number of samples in random catalogues with at least one measurement was subtracted from both samples. It is worth mentioning that the error in the measurements was estimated using a bootstrap technique as was previously explained. These results show that relaxed (unrelaxed) clusters have a lower (higher) probability of having at least one RF measurement within a radius R_{200} than a random sample.

In figure 4.8 the same analysis is shown but taking into account those systems with more and less than 40 galaxies members. It can be seen that in the sample of clusters with more than 40 galaxies the differences persist, while in the sample of systems with less than 40 galaxy members there are no significant differences between the relaxed (unrelaxed) samples with respect to the random catalogues.

It is worth noting that these results are consistent with those previously found, in which it was evident that unrelaxed systems have a greater dispersion towards central areas than relaxed systems.

4.4 Inferred magnetic fields.

The Faraday rotation originates from the integration of the magnetic fields in the line of sight and the value of σ_{RF} is affected by the characteristic length of the ICM cell in the line of sight $\Lambda = (L/l)^{1/2}$, where L is the length of the ICM column electron density and l is the characteristic length of the ICM cell with coherent magnetic field. From this, we are able to derive a relation between the σ_{RM}, the electron density N_e and the magnetic field $B_{\|}$ as [5]:

$$\frac{B_{||}}{1\mu\mathrm{G}} = 3.801 \times 10^{18} \frac{\sigma_{RM}}{\mathrm{rad\ m^{-2}}} \left(\frac{N_e}{\mathrm{cm^{-2}}}\right)^{-1} \Lambda \tag{4.2}$$

Since we do not have a reliable value for the electron density N_e, we use a constant value of $N_e = 10^{21} cm^2$. In addition, assuming that $L \sim 1$ Mpc and $l \sim 1$ Kpc we can infer a magnetic field value of $B_{\|} = 3.0 \pm 0.25\mu$G for the unrelaxed systems and $B_{\|} = 2.2 \pm 0.40\mu$G for the relaxed ones.

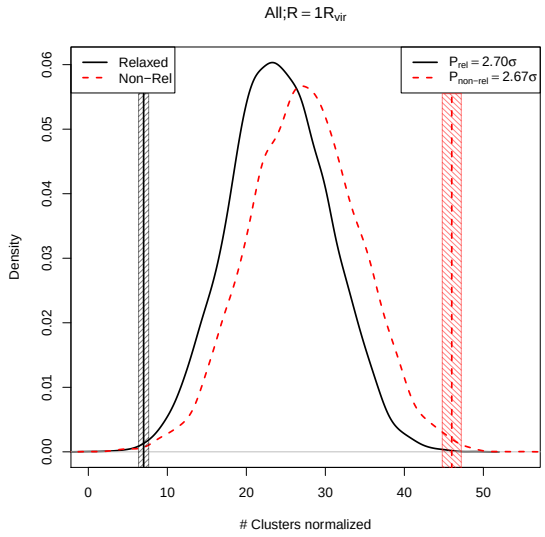

Figure 4.7: Probability distribution of having at least one RF measurement within a radius R_{200} for 1000 catalogues constructed randomly sorting the angular positions. In black and red are shown the measured values for the relaxed and unrelaxed systems respectively. The vertical values show the values measured in the real catalogue of *Wen & Han* [25]. For a better comparison we subtract the smaller number of clusters with at least one RF measurement within a radius of R_{200} found in the random catalogues for each sub-set (relaxed and not relaxed).

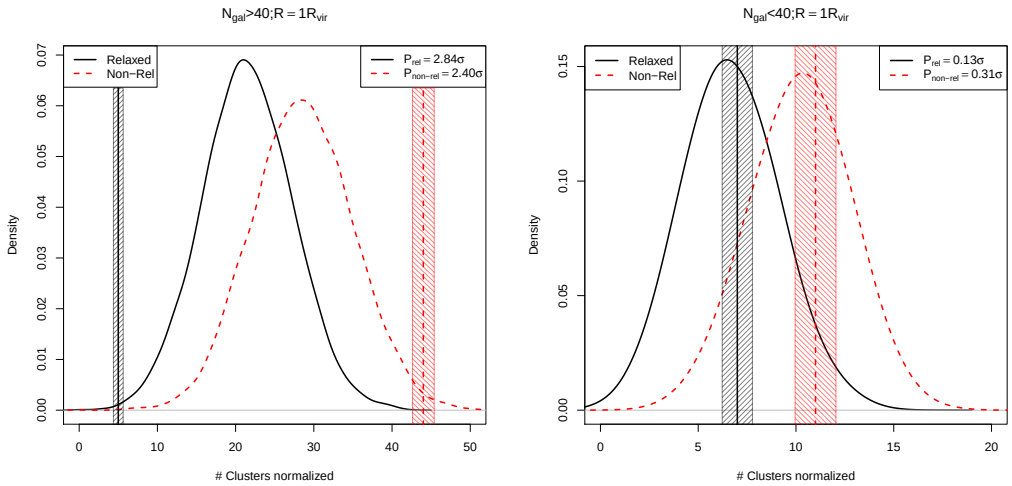

Figure 4.8: Probability distribution of having at least one RF measurement within a radius R_{200} for clusters with more (less) than 40 galaxy members in 1000 catalogues constructed randomly sorting the angular positions. In black and red are shown the measured values for the relaxed and unrelaxed systems respectively. The vertical values show the values measured in the real catalogue of *Wen & Han* [25]. For a better comparison we subtract the smaller number of clusters with at least one RF measurement within a radius of R_{200} found in the random catalogues for each sub-set (relaxed and not relaxed).

On the other hand if we assume that $l = 10$ Kpc for the unrelaxed systems and $l = 25$ Kpc for the relaxed ones the values for the magnetic field are $B_{\parallel} = 0.96 \pm 0.08$ μG and $B_{\parallel} = 0.45 \pm 0.08 \mu$G respectively.

It is worth mentioning that we assume a constant value of L because we did not find significant differences in the distributions of the fundamental properties of relaxed and unrelaxed systems (See figure 4.1). However, if we have in mind that turbulence is expected to be greater in unrelaxed systems, it is expected that these clusters will have a lower value of l.

It is important to note that if indeed the dispersion σ_{RF} is smaller in the relaxed clusters, it can be a consequence of the depolarisation effect [6], which would imply that we would not be able to observe the Faraday rotation even if there are magnetic fields. This phenomenon could explain the results found in section 4.3.1. While the *Taylor et al.* catalogue [21] was constructed based on data extrapolated from NVSS, the data were not designed for polarisation studies, so they may be affected by the effect of depolarisation. Nevertheless the fact that the difference between relaxed and unrelaxed is found in the analysis of different catalogues (*Taylor et al. [21]* and *Xu et al.* [26]) is an indication that the difference between relaxed and unrelaxed systems is real.

4.5 Conclusions.

Taking into account the hierarchical formation of structures of the standard model and the presence of magnetic fields in galaxy clusters, in this work we study the relationship between the dynamical state of the clusters and the Faraday rotation effect produced by them.

With this goal in mind, we use the galaxy cluster survey built by *Wen & Han* [25] which contains information about the dynamic status of each system, and correlated it with the sources of Faraday rotation.

We found that unrelaxed clusters have a greater standard deviation σ_{RF} of the Faraday rotation effect that the relaxed ones. This difference can be interpreted as the unrelaxed clusters having higher magnetic fields, what can be explained taking into account the dynamo effect caused by the mergers between clusters.

It is worth noting that the only difference between the relaxed and unrelaxed system

samples is their R_{mag} magnitude distribution (see figure 4.1), with the relaxed clusters being brighters. To verify that the difference in σ_{RF} is not a result of this magnitude difference, we separate the systems into bright and faint clusters without taking into account the dynamic state. When comparing the distribution of σ_{RF} of these sub-sets we did not found any difference, which reinforces the idea that the difference founded in section 4.3 is actually produced by the dynamic status of the systems.

In addition, we study the probability that a system belonging to a random sample of galaxy clusters has at least one RF measurement within a radius R_{200}. We found that unrelaxed groups are more likely to have at least one RF measurement within a radius R_{200} that a random sample, while a relaxed system has a lower probability than a random sample, which reinforces the hypothesis that the dynamical state of the systems significantly modifies the magnetic properties of galaxy clusters.

Taking into account the standard deviation σ_{RF} measured, and assuming typical values for the electronic density N_e, the length of the electronic density column of the ICM L and the characteristic length of the ICM cell with coherent fields we have been able to estimate a magnetic field value of $B_{\parallel} = 0.45\ 0.08 \mu\mathrm{G}$ for relaxed systems and $B_{\parallel} = 0.96\ 0.08\ \mu\mathrm{G}$ for the unrelaxed ones.

Bibliography

[1] A. A. Abdo, M. Ackermann, M. Ajello, A. Allafort, E. Antolini, W. B. Atwood, M. Axelsson, L. Baldini, J. Ballet, G. Barbiellini, D. Bastieri, B. M. Baughman, K. Bechtol, R. Bellazzini, B. Berenji, R. D. Blandford, E. D. Bloom, J. R. Bogart, E. Bonamente, A. W. Borgland, A. Bouvier, J. Bregeon, A. Brez, M. Brigida, P. Bruel, R. Buehler, T. H. Burnett, S. Buson, G. A. Caliandro, R. A. Cameron, A. Cannon, P. A. Caraveo, S. Carrigan, J. M. Casandjian, E. Cavazzuti, C. Cecchi, Ö. Çelik, A. Celotti, E. Charles, A. Chekhtman, A. W. Chen, C. C. Cheung, J. Chiang, S. Ciprini, R. Claus, J. Cohen-Tanugi, J. Conrad, L. Costamante, G. Cotter, S. Cutini, V. D'Elia, C. D. Dermer, A. de Angelis, F. de Palma, A. De Rosa, S. W. Digel, E. d. C. e. Silva, P. S. Drell, R. Dubois, D. Dumora, L. Escande, C. Farnier, C. Favuzzi, S. J. Fegan, E. C. Ferrara, W. B. Focke, P. Fortin, M. Frailis, Y. Fukazawa, S. Funk, P. Fusco, F. Gargano, D. Gasparrini, N. Gehrels, S. Germani, B. Giebels, N. Giglietto, P. Giommi, F. Giordano, M. Giroletti, T. Glanzman, G. Godfrey, P. Grandi, I. A. Grenier, M.-H. Grondin, J. E. Grove, S. Guiriec, D. Hadasch, A. K. Harding, M. Hayashida, E. Hays, S. E. Healey, A. B. Hill, D. Horan, R. E. Hughes, G. Iafrate, R. Itoh, G. Jóhannesson, A. S. Johnson, R. P. Johnson, T. J. Johnson, W. N. Johnson, T. Kamae, H. Katagiri, J. Kataoka, N. Kawai, M. Kerr, J. Knödlseder, M. Kuss, J. Lande, L. Latronico, C. Lavalley, M. Lemoine-Goumard, M. Llena Garde, F. Longo, F. Loparco, B. Lott, M. N. Lovellette, P. Lubrano, G. M. Madejski, A. Makeev, G. Malaguti, E. Massaro, M. N. Mazziotta, W. McConville, J. E. McEnery, S. McGlynn, P. F. Michelson, W. Mitthumsiri, T. Mizuno, A. A. Moiseev, C. Monte, M. E. Monzani, A. Morselli, I. V. Moskalenko, S. Murgia, P. L. Nolan, J. P. Norris, E. Nuss, M. Ohno, T. Ohsugi, N. Omodei, E. Orlando, J. F. Ormes, M. Ozaki, D. Paneque, J. H. Panetta, D. Parent, V. Pelassa, M. Pepe, M. Pesce-Rollins, S. Pi-

ranomonte, F. Piron, T. A. Porter, S. Rainò, R. Rando, M. Razzano, A. Reimer, O. Reimer, T. Reposeur, J. Ripken, S. Ritz, A. Y. Rodriguez, R. W. Romani, M. Roth, F. Ryde, H. F.-W. Sadrozinski, D. Sanchez, A. Sander, P. M. Saz Parkinson, J. D. Scargle, C.Sgrò, M. S. Shaw, E. J. Siskind, P. D. Smith, G. Spandre, P. Spinelli, J.-L. Starck, Ł. Stawarz, M. S. Strickman, D. J. Suson, H. Tajima, H. Takahashi, T. Takahashi, T. Tanaka, G. B. Taylor, J. B. Thayer, J. G. Thayer, D. J. Thompson, L. Tibaldo, D. F. Torres, G. Tosti, A. Tramacere, P. Ubertini, Y. Uchiyama, T. L. Usher, V. Vasileiou, N. Vilchez, M. Villata, V. Vitale, A. P. Waite, E. Wallace, P. Wang, B. L. Winer, K. S. Wood, Z. Yang, T. Ylinen, and M. Ziegler. The First Catalog of Active Galactic Nuclei Detected by the Fermi Large Area Telescope. *ApJ*, 715:429–457, May 2010. doi: 10.1088/0004-637X/715/1/429.

[2] R. J. Assef, D. Stern, G. Noirot, H. D. Jun, R. M. Cutri, and P. R. M. Eisenhardt. The WISE AGN Catalog. *The Astrophysical Journal Supplement Series*, 234:23, February 2018. doi: 10.3847/1538-4365/aaa00a.

[3] R. Beck. Galactic and extragalactic magnetic fields - a concise review. *Astrophysics and Space Sciences Transactions*, 5:43–47, October 2009. doi: 10.5194/astra-5-43-2009.

[4] H. Böhringer and N. Werner. X-ray spectroscopy of galaxy clusters: studying astrophysical processes in the largest celestial laboratories. *A&ARv*, 18:127–196, February 2010. doi: 10.1007/s00159-009-0023-3.

[5] H. Böhringer, G. Chon, and P. P. Kronberg. The Cosmic Large-Scale Structure in X-rays (CLASSIX) Cluster Survey. I. Probing galaxy cluster magnetic fields with line of sight rotation measures. *A&A*, 596:A22, November 2016. doi: 10.1051/0004-6361/201628873.

[6] A. Bonafede, F. Govoni, L. Feretti, M. Murgia, G. Giovannini, and M. Brüggen. Fractional polarization as a probe of magnetic fields in the intra-cluster medium. *A&A*, 530:A24, June 2011. doi: 10.1051/0004-6361/201016298.

[7] A. Bonafede, F. Vazza, M. Brüggen, T. Akahori, E. Carretti, S. Colafrancesco, L. Feretti, C. Ferrari, G. Giovannini, F. Govoni, M. Johnston-Hollitt, M. Murgia, A. Scaife,

V. Vacca, F. Govoni, L. Rudnick, and A. Scaife. Unravelling the origin of large-scale magnetic fields in galaxy clusters and beyond through Faraday Rotation Measures with the SKA. *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, art. 95, April 2015.

[8] C. L. Carilli and G. B. Taylor. Cluster Magnetic Fields. *Ann. Rev. Astron. Astrophys.*, 40:319–348, 2002. doi: 10.1146/annurev.astro.40.060401.093852.

[9] T. E. Clarke, P. P. Kronberg, and H. Böhringer. A New Radio-X-Ray Probe of Galaxy Cluster Magnetic Fields. *ApJL*, 547:L111–L114, February 2001. doi: 10.1086/318896.

[10] L. Feretti, G. Giovannini, F. Govoni, and M. Murgia. Clusters of galaxies: observational properties of the diffuse radio emission. *A&ARv*, 20:54, May 2012. doi: 10.1007/s00159-012-0054-z.

[11] S. Horiuchi, E. B. Fomalont, W. K. Taylor, A. R. Scott, J. E. J. Lovell, G. A. Moellenbrock, R. Dodson, Y. Murata, H. Hirabayashi, P. G. Edwards, L. I. Gurvits, and Z. Q. Shen. The VSOP 5 GHz Active Galactic Nucleus Survey. IV. The Angular Size/Brightness Temperature Distribution. *ApJ*, 616:110–122, November 2004. doi: 10.1086/424811.

[12] R. Kale, K. S. Dwarakanath, D. Vir Lal, J. Bagchi, S. Paul, S. Malu, A. Datta, V. Parekh, P. Sharma, and M. Pandey-Pommier. Clusters of Galaxies and the Cosmic Web with Square Kilometre Array. *Journal of Astrophysics and Astronomy*, 37:31, December 2016. doi: 10.1007/s12036-016-9406-9.

[13] A. V. Kravtsov and S. Borgani. Formation of Galaxy Clusters. *Ann. Rev. Astron. Astrophys.*, 50:353–409, September 2012. doi: 10.1146/annurev-astro-081811-125502.

[14] P. P. Kronberg. *Cosmic Magnetic Fields*. UK: Cambridge University Press, October 2016.

[15] F. Marinacci, M. Vogelsberger, R. Pakmor, P. Torrey, V. Springel, L. Hernquist, D. Nelson, R. Weinberger, A. Pillepich, J. Naiman, and S. Genel. First results from

the IllustrisTNG simulations: radio haloes and magnetic fields. *ArXiv e-prints*, July 2017.

[16] R. Pakmor, F. Marinacci, and V. Springel. Magnetic Fields in Cosmological Simulations of Disk Galaxies. *ApJL*, 783:L20, March 2014. doi: 10.1088/2041-8205/783/1/L20.

[17] R. Pakmor, F. A. Gómez, R. J. J. Grand, F. Marinacci, C. M. Simpson, V. Springel, D. J. R. Campbell, C. S. Frenk, T. Guillet, C. Pfrommer, and S. D. M. White. Magnetic field formation in the Milky Way like disc galaxies of the Auriga project. *MNRAS*, 469:3185–3199, August 2017. doi: 10.1093/mnras/stx1074.

[18] F. Stasyszyn, S. E. Nuza, K. Dolag, R. Beck, and J. Donnert. Measuring cosmic magnetic fields by rotation measure-galaxy cross-correlations in cosmological simulations. *MNRAS*, 408:684–694, October 2010. doi: 10.1111/j.1365-2966.2010.17166.x.

[19] F. A. Stasyszyn, K. Dolag, and A. M. Beck. A divergence-cleaning scheme for cosmological SPMHD simulations. *MNRAS*, 428:13–27, January 2013. doi: 10.1093/mnras/sts018.

[20] F. Stefani, T. Albrecht, R. Arlt, M. Christen, A. Gailitis, M. Gellert, A. Giesecke, O. Goepfert, J. Herault, O. Kirillov, G. Mamatsashvili, J. Priede, G. Ruediger, M. Seilmayer, A. Tilgner, and T. Vogt. Magnetic field dynamos and magnetically triggered flow instabilities. *ArXiv e-prints*, May 2017.

[21] A. R. Taylor, J. M. Stil, and C. Sunstrum. A Rotation Measure Image of the Sky. *ApJ*, 702:1230–1236, September 2009. doi: 10.1088/0004-637X/702/2/1230.

[22] F. Vazza, C. Ferrari, A. Bonafede, M. Brüggen, C. Gheller, R. Braun, and S. Brown. Filaments of the radio cosmic web: opportunities and challenges for SKA. *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, art. 97, April 2015.

[23] F. Vazza, G. Brunetti, M. Brüggen, and A. Bonafede. Resolved magnetic dynamo action in the simulated intracluster medium. *MNRAS*, 474:1672–1687, February 2018. doi: 10.1093/mnras/stx2830.

[24] M.-P Véron-Cetty and P. Véron. A catalogue of quasars and active nuclei: 13th edition. *A&A*, 518:A10, July 2010. doi: 10.1051/0004-6361/201014188.

[25] Z. L. Wen and J. L. Han. Substructure and dynamical state of 2092 rich clusters of galaxies derived from photometric data. *MNRAS*, 436:275–293, November 2013. doi: 10.1093/mnras/stt1581.

[26] J. Xu and J.-L. Han. A compiled catalog of rotation measures of radio point sources. *Research in Astronomy and Astrophysics*, 14:942-958, August 2014. doi: 10.1088/1674-4527/14/8/005.

Chapter 5

Individual study of the candidates for merging clusters $A2029/33$ and $A1204$.

In this chapter we will present the results of the dynamic and morphological study of the $A1204$ and $A2029/2033$ clusters that were previously identified as candidates for merging clusters with the `MeSsI` code. As was detailed in chapter 3, the merging galaxy groups present the ideal scenario to study the properties of dark matter, since, due to the merger, there is an offset between the dark matter distribution and the gas distribution [18]. Considering that the presence of dark matter produces distortions in the background galaxies, we perform a weak gravitational lens analysis on both systems in order to determine the density profiles and their masses. On the other hand, with the objective of studying the gas distribution in the ICM and, taking into account that at the typical temperatures that the ICM gas is found, it emits in X-rays, we carry out a study of this emission. In a complementary manner, an individual dynamic study was carried out using spectroscopic data from both galaxy clusters. Finally, we performed an individual dynamic study, as described in *Beers et al. 1982* [6], taking into account the 2 body model to decide if the substructures found are gravitationally linked.

The results of these studies were published in the international journal *Astronomy & Astrophysics* in the year 2017 [20].

It should be noted that this work was carried out in collaboration with Dra. Elizabeth Gonzales (weak gravitational lenses analysis), Dr. Gabriel Oio (X-rays emission analysis), Dr. Mariano Dominguez and Dr. Carlos Valotto (statistical analysis), Dr. Nilo Castellano,

ID.	α J200	δ J200	z	Program ID.	Filters	Exp. time [s]	*Seeing* [arcsec]
A1204	168.3845°	17.594°	0.1706	o10114	Subaru $i+$	240	0.74
				o11203	Subaru V	240	0.91
				WG931430P	ROSAT PSPCC	14699	
A2029	227.7446°	5.7616°	0.0775	06AC16	CFHT i.MP9701	500	0.52†
				03AC28	CFHT r.MP9601	480	0.65†
				US800249P	ROSAT PSPCB	12542	

Table 5.1: Specifications of the observational data used in the study of each cluster. In column (1) we detailed the ID of each system. In columns (2), (3) and (4), the coordinates and redshift of the centres of the systems are detailed. Finally, in columns (5), (6), (7) and (8) we present the specifications of the observational data used. († Average *seeing* in the mosaic.).

lic. Tania Tagliaferro, Dr. Héctor Cuevas and lic. Daniel Lang (data reduction).

5.1 Data Acquisition.

5.1.1 Photometric observations.

For the weak lensing analysis it is necessary to have deep photometric images with a small seeing (less than $1''$) to be able to estimate the ellipticity of the background galaxies. In table 5.1 we present the details of the observational data used for the study of each system. In the case of $A1204$ cluster, for the analysis of lenses we use images in the $i+$ and V Subaru bands obtained from the SMOKA database (Subaru Mitaka Okayama Kiso Archive). These observations were taken using the *Suprime-Cam* camera [28] mounted on the primary focus of the *Subaru* telescope as part of the COSMOS program (*Cosmic Evolution Survey*). This camera has a mosaic of 80 mega-pixels that cover an area of $34' \times 27'$ with a resolution of $0.202'/pixel$.

In the case of the $A2029/2033$ system, the photometric observations were obtained from the CADC database (*Canadian Astronomy Data Center*). These images were obtained with the MEGACAM camera in the $3.6m$ CFHT telescope in the r' and i' bands. This camera consists of 36 CCDs with $2048x4612$ pixels that cover $1x1$ square degree with a resolution of $0.187''$.

5.1.2 X-ray emission observations.

For the study of the ICM gas distribution we use data of the X-ray emission from the ROSAT [1] telescope that were downloaded from the HEASARC on-line repository (*High Energy Astrophysics Science Archive Research Center*). These data were reduced using the `ximage` software to eliminate sources of contamination and to subtract the signal from the background. For the case of the $A1204$ group we obtained an rms noise of $0.32(cts/pix)$, while for $A2029/2033$ we obtained an rms of $0.72(cts/pix)$. Finally, we used the images in the $0.1 - 2.4KeV$ band to measure the surface brightness of the X-ray emission in both system.

5.1.3 Spectroscopic Data.

In order to carry out an individual dynamic study with a greater precision it is also necessary to use the redshift of the galaxies, and so we have decided to work with spectroscopic surveys of both clusters.

For the case of $A1204$ we used the HeCS spectroscopic survey [35] (*Hectospec Cluster Survey*), built by *Rines et al. 2013* using the Hectospec [17] instrument mounted on the $6.5m$ MMT telescope. It should be noted that the clusters observed in this survey were previously selected based on the RASS survey [40] (*ROSAT All Sky Survey*) restricted to those systems with $0.1 \leq z \leq 0.3$.

In the case of the $A2029/2033$ cluster, we used the data from the SDSS DR7 survey [1, 44] restricted to galaxies with an r band apparent magnitude between 14.5 and 17.77.

5.2 Weak lensing analysis.

In order to obtain the distribution of projected mass density and the total mass of each substructure, we made, a weak lensing analysis applying the code presented in *Gonzalez et al. 2015* [19]. In this section we will briefly describe this code.

[1] See `http://www.mpe.mpg.de/xray/wave/rosat/index.php`

5.2.1 Identification and classification of sources.

To perform a weak gravitational lens study it is necessary to identify the background galaxies, that is, those galaxies that are behind the cluster we are analysing. For this we use the software `SExtractor` [7] in 2 stages. First, we compute the seeing and the saturation level of each image. Then, we ran again `SExtractor` but using the previously estimated seeing and saturation level as information. Finally, we classify the detected sources between galaxies and stars and we discard the false detections taking into account its position in the magnitude/central-flux diagram, its FWHM (full width at half maximum) with respect to the seeing and the stellarity index estimated by `SExtractor`. A visual inspection of the discarded sources shows that most of them are hot pixels, sources at the edges of the CCD, defects produced by saturated stars and cosmic rays.

Since we do not have an estimation of the redshift for all the identified sources we need a photometric criterion to identify background galaxies. We will consider that a galaxy is at the background if it has an apparent magnitude between $m_p < mag < m_{MAX} + 0.5$, where mag is the magnitude measured in the r' and $i+$ bands for the CFHT and SUBARU images respectively, m_p is the smallest magnitude for which the probability for a galaxy to be behind the cluster is greater than 0.7 and m_{MAX} corresponds to the peak of the mag distribution of all the identified galaxies. The upper cut ensures that we are not taking into account very faint galaxies that have higher uncertainties in the shape measurements. In the case of the SUBARU images, we obtained $m_{MAX} = 25.7$ for the band $i+$, while for the CFHT images we obtained $m_{MAX} = 22.7$ for the r' band. In addition to the magnitudes cut, we imposed a cut in the colour of the galaxy with the objective of discard blue galaxies that are likely to be foreground galaxies (i.e galaxies that are in front of the cluster under consideration) and, therefore, they dilute the true lens signal produced by the matter distribution of the galaxy cluster. Finally we are left with those galaxies with $r' - i' > -0.5$ and $V - i+ > -1.0$ identified in CFHT and SUBARU images respectively. The efficiency of the weak gravitational lens effect depends on the geometric factor $\beta := D_{LS}/D_S$, where D_{LS} and D_S are the angular diameter distances from the lens to the source and from the observer to the source, respectively. To estimate m_p and $\langle\beta\rangle$ (where $\langle ...\rangle$ express the average over the considered background galaxies for the lensing estimator) we use catalogues of photometric redshifts. In the case of the lensing analysis performed with CFHT frames we use the Coupon et al. [12] photometric catalogue, based

on the public release Deep Field 1 of the Canada–France–Hawaii Telescope Legacy Survey (CFHTLS Deep1), which has an 80% completeness limit of $m_{r'} \approx 26.$ and covers a sky region of roughly 1 deg^2 .

On the other hand, for the analysis with SUBARU frames we use the catalogue of photometric redshifts given by Laigle et al. [24] which contains precise photometric redshifts over the 2deg^2 COSMOS field and with a limiting magnitude $m_{i+} = 26.2$. We compute the fraction of galaxies with $z > z_{cluster}$ in magnitude bins of 0.25 mag and then we chose m_P as the lowest magnitude for which the fraction of galaxies was greater than 0.7, obtaining $m_P = 18.2$ and $m_P = 18.5$ for $i+$ and r' images, respectively. Then we apply the photometric selection criteria ($m_P < m_{mag} < m_{MAX} + 0.5$) to the catalogue and we compute β for the whole distribution of galaxies. To take into account the contamination by foreground galaxies given our selection criteria, we set $\beta(z_{phot} < z_{cluster}) = 0$ which outbalances the dilution of the shear signal by these unlensed galaxies.

To estimate the error in $\langle\beta\rangle$ regarding the cosmic variance, we divide CFHTLS Deep1 and COSMOS fields into 25 and 64 non-overlapping areas of ~ 144 arcmin^2 and ~ 160 arcmin^2, respectively. Then we compute $\langle\beta\rangle$ for each area considering the redshift of the analysed clusters according to the used images. The uncertainties in $\langle\beta\rangle$ due to cosmic variance are estimated according to the scatter among the values for each area, obtaining ~ 0.012 for CFHTLS DEEP1 field and ~ 0.010 for COSMOS field. These uncertainties were taken into account in the error estimation of the fitted parameters, and propagated to the resulting system masses.

In order to take into account the contamination of foreground galaxies in the catalogue, we compute for each galaxy the probability that it is behind the cluster. This probability is computed using the described photometric catalogues considering the fraction of galaxies with $z > z_{cluster}$ for each bin in magnitude, mag, and colour (*V* - *i+* and *r'* - *i'*). Hence, given the magnitude and the colour of each galaxy, we assign to it a weight, *w*, as the fraction of galaxies with $z > z_{cluster}$ in that bin. This weight is applied to compute the mass and the 2D density distribution.

5.2.2 Shape measurements.

Gravitational lensing effects are characterised by an isotropic stretching called convergence, κ, and an anisotropic distortion given by the complex-value lensing shear, $\gamma = \gamma_1 + i\gamma_2$. Using the second derivative of the projected gravitational potential to express the shear and convergence, one can show that for a lens with a circular-symmetric projected mass distribution, the tangential component of the shear, γ_T, is related to the convergence through [4]:

$$\gamma_T(r) = \bar{\kappa}(< r) - \langle\kappa\rangle(r), \tag{5.1}$$

where $\bar{\kappa}(< r)$ and $\langle\kappa\rangle(r)$ are the convergence averaged over the disc and circle of radius r, respectively. If lensing is weak ($\kappa \ll 1$), the image of a circular source appears elliptical with a and b as major and minor semi axes, respectively. The induced ellipticity could be directly related to the shear, $\gamma \approx e$. Here we define the ellipticity as a complex number, $e = e_1 + ie_2$, with magnitude $\mid e \mid = (a-b)/(a+b)$ and orientation angle determined by the direction of the major elliptical axis. If the source has an intrinsic ellipticity e_s the observed ellipticity would be $e \sim e_s + \gamma$ [5].

If we consider many sources with intrinsic ellipticities randomly orientated so that $\langle e_s\rangle = 0$, the ensemble average ellipticity after lensing gives an unbiased estimate of the shear: $\langle e\rangle \approx \gamma$, which can be related to the projected mass density of the cluster. Therefore, we can estimate the shear by averaging the shape of background sources that would be affected by the lensing effect.

It is important to take into account the roundness effects due to the atmosphere presence, as well as the distortion caused by the telescope effects. All of these are considered by the *Point Spread Function* (PSF), which is convoluted with the true galaxy intensity light distribution.
To obtain the shape parameters we use `im2shape` [9], which models the galaxies as a sum of Gaussians convoluted with a PSF.

To estimate the PSF at the position of each galaxy we average the shape parameters of the five closest sources classified as stars, since they are considered point-like sources[2].

[2]For more details about this process see the work of *Gonzalez et al. 2015* [19]

Finally to compute the lensing masses and the 2D density distributions we only kept the galaxies classified as background galaxies and with $FWHM > 6pix$, to ensure galaxies with a good pixel sampling and, therefore, with better-shaped parameter estimations. Also, we discard galaxies with $\sigma_e > 0.2$, where σ_e is defined as the error in the measured ellipticity and is computed as $\sigma_e = \sqrt{\sigma_{e_1}^2 + \sigma_{e_2}^2}$, where σ_{e_1} and σ_{e_2} are the errors in the ellipticity components provided by `im2shape`. With these criteria we obtain a density of background galaxies of ~ 17.4 galaxies/arcmin^2 for SUBARU image and ~ 3.5 galaxies/arcmin^2 for CFHT mosaic image.

5.2.3 Estimation of the individual masses of the substructures.

Once we obtain the shape parameters of the background galaxies we can compute the shear profile which could be fitted by assuming a density distribution for each structure identified with the code `MeSsI`. For this we assume spherical models and, averaging the ellipticities in radial intervals, we obtain a shear profile which is fitted by minimising the corresponding χ^2 function which contains the parameters of the model.

Spherical symmetry implies that the average in annular bins of the tangential component ellipticity, e_T, of the lensed galaxies traces the reduced shear. Since galaxies have an intrinsic ellipticity, the error in the shear estimator $\langle\gamma(r)\rangle$, obtained by averaging the tangential ellipticity component of the N galaxies at a distance $r \pm \delta r$ from the centre considered, $\langle e_T(r)\rangle$, would be [36]:

$$\sigma_\gamma \approx \frac{\sigma_\epsilon}{\sqrt{N}}, \tag{5.2}$$

where σ_ϵ is the dispersion of the intrinsic ellipticity distribution.

The brightest galaxy member identified from each structure, according to the membership presented in the catalogue of [14], is considered as the centre to build the profiles. Shear profiles are computed using non-overlapping logarithmic annuli, in order to have a similar signal-to-noise ratio (S/N) in each annuli.

Two mass models are used to fit the resultant shear profiles: a singular isothermal sphere (SIS) and a NFW profile [30]. The SIS profile is the simplest density model for describing a relaxed massive sphere with a constant value for the isotropic one dimensional velocity dispersion, σ_V. This is mainly described by the density distribution:

$$\rho(r) = \frac{\sigma_V^2}{2\pi G r^2} \tag{5.3}$$

where G is the gravitational constant. This model corresponds to a distribution of self-gravitating particles where the velocity distribution at all radii is a Maxwellian with one dimensional velocity dispersion, σ_V. From this equation, we can get the critical Einstein radius for the source sample as:

$$\theta_E = \frac{4\pi\sigma_V^2}{c^2}\langle\beta\rangle, \tag{5.4}$$

in terms of which one obtains:

$$\kappa(\theta) = \gamma(\theta) = \frac{\theta_E}{2\theta}, \tag{5.5}$$

where θ is the angular distance to the cluster centre. Hence, fitting the shear for a different radii, we can estimate the Einstein radius, and from that, we can obtain an estimation of the mass M_{200} as [26]:

$$M_{200} = \frac{2\sigma_V^3}{\sqrt{50}GH(z)}, \tag{5.6}$$

where $H(z)$ is the redshift dependent Hubble parameter.

The NFW profile is derived from fitting the density profile of numerical simulations of cold dark-matter haloes [30]. This profile depends on two parameters, the radius, R_{200}, that encloses a mean density equal to 200 times the critical density ($\rho_{crit} \equiv 3H^2(z)/8\pi G$), and a dimensionless concentration parameter, c_{200}:

$$\rho(r) = \frac{\rho_c \delta_c}{(r/r_s)(1 + r/r_s)^2} \tag{5.7}$$

where r_s is the scale radius, $r_s = R_{200}/c_{200}$ and δ_c is the characteristic overdensity of the halo,

$$\delta_c = \frac{200}{3}\frac{c_{200}^3}{\ln(1 + c_{200}) - c_{200}/(1 + c_{200})} \tag{5.8}$$

We use the lensing equation for the spherical NFW density profile from Wright and Brainerd [42]. If we fit the shear for different radius we can obtain an estimation of the

parameters c_{200} and R_{200}. Once we obtain R_{200} we can compute the M_{200} mass. Nevertheless, there is a well-known degeneracy between the parameters R_{200} and c_{200} when fitting the shear profile in the weak lensing regime. This is due to the lack of information on the mass distribution near the cluster centre, and only a combination of strong and weak lensing can raise it and provide useful constraints on the concentration parameter. Due to the lack of strong lensing modelling in our sample, we follow Kettula et al. [23], van Uitert et al. [39] and Pereira et al. [31], by fixing the concentration parameter according to the relation $c_{200}(M_{200}, z)$ given by Duffy et al. [15]:

$$c_{200} = 5.71(M_{200}/2 \times 10^{12} h^{-1})^{-0.084}(1+z)^{-0.47}. \tag{5.9}$$

The particular choice of this relation does not have a significant impact on the final mass values, with uncertainties dominated by the noise of the shear profiles. Thus we fit the profile with only one free parameter: R_{200}.
To derive the parameters of each mass model profile we perform a standard χ^2 minimisation:

$$\chi^2 = \sum_i^N \frac{(\langle\gamma(r_i)\rangle - \gamma(r_i, p))^2}{\sigma_\gamma^2(r_i)}, \tag{5.10}$$

where the sum runs over the N radial bins of the profile and the model prediction and p refers to either σ_V for the SIS profile, or R_{200} in the case of the NFW model.

5.2.4 Two-dimensional density distribution.

We obtain the 2D projected density distribution for the two analysed galaxy systems by using the `LensEnt2` code [8, 27]. This code applies maximum-entropy method algorithm for reconstructing the 2D density distribution from the two ellipticity components, e_1 and e_2, of the background galaxies and their respective uncertainties.

It is expected that galaxy clusters have smooth, extended projected mass distributions, therefore, an Intrinsic Correlation Function (ICF) is included in the analysis, so the physical projected density distribution, Σ, is expressed as the convolution with a broad kernel,

Galaxy system Id.	α (J2000)	δ (J2000)	z	$\langle\beta\rangle$	SIS σ_V $\mathrm{km\,s^{-1}}$	SIS M_{200} $\mathrm{h_{70}^{-1}}\,10^{14}M_\odot$	NFW c_{200}	NFW R_{200} $\mathrm{h_{70}^{-1}}$ Mpc	NFW M_{200} $\mathrm{h_{70}^{-1}}\,10^{14}M_\odot$
A1204 I	11h 13m 20.5s	+17° 35' 41.0"	0.1703	0.689	750 ± 80	3.7 ± 1.2	3.52	1.41 ± 0.22	3.8 ± 1.8
A1204 II	11h 14m 07.2s	+17° 27' 41.0"	0.1705	0.689	640 ± 100	2.3 ± 1.1	3.52	1.11 ± 0.23	1.9 ± 1.2
A2029	15h 10m 56.1s	+05° 44' 41.7"	0.0775	0.778	840 ± 110	5.3 ± 2.0	3.40	1.95 ± 0.40	9.1 ± 5.6
A2033	15h 11m 26.5s	+06° 20' 56.9"	0.0822	0.766	780 ± 120	4.3 ± 2.1	3.49	1.75 ± 0.36	6.5 ± 4.0

Table 5.2: Main results of the weak gravitational lens analysis. The ID of each system is presented in the 1^{st} column. In the columns (2), (3) and (4) we show the angular coordinates and redshift of the centres used for the analysis. In column (5) we show the geometric factor. In the columns (6) and (7) we shown the results of the fitting for the SIS profile, the velocity dispersion and M_{200} respectively. while in the columns (8), (9) and (10) the results for the fitting of an NFW profile are shown: c_{200}, R_{200} and the estimated mass M_{200} respectively.

given by the ICF.

Finally to obtain the projected density distribution, Σ, it is necessary to compute the critical density defined as:

$$\Sigma_c = \frac{c^2}{4\pi G}\frac{1}{D_L\langle\beta\rangle}, \tag{5.11}$$

where we adopt the parameters D_L and $\langle\beta\rangle$ for the primary component.

5.2.5 Results.

In this section we present the results obtained in the weak lensing analysis together with the X-ray emission data and the dynamic characterisation of the systems.

In Table 5.2 we show the obtained lensing masses for each structure according to the fitted shear profiles (Figs. 5.1 and 5.2). In figures 5.3 and 5.4 we shown the projected density distribution, together with the X-ray contours.

5.3 Dynamic analysis.

As was specified in chapter 3 the dynamic properties of the $A2029/2033$ and $A1204$ clusters were studied using the `MeSsI` code. In this section, we extend the study analysing these systems individually.

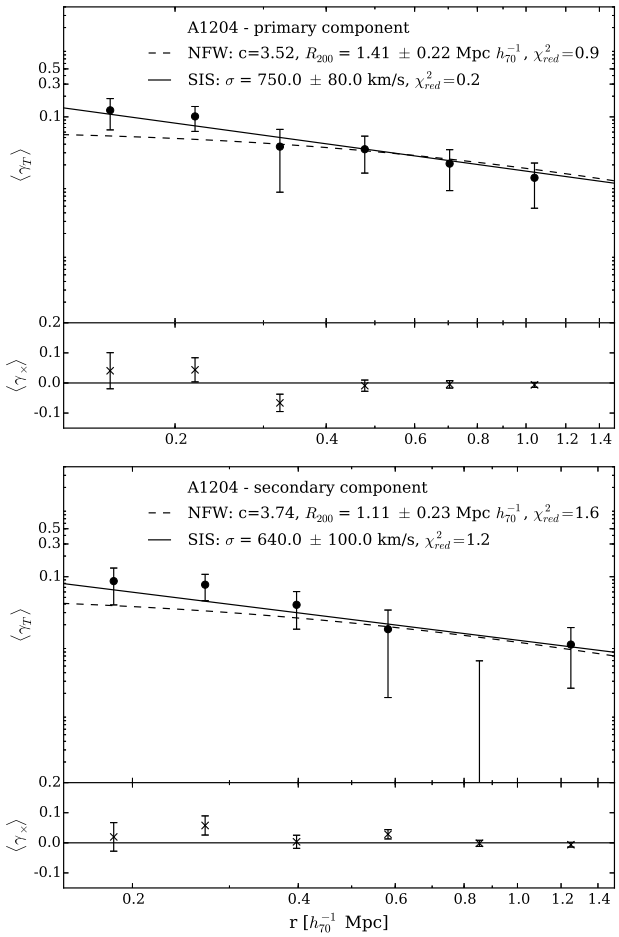

Figure 5.1: $A1204$: Radial shear profile as a function of the projected distance to the centre of the substructure. In the upper panel graph we show the main component, while the secondary component of $A1204$ is shown in the bottom panel. The solid and dotted lines represent the fitted SIS and NFW profiles respectively. The points and crossings show the tangential and cross ellipticity components of the selected background galaxies, averaged in annular bins, respectively.

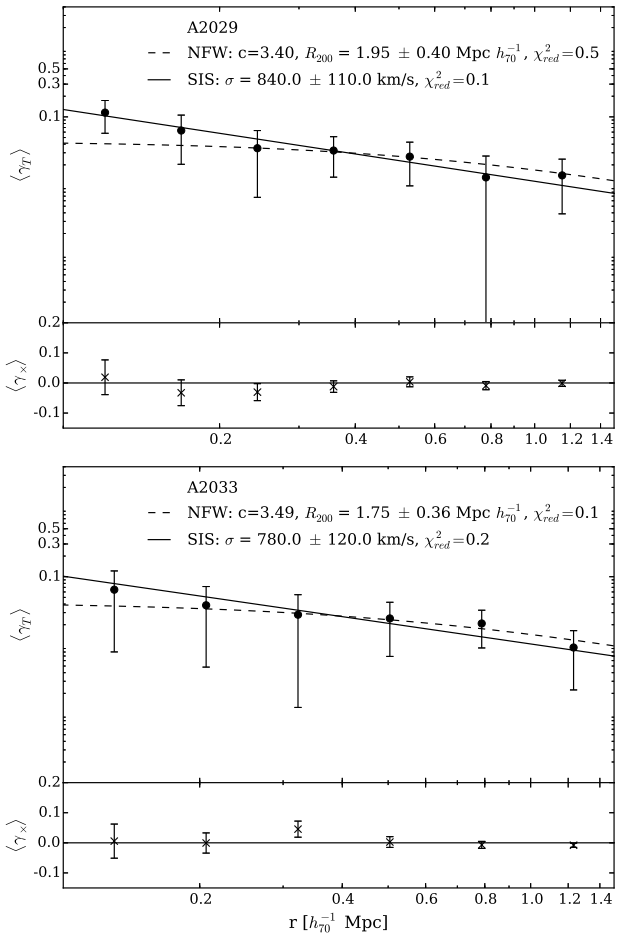

Figure 5.2: $A2029/2033$: Radial shear profile as a function of the projected distance to the center of the substructure. In the upper panel graph we show the main component ($A2029$), while the secondary component ($A2033$) is shown in the bottom panel. The solid and dotted lines represent the fitted SIS and NFW profiles respectively. The points and crossings show the tangential and cross ellipticity components of the selected background galaxies, averaged in annular bins, respectively.

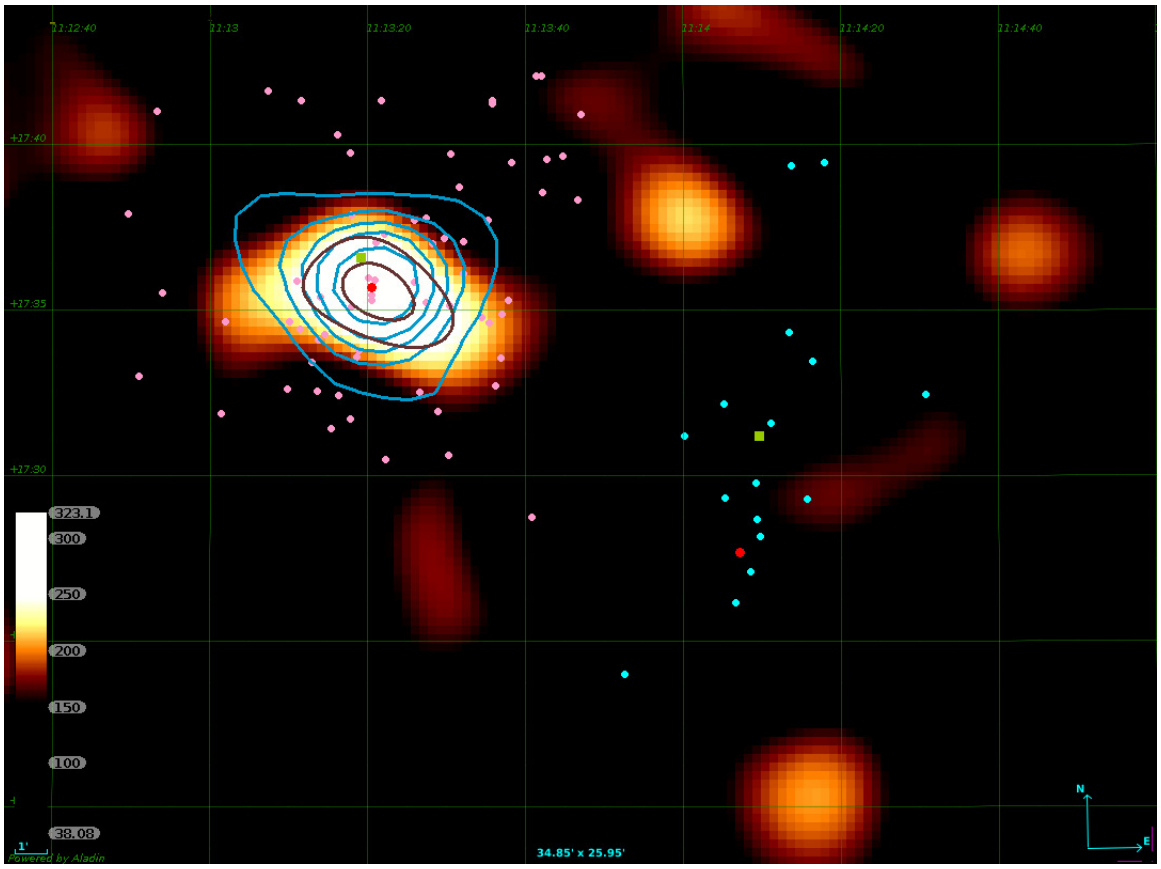

Figure 5.3: Projected density distribution in the field A1204 obtained with the weak lensing analysis. The scale, marked at bottom left, is given in $h_{70} M_{\odot}/pc^2$. Red contours corresponds to a projected density above 3σ significance level (250 $h_{70} M_{\odot}/pc^2$). X-ray contours are plotted in blue, the contour levels are (3, 5, 7, 9 and 12) times the rms noise. Red dots and green squares are the BCGs positions and the dynamical centres, respectively. Pink and light-blue points are the positions of the galaxies classified as members, for the primary and secondary component of A1204, respectively.

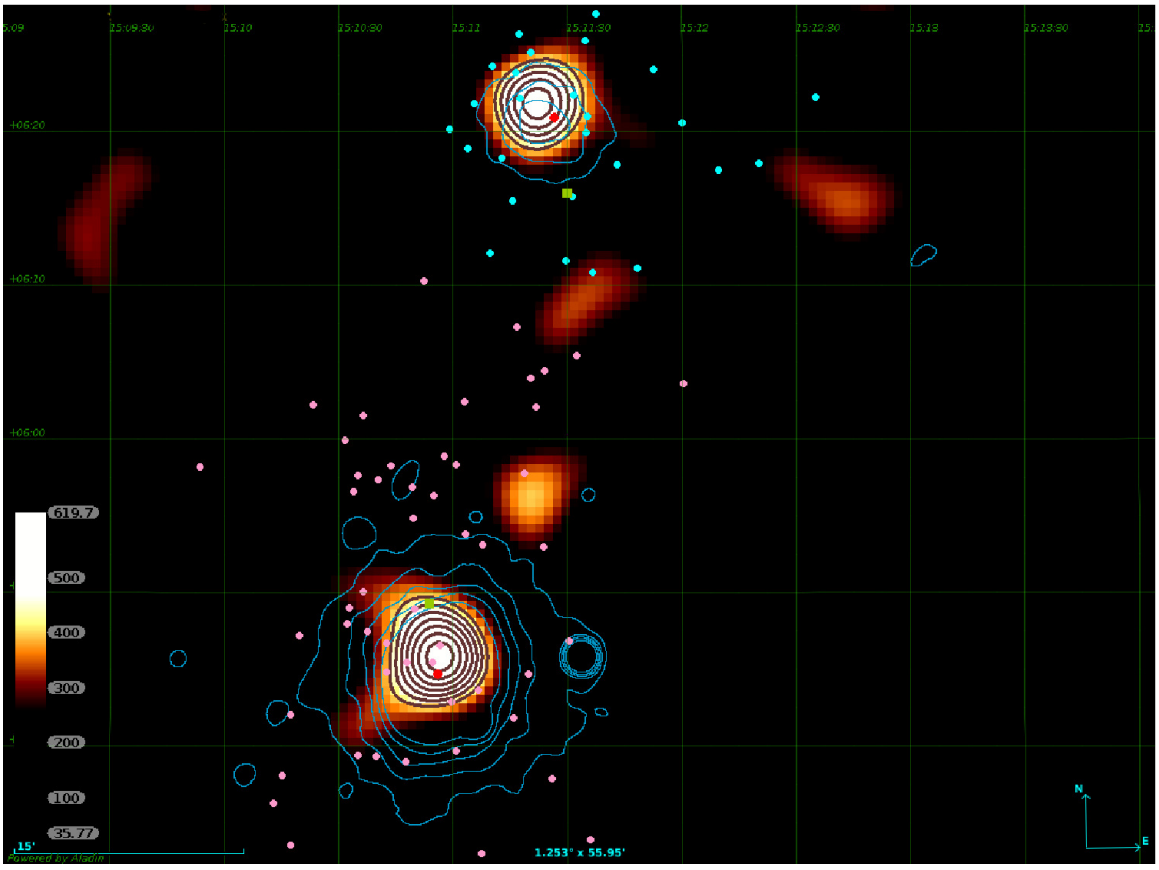

Figure 5.4: Projected density distribution in the field A2029 and A2033 obtained with the weak lensing analysis. The scale, marked at bottom left, is given in $h_{70} M_{\odot}/pc^2$. Red contours corresponds to a projected density above 3σ significance level (400 $h_{70} M_{\odot}/pc^2$). X-ray contours are plotted in blue, the contour levels are (3, 5, 7, 9 and 12) times the rms noise. Red dots and green squares are the BCGs positions and the dynamical centres, respectively. Pink and light-blue points are the positions of the galaxies classified as members, for A2029 and A2033, respectively.

In order to study the stability of the result obtained in chapter 3, we apply the same technique but deleting some randomly selected galaxies and then we analyse how the classification changes with the percentage of dropped galaxies. In particular, we perform 100 realisations for each value of completeness (99%, 95% and 90%).

In the case of the $A1204$ cluster in all cases, we classify the system as a merging cluster in $\sim 50\%$ of the realisations, suggesting that the classification is not stable, therefore if there is an interaction between the components, it will not be strong. On the other hand, for the $A2029/2033$ cluster in all the realisations, regardless of the value of completeness, we find that the system is in an interaction process, indicating a stable classification.

On the other hand it is known that there are clusters that are the interaction between more than 2 substructures [34]. In order to analyse if $A1204$ and/or $A2029/33$ are the product of multiple interactions we apply the `MeSsI` code in an iterative way. In the first instance we studied the clusters as specified in the chapter 3 and we find the substructures that can be in interaction. In a second instance we apply the `MeSsI` code to the two components identified by the algorithm in the first study.

In both cases, $A1204$ and $A2029/2033$, we found that each individual component are classified as not in an interaction process, indicating that these systems are not multiple mergers.

5.3.1 Two-body model.

In order to obtain information regarding the state of evolution of the studied systems, we apply the Newtonian gravitational binding criterion that established that the two-body system is bound if the potential energy of the system is equal to or greater than the kinetic energy. This dynamical model was described in detail by Beers et al. [6] and Gregory and Thompson [21] and, was also applied to the analysis of several bimodal galaxy systems [eg., 2, 10, 11, 22, 43]. This model assumes radial orbits for the identified structures, which start their evolution at time $t_0 = 0$ with separation $R_0 = 0$, and are moving apart or coming together for the first time in their history. With the obtained lensing masses, this method allows us to estimate the probability that (1) the system is bound but still expanding, (2) the system is collapsing, or (3) the two structures are not bound to each other but are merely close together on the sky. It is important to highlight that this model does not

consider the angular moment of the system, the distribution of matter inside each cluster, or the gravitational interaction of the infalling matter outside the cluster pair, since it is assumed that the masses are constant since their formation time [29]. Nevertheless, this model puts another constraint on the dynamical state and the evolution of the systems.

The Newtonian criterion for gravitational binding can be stated in terms of the projected separation, R_p, the radial velocity difference, V_r and the total mass of the combined system, M, as:

$$V_r^2 R_p \leq 2GMsin^2(\alpha)cos(\alpha), \tag{5.12}$$

where α is the projection angle between the plane of the sky and the line connecting the components of the system. We can express the true (3D) velocity, V, and separation, R, as:

$$R_p = R\,cos(\alpha), V_r = V\,sin(\alpha). \tag{5.13}$$

We compute V_r considering the redshifts of the galaxy members of each component within the R_{200} radius, computed according to SIS masses. R_p is computed as the distance between each BCG (which is the adopted centre for the lensing analysis) and the combined mass of the system is obtained by adding the SIS masses of each structure.

Using Beers et al. [6] equations of motion for unbound and bound systems, we can obtain the projected angle α as a function of the radial velocity difference V_r. The equation relates the time (which is assumed for each system as the age of the Universe at the mean redshift of the considered structures), the velocity, V, and the separation, R, with the developmental angle, χ, and the maximum separation, R_m, of the system components for the bound solution and the asymptotic expansion velocity for the unbound solution. The obtained $V_r - \alpha$ relations are plotted in Figure 5.5. The solid black curve separates the bound and unbound solutions according to the Newtonian criterion (Eq. 5.12), the blue curves are the solutions of the equations of motions for bound (solid line) and unbound systems (dashed line) and the solutions for each system are marked with open circles, according to the observed V_r. For both systems the solutions are bound, defined by the intersections between the solutions of the equations of motion and the observed V_r. For each solution, i, we compute the probabilities given by:

Galaxy system	Solution	α [deg]	R [h_{70}^{-1} Mpc]	R_m [h_{70}^{-1} Mpc]	V [km s^{-1}]	P %
A1204	Bound-outgoing	69.8	6.9	7.0	70	10
	Bound-incoming$_a$	66.9	6.1	6.1	70	19
	Bound-incoming$_b$	3.6	2.4	4.4	1000	71
A2029/2033	Bound-outgoing	75.7	13.6	64.6	690	100.

Table 5.3: Solutions for the 2-body model. Columns: (1) Cluster identification. (2) Solution class. Resultant values of α (3), R (4), R_m (5), V (6) and the computed probability P (7) for each solution.

$$p_i = \int_{\alpha_{inf,i}}^{\alpha_{sup,i}} cos(\alpha) d\alpha, \tag{5.14}$$

then we normalise the obtained probability to obtain $P_i = p_i/(\sum_i p_i)$.

Solutions are presented in Table 5.3. For A2029 we obtain one bound-outgoing solution, close to the unbound solution considering the errors in V_r. According to this, the components are expanding separated by a distance of $13.6\, h_{70}^{-1}$ Mpc.

On the other hand, for A1204 we obtain three bound solutions, two incoming and one outgoing. The most probable is the incoming solution for which the components have reached their maximum distance of $4.4\, h_{70}^{-1}$ Mpc and are now separated by $2.4\, h_{70}^{-1}$ Mpc collapsing with a velocity of $1000\,$km s^{-1}.

5.4 Conclusions.

In this chapter we present the results of the study of 2 candidates to merging clusters previously identified with the `MeSsI` algorithm. These results were published in the international journal *Astronomy & Astrophysics* [20].

Although the systems studied here were classified as merging candidates, the results show that there is no sufficient evidence to conclude that these systems are in an interaction process. However, it can not be ruled out that the systems are bounded and will interact in the future.

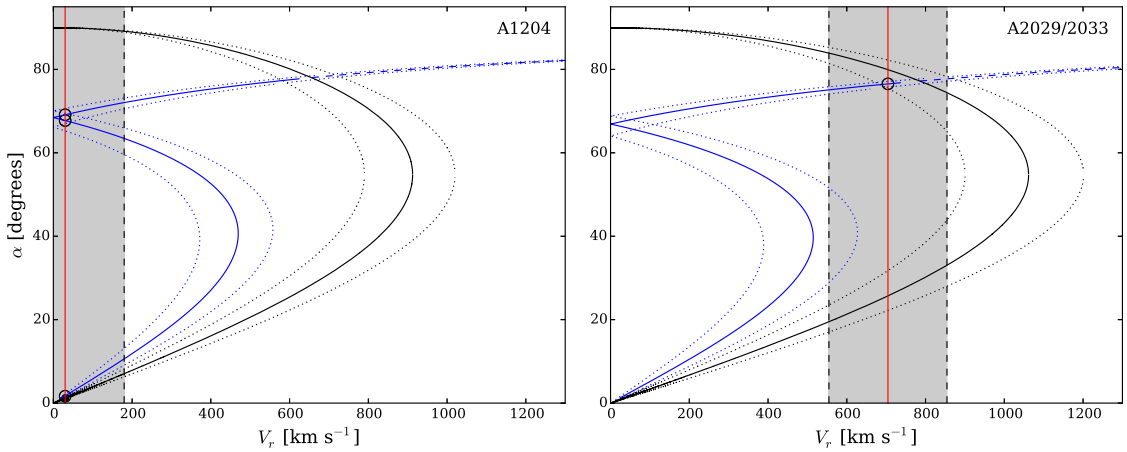

Figure 5.5: Projection angle (α) as a function of the relative radial velocity difference (V_r) computed according to the equations of motion for A1204 (*left panel*) and A2029/2033 (*right panel*). Solid black curves separate the bound and unbound solutions according to the Newtonian criterion. Blue curves are the solutions of the equations of motions for bound (solid line) and unbound systems (dashed line) and the solutions for each system are marked with open circles, according to the observed V_r marked with the red line, where the gray region corresponds to the uncertainty in these values delimited by the dashed black line. Point curves express the uncertainties in the computed curves considering the errors in the adopted lensing masses.

5.4.1 $A1204$.

For the galaxy system $A1204$, through weak lensing analysis, we were able to compute the masses of the substructures previously identified and we obtained a mass ratio of ~ 2. It is worth mentioning that in a previous work *Babyk et al. 2012* [3] estimated the mass for these system by an analysis of the X-rays emission and obtained a value of $M_{200} = 3.18^{+0.34}_{-0.24} \times 10^{14} h_{70}^{-1} M_{\odot}$, which is in agreement with the one found in our work ($M_{200} = 4.0 \pm 1.8 \times 10^{14} h_{70}^{-1} M_{\odot}$).

In the 2D density distribution (Fig. 5.3 we can distinguish only the primary structure. Taking into account the estimated mass by adjusting the *shear* profile to the second component, the density value is below the detection level of 3 σ corresponding to 250 $\mathrm{h}_{70} M_{\odot}/pc^2$.

Also, we do not detect significant X-ray emission in the secondary structure region above the threshold adopted to build the brightness contours. In order to establish if this is due to the observing detection limit, we consider the lowest detected flux of RASS-based catalogues according to *Piffaretti et al. 2011* [32], which contains the lowest X-ray emission clusters identified using this data. This corresponds to a flux F_{lim} = $1.5 \times 10^{-12}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$, in the 0.1-2.4 keV band of the SGP catalogue [13]. Considering this flux and the redshift of the secondary structure, we obtain a limiting luminosity of $1.5 \times 10^{42}\,\mathrm{h}_{70}^{-2}\,\mathrm{erg\,s^{-1}}$. Taking into account the obtained lensing mass, we compute the expected X-ray luminosity for the system according to *Leauthaud et al. 2007* [25] M_L_X relation, obtaining $(3.7 \pm 2.4) \times 10^{43}\,\mathrm{h}_{70}^{-2}\,\mathrm{erg\,s^{-1}}$, only 1.5σ above the detection limit. However, it is important to highlight that a low X-ray emissivity could be explained by a low density of the intra-cluster gas, which might be produced by a past interaction between the structures.

The cluster A1204 has recently been classified as a system with a strong cool core [45] which is in good correspondence with the X-ray contours found in this work. Thus, the cluster does not show any evidence of having suffered a recent merger event. Also, for this cluster the classification as a merging system is unstable as was stated in Section 5.3. Nevertheless, we detect lensing signal for the secondary structure through the lensing profile (Fig. 5.1) from which we obtain the total mass, and, also, the density contours exhibited in Figure 5.3 are elongated in the direction of the secondary component. Fur-

thermore, assuming a relative velocity for these systems similar to the velocities measured in other merging clusters $1000km/s$ [37], the time necessary for the structures to reach the observed distance between the centres of the identified substructures (~ 2.4 Mpc) is ≤ 3 Gyr. Hence, taking into account that there are no other signatures of a collision, it is not likely that this could be a past merger event. Nevertheless this scenario could not be discarded if we were to consider lower velocities for the components; we also do not discard the possibility that the interaction between these structures could be in process.

5.4.2 $A2029/A2033$

For this system we obtain the total mass for both structures which can be identified in the 2D projected density distribution. Both clusters show X-ray emission in good correspondence with the density distribution (Fig. 5.4). The obtained total mass for A2029 ($M_{200} = 9 \pm 6 \times 10^{14} h_{70}^{-1} M_{\odot}$) is in good agreement with that estimated by Walker et al. [41] using X-ray observations ($M_{200} = 10.1 \pm 0.6 \times 10^{14} h_{70}^{-1} M_{\odot}$).

It has a large cD galaxy [38] whose major axis is aligned in the NE to SW direction, in approximately the same direction as that joining it to nearby $A2033$. These two systems, together with $A2028$ and $A2066$ form a small supercluster [16]. Studies examining whether or not these systems are connected by a filament structure show that this is not the case [33, 41]. Therefore, there is no conclusive evidence showing that these systems have interacted in the past or that they are now interacting, given the observed relaxed state of both structures. Nevertheless, if we take into account that they belong to the same supercluster and that the dynamical classification as interacting system is stable (see Sect. 5.3), it could be expected that these clusters would interact in the future.

Bibliography

[1] K. N. Abazajian, J. K. Adelman-McCarthy, M. A. Agueros, S. S. Allam, C. Allende Prieto, D. An, K. S. J. Anderson, S. F. Anderson, J. Annis, N. A. Bahcall, and et al. The Seventh Data Release of the Sloan Digital Sky Survey,. *Astrophys. J. Suppl. Ser.*, 182:543, June 2009. doi: 10.1088/0067-0049/182/2/543.

[2] F. Andrade-Santos, C. Jones, W. R. Forman, S. S. Murray, R. P. Kraft, A. Vikhlinin, R. J. van Weeren, P. E. J. Nulsen, L. P. David, W. A. Dawson, M. Arnaud, E. Pointecouteau, G. W. Pratt, and J.-B. Melin. Chandra and Xmm-Newton Observations of the Bimodal Planck SZ-Detected Cluster Plckg345.40-39.34 (A3716) with High and Low Entropy Subcluster Cores. 803:108, April 2015. doi: 10.1088/0004-637X/803/2/108.

[3] I. Babyk, O. Melnyk, and A. Elyiv. The distribution of dark matter and intracluster gas in galaxy clusters. *Advances in Astronomy and Space Physics*, 2:56–59, March 2012.

[4] M. Bartelmann. Cluster mass estimates from weak lensing. 303:643, November 1995.

[5] M. Bartelmann and P. Schneider. Weak gravitational lensing. 340:291–472, January 2001. doi: 10.1016/S0370-1573(00)00082-X.

[6] T. C. Beers, M. J. Geller, and J. P. Huchra. Galaxy clusters with multiple components. I - The dynamics of Abell 98. 257:23–32, June 1982. doi: 10.1086/159958.

[7] E. Bertin and S. Arnouts. SExtractor: Software for source extraction. 117:393–404, June 1996.

[8] S. L. Bridle, M. P. Hobson, A. N. Lasenby, and R. Saunders. A maximum-entropy method for reconstructing the projected mass distribution of gravitational lenses. 299:895–903, September 1998. doi: 10.1046/j.1365-8711.1998.01877.x.

[9] S. L. Bridle, J.-P. Kneib, S. Bardeau, and S. F. Gull. Bayesian Galaxy Shape Estimation. In P. Natarajan, editor, *The Shapes of Galaxies and their Dark Halos*, pages 38–46, March 2002. doi: 10.1142/9789812778017_0006.

[10] T. Caglar and M. Hudaverdi. Dynamical history of a binary cluster: Abell 3653. 472: 2633–2642, December 2017. doi: 10.1093/mnras/stx2081.

[11] L. Cortese, G. Gavazzi, A. Boselli, J. Iglesias-Paramo, and L. Carrasco. Multiple merging in the Abell cluster 1367. 425:429–441, October 2004. doi: 10.1051/0004-6361:20040381.

[12] J. Coupon, O. Ilbert, M. Kilbinger, H. J. McCracken, and et al. Photometric redshifts for the CFHTLS T0004 deep and wide fields. 500:981–998, June 2009. doi: 10.1051/0004-6361/200811413.

[13] R. Cruddace, W. Voges, H. Böhringer, C. A. Collins, A. K. Romer, H. MacGillivray, D. Yentis, P. Schuecker, H. Ebeling, and S. De Grandi. The ROSAT All-Sky Survey: a Catalog of Clusters of Galaxies in a Region of 1 steradian around the South Galactic Pole. 140:239–264, June 2002. doi: 10.1086/324519.

[14] M. de los Rios, M. J. Domínguez R., D. Paz, and M. Merchán. The MeSsI (merging systems identification) algorithm and catalogue. *MNRAS*, 458:226–232, May 2016. doi: 10.1093/mnras/stw215.

[15] A. R. Duffy, J. Schaye, S. T. Kay, and C. Dalla Vecchia. Erratum: Dark matter halo concentrations in the Wilkinson Microwave Anisotropy Probe year 5 cosmology. 415:L85–L85, July 2011. doi: 10.1111/j.1745-3933.2011.01080.x.

[16] M. Einasto, J. Einasto, E. Tago, V. V. Müller, and H. Andernach. Optical and X-Ray Clusters as Tracers of the Supercluster-Void Network. I. Superclusters of Abell and X-Ray Clusters. 122:2222–2242, November 2001. doi: 10.1086/323707.

[17] D. Fabricant, R. Fata, J. Roll, E. Hertz, N. Caldwell, T. Gauron, J. Geary, B. McLeod, A. Szentgyorgyi, J. Zajac, M. Kurtz, J. Barberis, H. Bergner, W. Brown, M. Conroy, R. Eng, M. Geller, R. Goddard, M. Honsa, M. Mueller, D. Mink, M. Ordway, S. Tokarz, D. Woods, W. Wyatt, H. Epps, and I. Dell'Antonio. Hectospec, the MMT's 300 Optical Fiber-Fed Spectrograph. *Pub. Astron. Soc. Pac.*, 117:1411–1434, December 2005. doi: 10.1086/497385.

[18] S. R. Furlanetto and A. Loeb. Constraining the Collisional Nature of the Dark Matter through Observations of Gravitational Wakes. 565:854–866, February 2002. doi: 10.1086/324693.

[19] E. J. Gonzalez, G. Foëx, J. L. Nilo Castellón, M. J. Domínguez Romero, M. V. Alonso, D. García Lambas, O. Moreschi, and E. Gallo. Low X-ray luminosity galaxy clusters - III. Weak lensing mass determination at $0.18 < z < 0.70$. 452:2225–2235, September 2015. doi: 10.1093/mnras/stv787.

[20] E. J. Gonzalez, M. de los Rios, G. A. Oio, D. H. Lang, T. A. Tagliaferro, M. J. Domínguez R., J. L. N. Castellón, H. Cuevas L., and C. A. Valotto. Analysis of candidates for interacting galaxy clusters. I. A1204 and A2029/A2033. *A&A*, 611: A78, April 2018. doi: 10.1051/0004-6361/201732003.

[21] S. A. Gregory and L. A. Thompson. The A2197 and A2199 galaxy clusters. 286: 422–436, November 1984. doi: 10.1086/162617.

[22] H. S. Hwang and M. G. Lee. Galaxy activity in merging binary galaxy clusters. 397: 2111–2122, August 2009. doi: 10.1111/j.1365-2966.2009.15100.x.

[23] K. Kettula, S. Giodini, E. van Uitert, H. Hoekstra, A. Finoguenov, M. Lerchster, T. Erben, C. Heymans, H. Hildebrandt, T. D. Kitching, A. Mahdavi, Y. Mellier, L. Miller, M. Mirkazemi, L. Van Waerbeke, J. Coupon, E. Egami, L. Fu, M. J. Hudson, J. P. Kneib, K. Kuijken, H. J. McCracken, M. J. Pereira, B. Rowe, T. Schrabback, M. Tanaka, and M. Velander. CFHTLenS: weak lensing calibrated scaling relations for low-mass clusters of galaxies. 451:1460–1481, August 2015. doi: 10.1093/mnras/stv923.

[24] C. Laigle, H. J. McCracken, O. Ilbert, B. C. Hsieh, I. Davidzon, P. Capak, G. Hasinger, J. D. Silverman, C. Pichon, J. Coupon, H. Aussel, D. Le Borgne, K. Caputi, P. Cassata, Y.-Y. Chang, F. Civano, J. Dunlop, J. Fynbo, J. S. Kartaltepe, A. Koekemoer, O. Le Fèvre, E. Le Floc'h, A. Leauthaud, S. Lilly, L. Lin, S. Marchesi, B. Milvang-Jensen, M. Salvato, D. B. Sanders, N. Scoville, V. Smolcic, M. Stockmann, Y. Taniguchi, L. Tasca, S. Toft, M. Vaccari, and J. Zabl. The COSMOS2015 Catalog: Exploring the $1 < z < 6$ Universe with Half a Million Galaxies. 224:24, June 2016. doi: 10.3847/0067-0049/224/2/24.

[25] A. Leauthaud, R. Massey, J.-P. Kneib, J. Rhodes, D. E. Johnston, P. Capak, C. Heymans, R. S. Ellis, A. M. Koekemoer, O. Le Fèvre, Y. Mellier, A. Réfrégier, A. C. Robin, N. Scoville, L. Tasca, J. E. Taylor, and L. Van Waerbeke. Weak Gravitational Lensing with COSMOS: Galaxy Selection and Shape Measurements. 172:219–238, September 2007. doi: 10.1086/516598.

[26] A. Leonard and L. J. King. A new tool to determine masses and mass profiles using gravitational flexion. 405:1854–1866, July 2010. doi: 10.1111/j.1365-2966.2010.16573.x.

[27] P. J. Marshall, M. P. Hobson, S. F. Gull, and S. L. Bridle. Maximum-entropy weak lens reconstruction: improved methods and application to data. 335:1037–1048, October 2002. doi: 10.1046/j.1365-8711.2002.05685.x.

[28] S. Miyazaki, Y. Komiyama, M. Sekiguchi, S. Okamura, M. Doi, H. Furusawa, M. Hamabe, K. Imi, M. Kimura, F. Nakata, N. Okada, M. Ouchi, K. Shimasaku, M. Yagi, and N. Yasuda. Subaru Prime Focus Camera – Suprime-Cam. 54:833–853, December 2002. doi: 10.1093/pasj/54.6.833.

[29] R. S. Nascimento, A. L. B. Ribeiro, M. Trevisan, E. R. Carrasco, H. Plana, and R. Dupke. Dynamical analysis of the cluster pair: A3407 + A3408. 460:2193–2206, August 2016. doi: 10.1093/mnras/stw1114.

[30] J. F. Navarro, C. S. Frenk, and S. D. M. White. A Universal Density Profile from Hierarchical Clustering,. *ApJ*, 490:493, December 1997. doi: 10.1086/304888.

[31] M. E. S. Pereira, M. Soares-Santos, M. Makler, J. Annis, H. Lin, A. Palmese, A. Z. Vitorelli, B. Welch, G. B. Caminha, T. Erben, B. Moraes, and H. Shan. Weak-lensing calibration of a stellar mass-based mass proxy for redMaPPer and Voronoi Tessellation clusters in SDSS Stripe 82. *ArXiv e-prints*, August 2017.

[32] R. Piffaretti, M. Arnaud, G. W. Pratt, E. Pointecouteau, and J.-B. Melin. The MCXC: a meta-catalogue of x-ray detected clusters of galaxies. 534:A109, October 2011. doi: 10.1051/0004-6361/201015377.

[33] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. Balbi, A. J. Banday, and et al. Planck intermediate results. VIII. Filaments between interacting clusters. 550:A134, February 2013. doi: 10.1051/0004-6361/201220194.

[34] B. Ragozzine, D. Clowe, M. Markevitch, A. H. Gonzalez, and M. Bradač. Weak-lensing Results for the Merging Cluster A1758. *ApJ*, 744:94, January 2012. doi: 10.1088/0004-637X/744/2/94.

[35] K. Rines, M. J. Geller, A. Diaferio, and M. J. Kurtz. Measuring the Ultimate Halo Mass of Galaxy Clusters: Redshifts and Mass Profiles from the Hectospec Cluster Survey (HeCS). *ApJ*, 767:15, April 2013. doi: 10.1088/0004-637X/767/1/15.

[36] P. Schneider, L. King, and T. Erben. Cluster mass profiles from weak lensing: constraints from shear and magnification information. 353:41–56, January 2000.

[37] R. Thompson and K. Nagamine. Pairwise velocities of dark matter haloes: a test for the Λ, cold dark matter model using the bullet cluster,. *MNRAS*, 419:3560–3570, February 2012. doi: 10.1111/j.1365-2966.2011.20000.x.

[38] J. M. Uson, S. P. Boughn, and J. R. Kuhn. Diffuse light in dense clusters of galaxies. I - R-band observations of Abell 2029. 369:46–53, March 1991. doi: 10.1086/169737.

[39] E. van Uitert, H. Hoekstra, T. Schrabback, D. G. Gilbank, M. D. Gladders, and H. K. C. Yee. Constraints on the shapes of galaxy dark matter haloes from weak gravitational lensing. 545:A71, September 2012. doi: 10.1051/0004-6361/201219295.

[40] W. Voges, B. Aschenbach, T. Boller, H. Bräuninger, U. Briel, W. Burkert, K. Dennerl, J. Englhauser, R. Gruber, F. Haberl, G. Hartner, G. Hasinger, M. Kürster, E. Pfeffermann, W. Pietsch, P. Predehl, C. Rosso, J. H. M. M. Schmitt, J. Trümper, and H. U. Zimmermann. The ROSAT all-sky survey bright source catalogue. *A&A*, 349: 389–405, September 1999.

[41] S. A. Walker, A. C. Fabian, J. S. Sanders, M. R. George, and Y. Tawara. X-ray observations of the galaxy cluster Abell 2029 to the virial radius. *MNRAS*, 422: 3503–3515, June 2012. doi: 10.1111/j.1365-2966.2012.20860.x.

[42] C. O. Wright and T. G. Brainerd. Gravitational Lensing by NFW Halos. 534:34–40, May 2000. doi: 10.1086/308744.

[43] P.-F. Yan, Q.-R. Yuan, L. Zhang, and X. Zhou. Multicolor Photometry of the Merging Galaxy Cluster A2319: Dynamics and Star Formation Properties. 147:106, May 2014. doi: 10.1088/0004-6256/147/5/106.

[44] D. G. York, J. Adelman, J. E. Anderson, Jr., S. F. Anderson, J. Annis, and et al. The Sloan Digital Sky Survey: Technical Summary. 120:1579–1587, September 2000. doi: 10.1086/301513.

[45] C. Zhang, H. Xu, Z. Zhu, W. Li, D. Hu, J. Wang, J. Gu, L. Gu, Z. Zhang, C. Liu, J. Zhu, and X.-P. Wu. A Chandra Study of the Image Power Spectra of 41 Cool Core and Non-cool Core Galaxy Clusters. 823:116, June 2016. doi: 10.3847/0004-637X/823/2/116.

Chapter 6

Individual study of merging cluster candidate $A267$.

As it was introduced in chapter 3 in a ΛCDM universe the structure formation proceeds in a hierarchical way, i.e. the smaller structures are formed first and then, through their mergers, the larger structures are formed. In this context, we develop the software `MeSsI`, which identifies candidates for merging clusters. In this chapter, we present a spectroscopic, dynamic and morphological analysis of the candidate to merging cluster $A267$.

This cluster, located at $z \sim 0.230$, has a mass of $M_{200} = 3.23^{+0.82}_{-0.69} h^{-1} 10^{14} M_{\odot}$ [21], a radius of $R_{200} = 1.19$ Mpc [23] and was classified as not relaxed by various authors [2, 22, 27]. It presents an offset of $\sim 33 kpc$ between the position of the X-ray emission peak and the BCG and an offset of $\sim 90 kpc$ between the X-ray peak and the mass distribution peak according to the gravitational lens map [22, 27]. On the other hand, $A267$ presents a cD galaxy in the centre [6], an elliptic X-ray emission distribution [2, 16, 27] and a temperature of $8.7 keV$ [5]. Also *Kale et al.* [18] and *Pratt et al.* [22], taking into account the radio emission, arrive to the conclusion that it does not have a cold core. It is important to note that this system was classified as a fossil group [12, 31] with a magnitude gap between the first and the second galaxy of $\Delta m_{12} > 2.12$. Considering that many works predict that fossil groups were formed at high redshift ($z > 1$) [7, 9, 10] it is expected that these groups will be in a relaxed dynamic state and will not exhibit a high rate of substructures. Nevertheless, *Zarattini et al.* [32] find a high fraction of substructures in a sample of real fossil groups. In this context we study the probability of finding fossil

groups which have had a recent major merger in numerical simulations.

It should be noted that this work was carried out in collaboration with Dra. Elizabeth Gonzales (weak lensing analysis), Lic. María José Kanagusuku (fossil groups analysis), Dr. Gabriel Oio (X-ray emission analysis), Dr. Mariano Dominguez and Dr. Carlos Valotto (statistical analysis), Dr. Nilo Castellano, Lic. Tania Tagliaferro, Dr. Héctor Cuevas and Lic. Daniel Lang (data reduction).

6.1 Data Acquisition.

6.1.1 Spectroscopic and photometric catalogues.

In order to carry out the dynamic analysis, we use the spectroscopic survey Hectospec [23].

On the other hand, for the photometric analysis we use data from the V and $i+$ filters of the Suprime-cam [19] mounted on the $8.2m$ Subaru telescope. It is worth mentioning that these are public data and can be downloaded from the SMOKA archive (*Subaru Mitaka Okayama Kiso Archive*).

The data reduction was carried out using the software `THELI` [24] following the procedure described in *Schirmer et al. 2015* [25]. In order to detect galaxies and eliminate spurious detections, we use the `SExtractor` software [3]. In addition, to determine the zero point of the photometry, we use the sources classified as stars (following the procedure described in *Castellón et al. 2016* [4]) and compare them with the SDSS-DR12 survey [1]. Aperture photometry for stars in the $i+$ (V) bands are compared to SDSS aperture magnitudes, after applying the magnitude system transformation described by Yagi et al. [30] to obtain g and r SDSS magnitudes from the Subaru's V and $i+$ magnitudes.

6.1.2 X-ray observations and data reduction.

In order to analyse the spatial distribution of the gas present in the ICM, we studied the X-ray emission using the data from the public archive of the *Chandra* telescope. This cluster was observed (Obsid 3580) with a exposure time of $20,143Ks$ with the ACIS-I CCD (*Advanced CCD Imaging Spectrometer array*) [13]. The reduction of the data was carried out using the `CIAO` (Chandra Interactive Analysis of Observations) software

version `4.7.8`. In addition, the contamination sources were detected and removed using the code `wavdetect` belonging to `CIAO`.

6.1.3 Simulated data.

In order to study the probability that a fossil group had a recent major merger, we analysed data from the semianalytic model for galaxy formation developed by *Guo et al.* [14] applied to the *Millenium I* cosmological simulation [29]. This simulation follows the evolution of 2160^3 dark matter particles in a volume of $500Mpch^{-1}$ in a ΛCDM universe with cosmological parameters consistent with those reported by WMAP1 [28]($\Omega_m = 0.25$, $\Omega_b = 0.045$, $\Omega_\Lambda = 0.75$, $h = 0.73$, $\Omega_k = 0$, $n_s = 1$ y $\sigma_8 = 0.9$)

6.2 Dynamic and morphological analysis.

As was specified previously, the $A267$ cluster was classified as a candidate for a merging cluster using the `MeSsI` code [8]. This software found that the most likely scenario was that the interaction was in the line of sight. The figure 6.1 shows the 2 interacting substructures found by `MeSsI`. As can be seen, the main substructure is composed by the central galaxies of the cluster ($< R_{200}$) and has an average redshift of $\langle z \rangle = 0.228$, while the secondary substructure is composed by the external galaxies ($> R_{200}$) and has an elongated shape with redshift from $z \approx 0.225$ to $z \approx 0.240$.

In order to quantify the stability of these results, we made 100 realisations of the galaxy cluster by randomly discarding the 10% of the member galaxies in each realisation. Then, using the software `MeSsI`, we analyse each one to determine its dynamic state. We found that the system is classified as interaction in all random samples, indicating that the classification is stable.

Focusing our analysis on the main component of the system, we study the dynamic state by restricting the analysis to galaxies within R_{200}. We found that this component has a high probability ($\approx 60\%$) of being in fusion. However, when performing a stability analysis similar to the one previously described, we found that only 47 of the 100 random realisations of this component is classified as in interaction. This indicates that such classification is not stable.

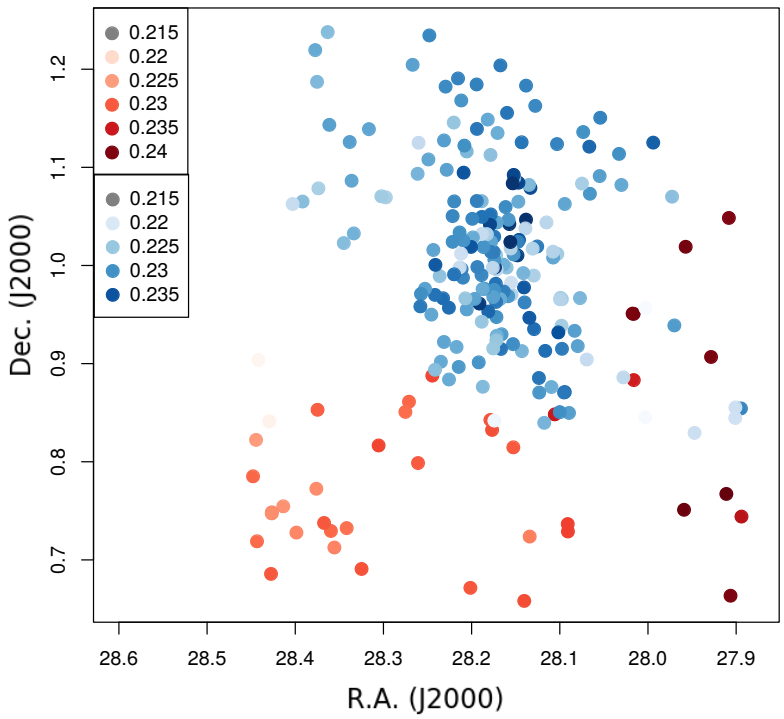

Figure 6.1: Distribution of the member galaxies of the substructures identified by `MeSsI`. The color scale shows the redshift of each galaxy.

Continuing with the analysis of the central component, we performed the *Dressler-Shectman* test [11] (See section 3.2.2) to the galaxies of this component. In figure 6.2 we plot the positions of the galaxies within R_{200}. We show in colours the radial velocity difference with respect to the average of the cluster, while the size of the circles represents the δ statistics of each galaxy.

We found a $\Delta/ngal = 1.2704$ parameter for the cluster, which implies a value $p = 0.022$ estimated by the analysis of 1000 Monte Carlo realisations. It can be seen that there is a substructure to a redshift smaller than the cluster average to the west of the centre.

Taking into account these results, and in order to identify the substructures present within R_{200}, we made a 3-dimensional Mixture of Gaussians (See section 2.2.1) (angular coordinates + redshift) weighting each galaxy by its luminosity. For this we use the `mclust` code of the language `R` [26]. We find 2 substructures that match those identified by the *Dressler-Shectman* test.

Using the spectroscopic data of the member galaxies of each substructure, we estimate

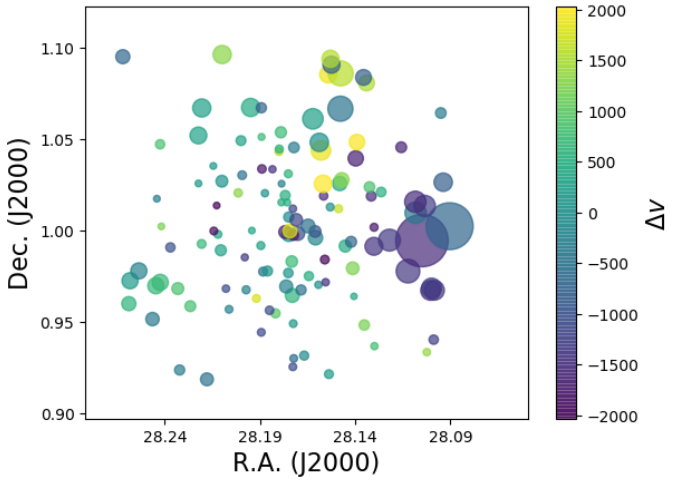

Figure 6.2: Results of the Dressler-Shectman test for the $A267$ core. We plot the positions of the galaxies within a box of $R_{200} \times R_{200}$, with the colourbar and size coding their velocities and their δ parameter, respectively.

Table 6.1: Main characteristics of the substructures identified within R_{200}.

Str.	n	z	σ_V (km s^{-1})	M_{vir} ($\times 10^{14}$ M$\odot$)	M_{200} ($\times 10^{14}$ M$\odot$)
1	100	0.2308	832 ± 77	4.93 ± 0.91	4.88 ± 1.35
2	25	0.2236	306 ± 57	0.45 ± 0.17	0.24 ± 0.14

their virial masses:

$$M_{vir} = \frac{3\pi}{2}\frac{\sigma_V^2 R_H}{G}, \; R_H = \frac{N(N-1)}{2}\frac{1}{\sum_{j>i}|r_{ij}|^{-1}} \tag{6.1}$$

where $|r_{ij}|$ is the distance between the galaxies i and j, R_H is the mean harmonic radius and σ_V is the velocity dispersion estimated by the *gapper* method [15].

In table 6.1 we present the main characteristics of both substructures.

The figure 6.3 shows the galaxies of both substructures in phase space, where Δv is the peculiar velocity of each galaxy estimated as $\Delta v_i = c(z_i - z_1)/(1 + z_c)$, where z_i, z_1 and z_c correspond to the redshift of the galaxy i, the redshift of the main substructure and the redshift of the cluster ($z_c = 0.23$), while σ corresponds to the velocity dispersion of the main substructure. On the other hand, the black dotted lines correspond to the constant values of $(\Delta/\sigma) \times (R/R_{200})$, defined by *Noble et al. 2013* [20], that divide the phase

space into a virialized zone (lower zone), a mixed area (region between both curves) and an infalling zone corresponding to the external region.

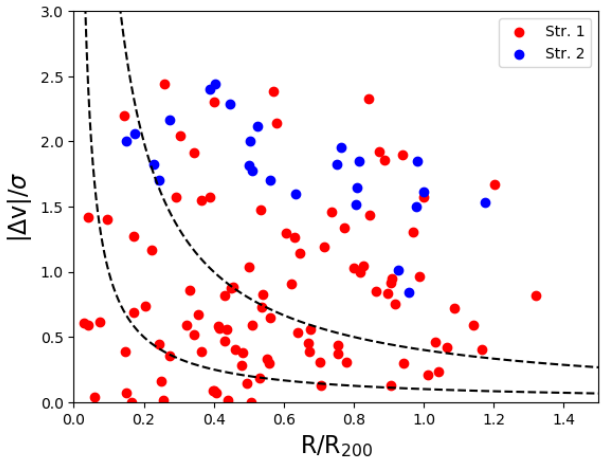

Figure 6.3: Phase-space diagram for A267. Red and blue circles represent the two structures found with MCLUST. Dashed lines correspond to constant values of $(|\Delta(v)|/\sigma_1) \times (r/r_{200})$ following *Noble et al. 2013* [20]

As can be seen, the galaxies that are members of the main substructure are in the virialized zone, whereas the member galaxies of the secondary substructure are mostly in the area corresponding to infalling. This reinforces the hypothesis that this system presents a dynamic state that is not relaxed and is in a merging process.

On the other hand, taking into account the photometric magnitudes according to what is described in the section 6.1, we plot in the left panel of the figure 6.4 the colour-magnitude diagram for the central region of the cluster. In red and blue points are distinguished the galaxies belonging to the 2 identified substructures, in black points are plotted other galaxies that are found at the same redshift and in black squares are highlight those galaxies that are within $<= 0.5R_{200}$ from the centre of the cluster. It is important to note that the difference in magnitudes between the first and the second galaxy within $0.5R_{200}$ is $\Delta mag_{12} = 2.279$, which confirms the classification of this system as a fossil group.

On the right panel of Fig. 6.4 we show the colour histogram for both structures. There is a tendency for red galaxies belonging to the main structure to have a redder colour than those from the infalling structure.

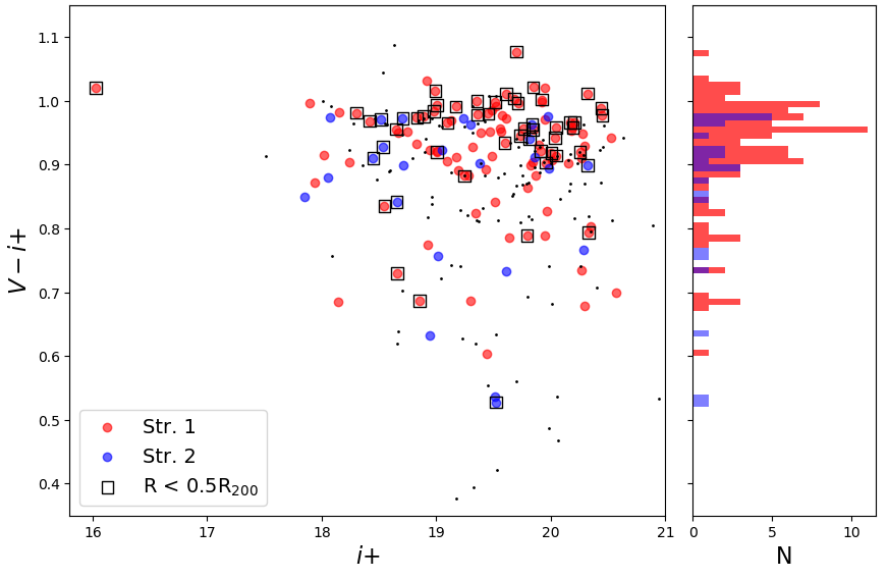

Figure 6.4: *Left*: CMD diagram for A267, where the red, blue and black dots represent galaxies belonging to the main substructure, the infalling substructure and the outskirts, respectively. Black squares enclose galaxies with distances to the cluster's centre $< 0.5R_{200}$. *Right*: Colour histogram of the structures.

To study the ICM gas distribution, we analysed the maps of the X-rays brightness contours. Convergence map (made by *Okabe et al.* [21]) together with X-ray brightness contours and galaxies of each identified substructure are plotted in figure 6.5. Although there are no substructures in the projected density map, it can be seen that there are 2 peaks in the X-ray emission contours aligned in the NE-SO direction. It is important to note that the elongation of the density map and the distribution of galaxies of the main component identified by `MeSsI` are in the same direction.

6.3 Fossil groups in cosmological simulations.

Using the simulation described in 6.1 we identify FGs at $z = 0$ considering two different criteria. On the one hand we consider the criterion of *Jones et al. 2003* [17] that establishes that a fossil group (FG_{12}) must have a difference of magnitude between the first and the second brightest galaxies of more than $\Delta m_{12} > 2.0$. On the other hand, we consider the criterion established by *Dariush et al. 2010* [7] which states that fossil groups (FG_{14})

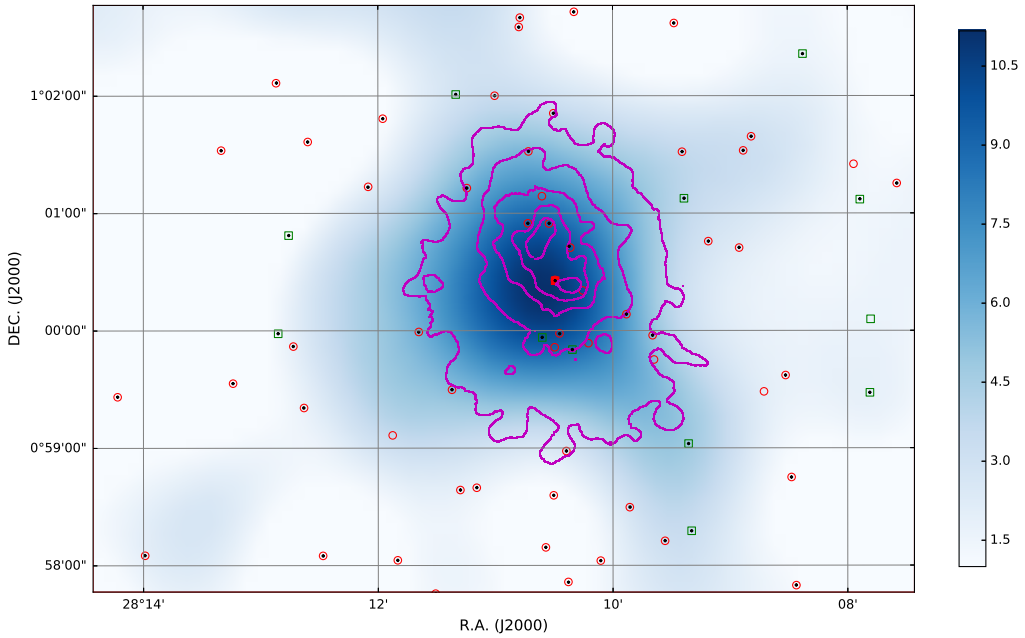

Figure 6.5: Projected density map made by *Okabe et al.* [21] superimposed to the X-ray emission contours. In red circles and green squares are plotted the galaxies belonging to the substructures identified by *MeSsI* in this region.

must have a difference of magnitude between the first and the fourth brightest galaxies of more than $\Delta m_{14} > 2.5$. In both criteria only those galaxies within $0.5 \times R_{200}$ are taken into account. In addition, taking into account that fossil groups must have an X-ray emission, we restrict ourselves to studying haloes with masses greater than $M_{FG} = 10^{13.5} M_{\odot} h^{-1}$. With these criteria we obtain a final sample of 2146 FF_{12} and 1130 FG_{14}.

We study the merger history of the identified FGs and we determine the redshift, z_{merger}, defined as the redshift at which the group pass through its last major merger event in which its total mass is increased by more than a factor $\Delta M = 4/3$. This corresponds to a merger between the original fossil group and another system with more than $1/3$ of the total mass of the fossil group at this redshift. Results for FG_{12} and FG_{14} are in good agreement and show that $\sim 43\%$ have the last major merger event at $z_{merger} < 0.8$.
We also analyse the fraction of groups that had a recent major merger ($z_{merger} < 0.3$), obtaining that 25% of the identified fossil groups satisfies this condition. If we consider only massive groups ($M_{FG} > 10^{14}$ $M_{\odot}$ h^{-1}), the fraction of groups that had a recent merger event is reduced to a 15%. This could be observed in Figure 6.6 (*upper panel*), in which

the median M_{FG} computed for each z_{merger} is obtained. As it can be seen, massive systems tend to present larger z_{merger} values. Moreover, this trend is also observed for ΔM (see figure 6.6 *upper panel*), suggesting that important merger events are more likely at larger redshifts.

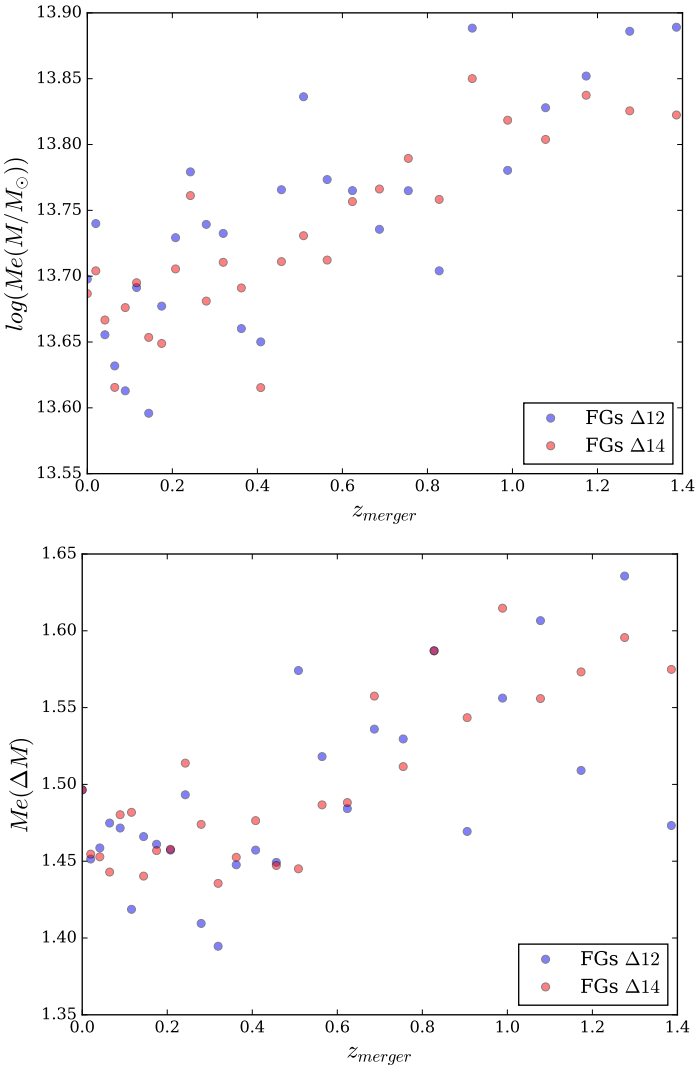

Figure 6.6: Results obtained from simulated data. *Upper panel:* Logarithmic of the median mass of the FGs at $z = 0$ related to the redshift of the last major merger event, z_{merger}. *Lower panel:* Median factor of the mass increment after the last merger event as a function of z_{merger}. In blue are shown FG_{12} and in red FG_{14}.

6.4 Conclusions.

In this work, a dynamic and morphological study of the $A267$ cluster was carried out. This system was classified as a fossil group and as a non-relaxed group by different authors. However, it is expected that fossil groups will have a relaxed dynamical state and few substructures.

In the first part of the work we performed a dynamic study of the system using the software `MeSsI` in which it was possible to confirm the presence of 2 major substructures. Continuing with the dynamic study, we focus our analysis on the main substructure, which is formed by galaxies within R_{200}. Analysing the spectroscopic and morphological data of the galaxies within this region we found 2 substructures, one of which is in the virialized region of the cluster while the other is in the region of infalling (figure 6.3).

The presence of substructures within R_{200} was also confirmed by the bimodal distribution of the X-ray emission and the elongation of the gravitational lens contours estimated by *Okabe et al.* [21].

In order to study the probability that a fossil group presents substructures, we analyse the merger trees of 2 samples of fossil groups constructed from the semianalytical model of *Guo et al.* [14] applied to the data of the *Millenium* cosmological simulation [29]. We find that approximately 15% of the massive fossil groups had a recent major merger ($z_{merger} < 0.3$), whereas if we analyse the smaller groups ($< 10^{14} M_{\odot} h^{-1}$) the fraction increases up to 25%. Considering these results is to be expected that some fossil groups have a significant fraction of substructures, being the $A267$ cluster a particular case of this phenomenon.

Bibliography

[1] Shadab Alam, Franco D. Albareti, Carlos Allende Prieto, F. Anders, Scott F. Anderson, Timothy Anderton, Brett H. Andrews, Eric Armengaud, Éric Aubourg, Stephen Bailey, Sarbani Basu, Julian E. Bautista, Rachael L. Beaton, Timothy C. Beers, Chad F. Bender, Andreas A. Berlind, Florian Beutler, Vaishali Bhardwaj, Jonathan C. Bird, Dmitry Bizyaev, Cullen H. Blake, Michael R. Blanton, Michael Blomqvist, John J. Bochanski, Adam S. Bolton, Jo Bovy, A. Shelden Bradley, W. N. Brandt, D. E. Brauer, J. Brinkmann, Peter J. Brown, Joel R. Brownstein, Angela Burden, Etienne Burtin, Nicolás G. Busca, Zheng Cai, Diego Capozzi, Aurelio Carnero Rosell, Michael A. Carr, Ricardo Carrera, K. C. Chambers, William James Chaplin, Yen-Chi Chen, Cristina Chiappini, S. Drew Chojnowski, Chia-Hsun Chuang, Nicolas Clerc, Johan Comparat, Kevin Covey, Rupert A. C. Croft, Antonio J. Cuesta, Katia Cunha, Luiz N. da Costa, Nicola Da Rio, James R. A. Davenport, Kyle S. Dawson, Nathan De Lee, Timothée Delubac, Rohit Deshpande, Saurav Dhital, Letícia Dutra-Ferreira, Tom Dwelly, Anne Ealet, Garrett L. Ebelke, Edward M. Edmondson, Daniel J. Eisenstein, Tristan Ellsworth, Yvonne Elsworth, Courtney R. Epstein, Michael Eracleous, Stephanie Escoffier, Massimiliano Esposito, Michael L. Evans, Xiaohui Fan, Emma Fernández-Alvar, Diane Feuillet, Nurten Filiz Ak, Hayley Finley, Alexis Finoguenov, Kevin Flaherty, Scott W. Fleming, Andreu Font- Ribera, Jonathan Foster, Peter M. Frinchaboy, J. G. Galbraith-Frew, Rafael A. García, D.A. García-Hernández, Ana E. García Pérez, Patrick Gaulme, Jian Ge, R. Génova-Santos, A. Georgakakis, Luan Ghezzi, Bruce A. Gillespie, Léo Girardi, Daniel Goddard, Satya Gontcho A. Gontcho, Jonay I. González Hernández, Eva K. Grebel, Paul J. Green, Jan Niklas Grieb, Nolan Grieves, James E. Gunn, Hong Guo, Paul Harding, Sten Hasselquist, Suzanne L. Hawley, Michael Hayden, Fred R. Hearty,

Saskia Hekker, Shirley Ho, David W. Hogg, Kelly Holley-Bockelmann, Jon A. Holtzman, Klaus Honscheid, Daniel Huber, Joseph Huehnerhoff, Inese I. Ivans, Linhua Jiang, Jennifer A. Johnson, Karen Kinemuchi, David Kirkby, Francisco Kitaura, Mark A. Klaene, Gillian R. Knapp, Jean-Paul Kneib, Xavier P. Koenig, Charles R. Lam, Ting-Wen Lan, Dustin Lang, Pierre Laurent, Jean-Marc Le Goff, Alexie Leauthaud, Khee-Gan Lee, Young Sun Lee, Timothy C. Licquia, Jian Liu, Daniel C. Long, Martín López-Corredoira, Diego Lorenzo-Oliveira, Sara Lucatello, Britt Lundgren, Robert H. Lupton, III Mack, Claude E., Suvrath Mahadevan, Marcio A. G. Maia, Steven R. Majewski, Elena Malanushenko, Viktor Malanushenko, A. Manchado, Marc Manera, Qingqing Mao, Claudia Maraston, Robert C. Marchwinski, Daniel Margala, Sarah L. Martell, Marie Martig, Karen L. Masters, Savita Mathur, Cameron K. McBride, Peregrine M. McGehee, Ian D. McGreer, Richard G. McMahon, Brice Ménard, Marie-Luise Menzel, Andrea Merloni, Szabolcs Mészáros, Adam A. Miller, Jordi Miralda- Escudé, Hironao Miyatake, Antonio D. Montero-Dorta, Surhud More, Eric Morganson, Xan Morice- Atkinson, Heather L. Morrison, Benôit Mosser, Demitri Muna, Adam D. Myers, Kirpal Nandra, Jeffrey A. Newman, Mark Neyrinck, Duy Cuong Nguyen, Robert C. Nichol, David L. Nidever, Pasquier Noterdaeme, Sebastián E. Nuza, Julia E. O'Connell, Robert W. O'Connell, Ross O'Connell, Ricardo L. C. Ogando, Matthew D. Olmstead, Audrey E. Oravetz, Daniel J. Oravetz, Keisuke Osumi, Russell Owen, Deborah L. Padgett, Nikhil Padmanabhan, Martin Paegert, Nathalie Palanque- Delabrouille, Kaike Pan, John K. Parejko, Isabelle Pâris, Changbom Park, Petchara Pattarakijwanich, M. Pellejero-Ibanez, Joshua Pepper, Will J. Percival, Ismael Pérez-Fournon, Ignasi rez-Ra'fols, Patrick Petitjean, Matthew M. Pieri, Marc H. Pinsonneault, Gustavo F. Porto de Mello, Francisco Prada, Abhishek Prakash, Adrian M. Price-Whelan, Pavlos Protopapas, M. Jordan Raddick, Mubdi Rahman, Beth A. Reid, James Rich, Hans-Walter Rix, Annie C. Robin, Constance M. Rockosi, Thaíse S. Rodrigues, Sergio Rodríguez-Torres, Natalie A. Roe, Ashley J. Ross, Nicholas P. Ross, Graziano Rossi, John J. Ruan, J.A. Rubiño-Martín, Eli S. Rykoff, Salvador Salazar-Albornoz, Mara Salvato, Lado Samushia, Ariel G. Sánchez, Basílio Santiago, Conor Sayres, Ricardo P. Schiavon, David J. Schlegel, Sarah J. Schmidt, Donald P. Schneider, Mathias Schultheis, Axel D. Schwope, C.G. Scóccola, Caroline Scott, Kris Sellgren,

Hee-Jong Seo, Aldo Serenelli, Neville Shane, Yue Shen, Matthew Shetrone, Yiping Shu, V. Silva Aguirre, Thirupathi Sivarani, M. F. Skrutskie, Anže Slosar, Verne V. Smith, Flávia Sobreira, Diogo Souto, Keivan G. Stassun, Matthias Steinmetz, Dennis Stello, Michael A. Strauss, Alina Streblyanska, Nao Suzuki, Molly E. C. Swanson, Jonathan C. Tan, Jamie Tayar, Ryan C. Terrien, Aniruddha R. Thakar, Daniel Thomas, Neil Thomas, Benjamin A. Thompson, Jeremy L. Tinker, Rita Tojeiro, Nicholas W. Troup, Mariana Vargas-Magaña, Jose A. Vazquez, Licia Verde, Matteo Viel, Nicole P. Vogt, David A. Wake, Ji Wang, Benjamin A. Weaver, David H. Weinberg, Benjamin J. Weiner, Martin White, John C. Wilson, John P. Wisniewski, W. M. Wood-Vasey, Christophe Ye'che, Donald G. York, Nadia L. Zakamska, O. Zamora, Gail Zasowski, Idit Zehavi, Gong-Bo Zhao, Zheng Zheng, Xu Zhou, Zhimin Zhou, Hu Zou, and Guangtun Zhu. The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III. *The Astrophysical Journal Supplement Series*, 219:12, July 2015. doi: 10.1088/0067-0049/219/1/12.

[2] S. Bardeau, G. Soucail, J.-P. Kneib, O. Czoske, H. Ebeling, P. Hudelot, I. Smail, and G. P. Smith. A CFH12k lensing survey of X-ray luminous galaxy clusters. II. Weak lensing analysis and global correlations. 470:449–466, August 2007. doi: 10.1051/0004-6361:20077443.

[3] E. Bertin and S. Arnouts. SExtractor: Software for source extraction. 117:393–404, June 1996.

[4] José Luis Nilo Castellón, M. Victoria Alonso, Diego García Lambas, Carlos Valotto, Ana Laura O'Mill, Héctor Cuevas, Eleazar R. Carrasco, Amelia Ramírez, José M. Astudillo, Felipe Ramos, Marcelo Jaque Arancibia, Natalie Ulloa, and Yasna Órdenes. Low X-Ray Luminosity Galaxy Clusters: Main Goals, Sample Selection, Photometric and Spectroscopic Observations. *AJ*, 151:151, June 2016. doi: 10.3847/0004-6256/151/6/151.

[5] K. W. Cavagnolo, M. Donahue, G. M. Voit, and M. Sun. Intracluster Medium Entropy Profiles for a Chandra Archival Sample of Galaxy Clusters. 182:12–32, May 2009. doi: 10.1088/0067-0049/182/1/12.

[6] H. Dahle, N. Kaiser, R. J. Irgens, P. B. Lilje, and S. J. Maddox. Weak Gravitational

Lensing by a Sample of X-Ray Luminous Clusters of Galaxies. I. The Data Set. 139: 313–368, April 2002. doi: 10.1086/338678.

[7] A. A. Dariush, S. Raychaudhury, T. J. Ponman, H. G. Khosroshahi, A. J. Benson, R. G. Bower, and F. Pearce. The mass assembly of galaxy groups and the evolution of the magnitude gap. 405:1873–1887, July 2010. doi: 10.1111/j.1365-2966.2010.16569.x.

[8] M. de los Rios, M. J. Domínguez R., D. Paz, and M. Merchán. The MeSsI (merging systems identification) algorithm and catalogue. *MNRAS*, 458:226–232, May 2016. doi: 10.1093/mnras/stw215.

[9] E. D'Onghia and G. Lake. Cold Dark Matter's Small-Scale Crisis Grows Up. 612: 628–632, September 2004. doi: 10.1086/422794.

[10] E. D'Onghia, J. Sommer-Larsen, A. D. Romeo, A. Burkert, K. Pedersen, L. Portinari, and J. Rasmussen. The Formation of Fossil Galaxy Groups in the Hierarchical Universe. 630:L109–L112, September 2005. doi: 10.1086/491651.

[11] A. Dressler and S. A. Shectman. Evidence for substructure in rich clusters of galaxies from radial-velocity measurements. *AJ*, 95:985–995, April 1988. doi: 10.1086/114694.

[12] P. Eigenthaler and W. W. Zeilinger. The properties of fossil groups of galaxies. *Astronomische Nachrichten*, 330:978, December 2009. doi: 10.1002/asna.200911274.

[13] G. P. Garmire, M. W. Bautz, P. G. Ford, J. A. Nousek, and G. R. Ricker, Jr. Advanced CCD imaging spectrometer (ACIS) instrument on the Chandra X-ray Observatory. In J. E. Truemper and H. D. Tananbaum, editors, *X-Ray and Gamma-Ray Telescopes and Instruments for Astronomy.*, volume 4851 of *Proceedings of SPIE*, pages 28–44, March 2003. doi: 10.1117/12.461599.

[14] Q. Guo, S. White, M. Boylan-Kolchin, G. De Lucia, G. Kauffmann, G. Lemson, C. Li, V. Springel, and S. Weinmann. From dwarf spheroidals to cD galaxies: simulating the galaxy population in a ΛCDM cosmology. 413:101–131, May 2011. doi: 10.1111/j.1365-2966.2010.18114.x.

[15] A. Hou, L. C. Parker, W. E. Harris, and D. J. Wilman. Statistical Tools for Classifying Galaxy Group Dynamics. *ApJ*, 702:1199–1210, September 2009. doi: 10.1088/0004-637X/702/2/1199.

[16] T. E. Jeltema, C. R. Canizares, M. W. Bautz, and D. A. Buote. The Evolution of Structure in X-Ray Clusters of Galaxies. 624:606–629, May 2005. doi: 10.1086/428940.

[17] L. R. Jones, T. J. Ponman, A. Horton, A. Babul, H. Ebeling, and D. J. Burke. The nature and space density of fossil groups of galaxies. 343:627–638, August 2003. doi: 10.1046/j.1365-8711.2003.06702.x.

[18] R. Kale, T. Venturi, S. Giacintucci, D. Dallacasa, R. Cassano, G. Brunetti, G. Macario, and R. Athreya. The Extended GMRT Radio Halo Survey. I. New upper limits on radio halos and mini-halos. 557:A99, September 2013. doi: 10.1051/0004-6361/201321515.

[19] S. Miyazaki, Y. Komiyama, M. Sekiguchi, S. Okamura, M. Doi, H. Furusawa, M. Hamabe, K. Imi, M. Kimura, F. Nakata, N. Okada, M. Ouchi, K. Shimasaku, M. Yagi, and N. Yasuda. Subaru Prime Focus Camera – Suprime-Cam. 54:833–853, December 2002. doi: 10.1093/pasj/54.6.833.

[20] A. G. Noble, T. M. A. Webb, A. Muzzin, G. Wilson, H. K. C. Yee, and R. F. J. van der Burg. A Kinematic Approach to Assessing Environmental Effects: Star-forming Galaxies in a z ˜ 0.9 SpARCS Cluster Using Spitzer 24 μm Observations. *ApJ*, 768:118, May 2013. doi: 10.1088/0004-637X/768/2/118.

[21] N. Okabe, M. Takada, K. Umetsu, T. Futamase, and G. P. Smith. LoCuSS: Subaru Weak Lensing Study of 30 Galaxy Clusters. 62:811–870, June 2010. doi: 10.1093/pasj/62.3.811.

[22] G. W. Pratt, E. Pointecouteau, M. Arnaud, and R. F. J. van der Burg. The hot gas content of fossil galaxy clusters. 590:L1, May 2016. doi: 10.1051/0004-6361/201628462.

[23] K. Rines, M. J. Geller, A. Diaferio, and M. J. Kurtz. Measuring the Ultimate Halo Mass of Galaxy Clusters: Redshifts and Mass Profiles from the Hectospec Cluster Survey (HeCS). 767:15, April 2013. doi: 10.1088/0004-637X/767/1/15.

[24] M. Schirmer. THELI: Convenient Reduction of Optical, Near-infrared, and Mid-infrared Imaging Data. *Astrophys. J. Suppl. Ser.*, 209:21, December 2013. doi: 10.1088/0067-0049/209/2/21.

[25] M. Schirmer, E. R. Carrasco, P. Pessev, V. Garrel, C. Winge, B. Neichel, and F. Vidal. Gemini Frontier Fields: Wide-field Adaptive Optics K_s-band Imaging of the Galaxy Clusters MACS J0416.1-2403 and Abell 2744. *Astrophys. J. Suppl. Ser.*, 217:33, April 2015. doi: 10.1088/0067-0049/217/2/33.

[26] Luca Scrucca, Michael Fop, Thomas Brendan Murphy, and Adrian E. Raftery. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal*, 8(1):205–233, 2016.

[27] G. P. Smith, J.-P. Kneib, I. Smail, P. Mazzotta, H. Ebeling, and O. Czoske. A Hubble Space Telescope lensing survey of X-ray luminous galaxy clusters - IV. Mass, structure and thermodynamics of cluster cores at z= 0.2. 359:417–446, May 2005. doi: 10.1111/j.1365-2966.2005.08911.x.

[28] D. N. Spergel, L. Verde, H. V. Peiris, E. Komatsu, M. R. Nolta, C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters. 148:175–194, September 2003. doi: 10.1086/377226.

[29] V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, J. A. Peacock, S. Cole, P. Thomas, H. Couchman, A. Evrard, J. Colberg, and F. Pearce. Simulations of the formation, evolution and clustering of galaxies and quasars. 435:629–636, June 2005. doi: 10.1038/nature03597.

[30] M. S. Yagi, Nao, H. Yamanoi, H. Furusawa, F. Nakata, and Y. Komiyama. Re-

Calibration of SDF/SXDS Photometric Catalogs of Suprime-Cam with SDSS Data Release 8. *Pub. Astron. Soc. Jap.*, 65:22, February 2013. doi: 10.1093/pasj/65.1.22.

[31] S. Zarattini, R. Barrena, M. Girardi, N. Castro-Rodriguez, W. Boschin, J. A. L. Aguerri, J. Méndez-Abreu, R. Sánchez-Janssen, C. Catalán-Torrecilla, E. M. Corsini, C. del Burgo, E. D'Onghia, N. Herrera-Ruiz, J. Iglesias-Páramo, E. Jimenez Bailon, M. Lozada Muoz, N. Napolitano, and J. M. Vilchez. Fossil group origins. IV. Characterization of the sample and observational properties of fossil systems. 565:A116, May 2014. doi: 10.1051/0004-6361/201323351.

[32] S. Zarattini, M. Girardi, J. A. L. Aguerri, W. Boschin, R. Barrena, C. del Burgo, N. Castro-Rodriguez, E. M. Corsini, E. D'Onghia, A. Kundert, J. Méndez-Abreu, and R. Sánchez-Janssen. Fossil group origins. VII. Galaxy substructures in fossil systems. 586:A63, February 2016. doi: 10.1051/0004-6361/201527175.

Chapter 7

Studies on the angular distribution of the cosmological parameters.

In this chapter we will present the results of the study of the cosmic microwave background through supervised and unsupervised machine learning techniques. These techniques allowed studying the CMB data in a new and independent way. These studies give place to the `CosmoML` software, which allows us to estimate the cosmological parameters using the temperature power spectrum of the CMB as feature. Although this algorithm does not improve the precision of the measurements compared with the traditional methods, it reduces significantly the computation time. Another advantage of this new approach is the ability to easily add new features from different cosmological observations (BAO, supernovas, Sunyaev-Zeldovich effect, etc.) and to analyse them in a homogeneous way.

7.1 On the temperature power spectrum of the cosmic microwave background.

In the chapter 1 we introduced the standard cosmological model, which describes the main characteristics of the universe assuming homogeneity and isotropy, and that the different fluids that composed the universe do not interact with each other. However, it is important to emphasise that these are approximations and that in order to describe some phenomena with greater precision, as for example the cosmic microwave background, it is necessary to

take into account the small inhomogeneities in the matter and energy distribution and the interactions between the different fluids. Taking into account the small inhomogeneities in the fluids distribution, these will produce inhomogeneities in the metric, which can be modelled by adding a perturbative term to the FRW metric (see eq. 1.2). If we work in the Newtonian gauge, also called conformal gauge, such perturbations are coded into 2 functions ψ and ϕ which depends on the spatial coordinates and on the time [9]:

$$ds^2 = a^2(t)\left[-(1+2\psi)dt^2 + (1-2\phi)dx^i dx^j\right] \tag{7.1}$$

On the other hand, in order to model the interactions between the different fluids, it is necessary to study the Boltzmann equations, which describe the temporal variation of the phase space distribution function of each fluid, taking into account the interactions between them:

$$\frac{df_s(x^i, P_i, t)}{dt} = \frac{\partial f_s}{\partial t} + \frac{dx^i}{dt}\frac{\partial f_s}{\partial x^i} + \frac{dP_i}{dt}\frac{\partial f_s}{\partial P_i} = C[f_s] \tag{7.2}$$

where $f_s(x^i, P_i, t)$ is the distribution function of a given fluid s which describes the distribution of that fluid in the phase space and $C[f_s]$ is a term that takes into account the interactions of that fluid with the other components present in the universe.

Following the definitions introduced in *Ma & Bertschinger 1995* [9], it is convenient to write the distribution function of a fluid as the distribution of the 0^{th} order function that will have the fluid in a homogeneous and isotropic universe without taking into account the interactions with other components, plus a perturbative term $\Psi(x^i, P_i, t)$:

$$f_s(x^i, P_j, t) = f_{s0}(P)[1 + \Psi(x^i, P_j, t)] \tag{7.3}$$

Taking into account this distribution function, we can write the energy-moment tensor as:

$$T_{\mu,\nu} = \int dP_1 dP_2 dP_3 (-g)^{-1/2} \frac{P_\mu P_\nu}{P^0} f(x^i, P_j, t) \tag{7.4}$$

Defining $q_i \equiv P_i/(1-\phi) = q n_i$, with $n^i n_i = \delta_{ij} n^i n^j = 1$ and taking into account the angular dependencies, this tensor can be rewritten in a simplified form as:

$$T_0^0 = -a^{-4}\int q^2 dq d\Omega \sqrt{q^2+m^2a^2} f_0(q)(1+\Psi) \tag{7.5}$$

$$T_i^0 = a^{-4}\int q^2 dq d\Omega q n_i f_0(q)\Psi \tag{7.6}$$

$$T_j^i = a^4\int q^2 dq d\Omega \frac{q^2 n_i n_j}{\sqrt{q^2+m^2a^2}} f_0(q)(1+\Psi) \tag{7.7}$$

Finally, we can relate the different components of the energy-momentum tensor with the physical properties of a fluid in a similar way to what is done in equation 1.3 for an homogeneous and isotropic universe

$$T_0^0 = -(\rho+\delta\rho) \tag{7.8}$$

$$T_i^0 = (\rho+P)v_i \tag{7.9}$$

$$T_j^i = (P+\delta P)\delta_j^i + \Sigma_j^i, \qquad \Sigma_i^i = 0 \tag{7.10}$$

Working in Fourier space we can define the variables θ and σ in the following way:

$$(\rho+P)\theta \equiv ik^j T_j^0 \tag{7.11}$$

$$(\rho+P)\sigma \equiv -(\hat{k}_i\hat{k}_j - \frac{1}{3}\delta_{ij})\Sigma_j^i \tag{7.12}$$

Solving the equations 7.2 for each fluid (dark matter, photons, neutrinos and baryons) taking into account the interactions between baryons and photons by Compton scattering, and assuming that dark matter does not present interactions with any particle of the standard model, the distributions of the different components result:

$$
\begin{aligned}
&Photons && \\
\dot{\delta_\gamma} &= -\frac{4}{3}\theta_\gamma + 4\dot{\phi} \\
\dot{\theta_\gamma} &= k^2(\frac{\delta_\gamma}{4} - \sigma_\gamma) + k^2\psi + \kappa^2(\theta_b - \theta_\gamma) \\
&Neutrinos && \\
\dot{\delta_\nu} &= -\frac{4}{3}\theta_\nu + 4\dot{\phi} \\
\dot{\theta_\nu} &= k^2(\frac{\delta_\nu}{4} - \sigma_\nu) + k^2\psi \\
&Dark\ Matter && \\
\dot{\delta}_{DM} &= -3\dot{\phi} - \theta_{DM} \\
\dot{\theta}_{DM} &= -\frac{\dot{a}}{a}\theta_{DM} + k^2\psi \\
&Baryonic\ Matter && \\
\dot{\delta_b} &= -3\dot{\phi} - \theta_b \\
\dot{\theta_b} &= -\frac{\dot{a}}{a}\theta_b + k^2\psi + k^2c_s^2\delta_b - R\dot{\kappa}(\theta_b - \theta_\gamma)
\end{aligned}
\tag{7.13}
$$

It is interesting to note the similarity between the equations of baryons and dark matter. This is because both components are nonrelativistic and present a state equation $w = 0$. However, we can see that in the equation for baryons there is an extra term ($k^2c_s^2\delta_b - R\dot{\kappa}(\theta_b - \theta_\gamma)$) that comes from the interaction between baryons and photons by Compton scattering.

Similarly, the equations for photons are very similar to the equations for neutrinos, because both components are relativistic and have a state equation $w = 1/3$. Again the only difference lies in the extra term in the case of the photons ($k^2\psi + \kappa^2(\theta_b - \theta_\gamma)$) that comes from their interaction with the baryons.

Finally, given a set of cosmological parameters and solving the set of equations 7.13, we can obtain the perturbations to the photons distribution, which will result in small perturbations in the temperature Δ_T of these photons. Finally, it these perturbations will translate to anisotropies in temperature and polarisation that we will observe in the cosmic microwave background. Taking into account the spherical topology of the problem, we can describe the temperature anisotropies by means of spherical harmonics:

$$\Delta_T(\vec{x}) = \sum_{l=1}^{\infty} \sum_{m=-l}^{l} a_{lm}(\vec{x}) Y_{lm} \tag{7.14}$$

Taking into account the statistical nature of this analysis, we can not exactly predict the temperature that will have a photon coming from a given direction. However, we can predict the power spectrum C_l of that distribution, which is defined by:

$$< a_{lm} a^*_{l'm'} >= \delta_{ll'} \delta_{mm'} C_l \tag{7.15}$$

Where $<>$ refers to an angular average. Finally this power spectra (C_l) will be the fundamental observable with which we will work.

7.2 Analysis of the main components of the temperature power spectra.

As it was introduced in chapter 2, to work with machine learning methods it is necessary to have a data set that contains all the relevant information for our problem. For this work we estimate the power spectra for 5000 random cosmological models varying the cosmological parameters $\Omega_m h^2$, $\Omega_b h^2$, Ω_k, H_0, n, A_s and τ within an interval around the parameters estimated by Planck [11]. In addition we will add as an independent parameter $10^9 A_s e^{-2\tau}$.

Given a set of cosmological parameters, we can solve the set of equations 7.13 and obtain the temperature power spectrum that would be observed in a universe with those parameters (see equation 7.15). There are several programs that perform these calculations in a numerical way. In this chapter we will use the software `CAMB (Code for Anisotropies in the Microwave Background)` [1] to calculate the power spectra corresponding to different cosmological models.

In figure 7.1 you can see examples of some spectra calculated for different cosmologies.

In table 7.1 we present the intervals used in each parameter to build the training set.

It is important to note that these spectra are theoretical, that is, they come from the

[1] https://camb.info/

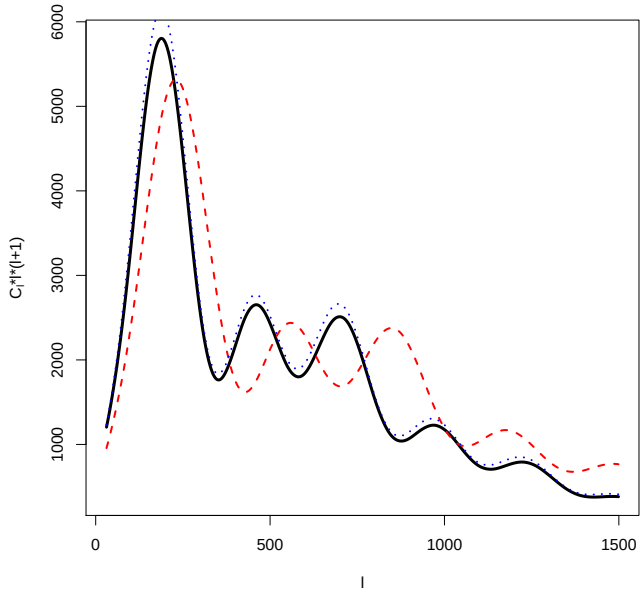

Figure 7.1: Examples of power spectra computed with `CAMB` for different sets of cosmological parameters.

numerical resolution of the Boltzmann-Einstein (Equations 7.13). However, the spectrum estimated from the measurement of the temperature anisotropies on the sky suffers from various observational problems such as, the contamination from secondary anisotropies, the resolution of the telescope, etc. that introduce errors in the measurement. To have a set of simulated spectra similar to those observed, we construct, for each one of the numerically calculated spectra, a simulated temperature map with the same resolution as the Planck maps using the `sphtfunc.synfast` function of the `healpy` software [2] [4]. Then, using the `anafast` function from the same software, we calculate the spectrum corresponding to the temperature map. In this way we simulate the errors introduced by the resolution of the instrument. Finally, to minimize the errors introduced by the resolution of the instrument is common to work with a binned spectrum. In our case we will use 80 bins, because if we use a smaller number of bins, relevant information is lost, whereas if we increase too much the number of bins we begin to model the errors produced by the resolution of the instrument. In figure 7.2 it can be seen an example of a power spectrum

[2] http://healpix.sourceforge.net

Parameter	Minimum	Maximum	Planck
$\Omega_m h^2$	0.1131	0.1263	0.1197
$\Omega_b h^2$	0.02131	0.02269	0.022
Ω_k	-0.1	0.1	0
H_0	62.31	72.31	67.31
n	0.9469	0.9841	0.9655
A_s	$1.988 * 10^{-9}$	$2.408 * 10^{-9}$	$2.198 * 10^{-9}$
τ	0.021	0.1349	0.078

Table 7.1: Studied intervals for each cosmological parameter.

measured in the whole sky and its corresponding binned spectra.

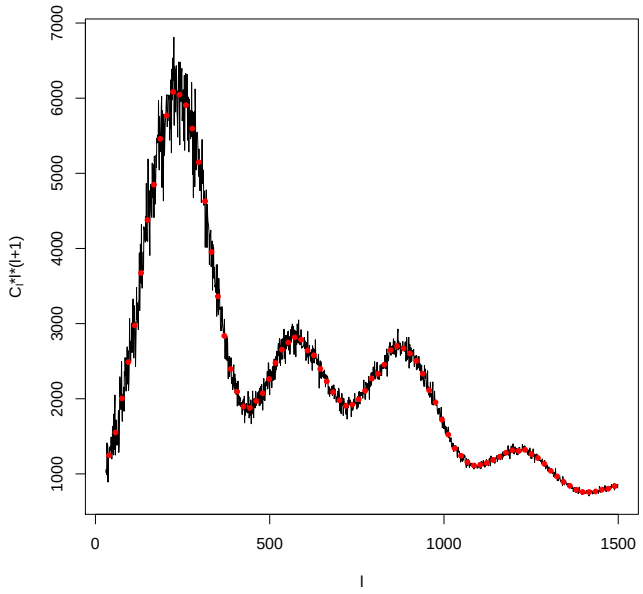

Figure 7.2: Example of a full-sky spectrum (black line) superimposed to the estimated bins (red dots).

It is worth mentioning that the Planck maps that we will use in section 7.4.3 were previously reduced in order to substract the contamination of secondary anisotropies, so it is not necessary to introduce such errors in our simulated maps.

In figure 7.3 it can be seen the final sample of simulated spectra for the set of cosmologies previously described.

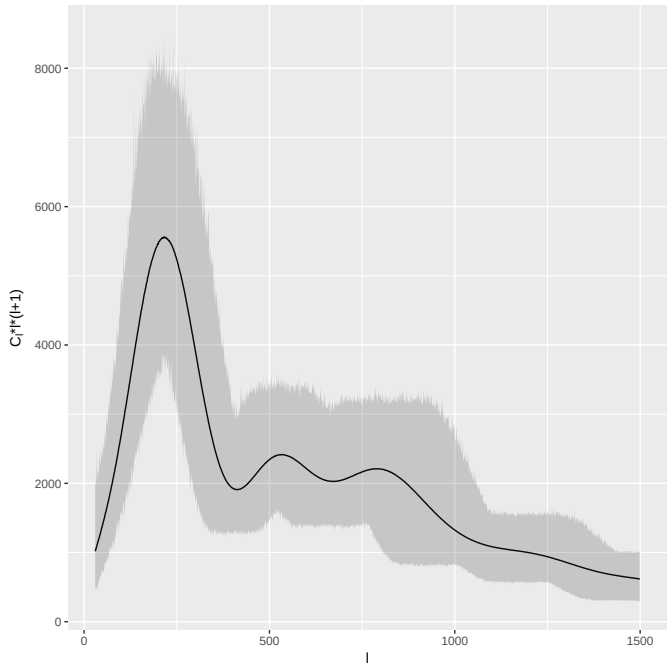

Figure 7.3: Power spectrum of the different cosmological models that will be used in the analysis with machine learning methods.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
% Variance	.641	.239	.0894	.016	.0026	.0019	.000702	.00068	.00064	.00063
Accumulative Variance	.641	.88	.9694	.9854	.988	.9899	.9906	.9913	.99194	.99257

Table 7.2: Relative importance of the principal components of the spectra.

As a first step to study a data set, it is common to performed a principal components analysis (see section 2.2.1), which gives us information between the correlations of the different features (in this case the bins of the power spectra) and can help us to reduce the dimensionality of the problem.

When doing this analysis to the set of power spectra, we find that only the first 8 components are important, representing more than the 99% of the variance of the data. In other words, with only the values of the first 8 components any spectrum (within the studied ranges) can be reconstructed with less than 1% of error. In table 7.2 you can see the relative importance of the first 10 principal components.

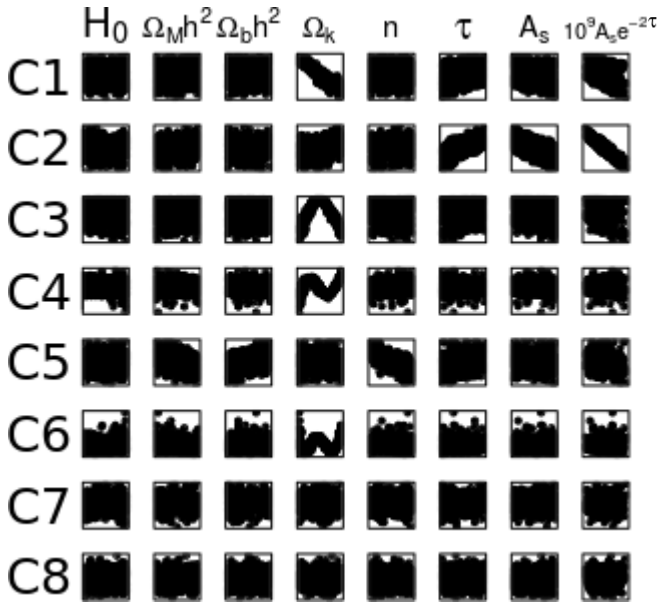

Figure 7.4: Correlation between the 8 first principal components and the cosmological parameters.

7.2.1 Unsupervised learning techniques applied to the principal components.

Having found the principal components that describe our data set, it is natural to analyse the correlations between these components and the cosmological parameters.

The figure 7.4 shows, schematically, the relationships between the first 8 components and the cosmological parameters. As you can see, there are parameters that have a strong correlation with some of the components (see for example the correlation between the first component $C1$ and Ω_k), while there are other cosmological parameters (H_0) that seem to have no relationship with these components. This type of analysis shows us quickly that these components can be use as features to predict some of the cosmological parameters.

Continuing with the analysis of the spectra by unsupervised learning techniques, it is natural to ask whether the spectra corresponding to different cosmological models are grouped in different areas in the feature space. In order to do this we perform a mixture of Gaussians (see section 2.2.1), with the package `Mclust` [14] of `R`, in the feature space without using information about the cosmological parameters. In figure 7.5 we show the cosmological models in the $\Omega_m h^2 - \Omega_k$ space, distinguishing by colour the groups found

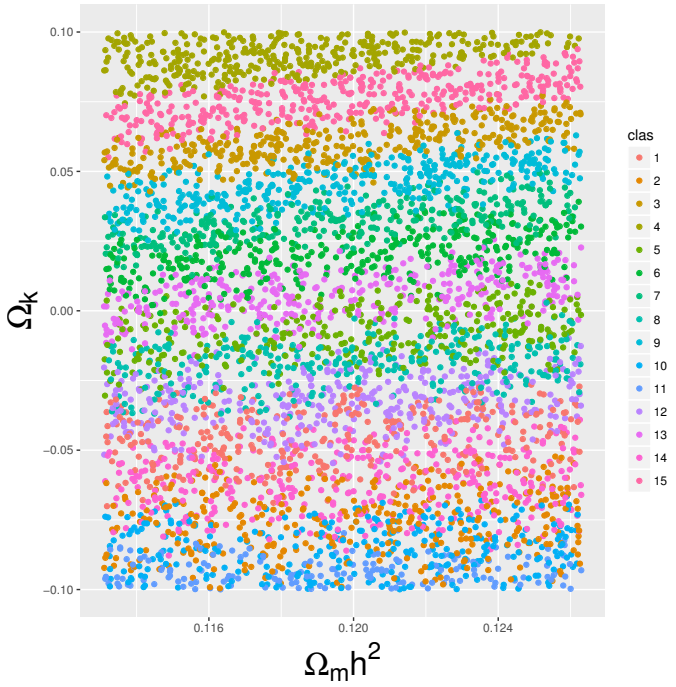

Figure 7.5: Cosmological models belonging to the same groups found by `mclust` in the $\Omega_m h^2$-Ω_k space.

by the Gaussian mixture. It is important to note that this algorithm only studies the principal components calculated from the spectra without information from the cosmological parameters that gave rise to those spectra. You can see that cosmologies that share the same Ω_k are grouped. This can be understood since there is a strong correlation between this parameter and the principal components, especially with the first one $C1$. As will be seen later, these correlations imply that the curvature is the cosmological parameter predicted with the best performance using machine learning algorithms.

7.3 *CosmoML*: Estimation of the cosmological parameters through supervised machine learning methods.

As seen in the previous section, there are different correlations between the principal components and the cosmological parameters. This means that by studying such relationships it is possible to predict the cosmological parameters using the principal components of a

given spectrum.

The traditional way of estimating the cosmological parameters using the CMB spectrum is through an extensive exploration in the cosmological parameter space using Monte Carlo Markov chains. This method has the advantage that it does not only estimate the parameters that best fit a given spectrum, but also give us information about the covariances of these parameters around the best fit values. However, the main disadvantage of these algorithms is that they consume a lot of time, especially when they want to explore multidimensional spaces such as those of cosmologic interest in which is common to work with more than 6 variables.

Another disadvantage of traditional methods is the difficulty of combining data from different cosmological tests (BAO, CMB, supernovas, etc.) since these techniques uses the Bayes theorem to estimate the probability of a given observation to come from a given set of cosmological parameters. To estimate this probability, it is necessary to know the combined *likelihood* of the different data sets, what is very difficult to obtain analytically and usually requires of big simulations for it estimation.

With this in mind, we decided to perform a supervised machine learning algorithm that estimates the cosmological parameters using as features the principal components of the CMB power spectrum. These techniques have the ability to estimate the parameters that best fit a given spectrum in a computationally more efficient way than the traditional methods, but suffers the disadvantage of not being able to provide information on the covariances around the best fit values. On the other hand, taking into account the way in which machine learning methods work, it is very easy to add features that contain information from different cosmological probes.

In section 7.3.1 we will analyse a 'toy model' to understand the improvement on the estimation of the cosmological parameters when we add features that contain information about the Hubble diagram of supernovas.

As we saw in chapter 2 the starting point of any supervised machine learning method is a data set in which we know both the predictor variables or features (in this case the principal components of each spectrum) and the variables that we want to predict (in this case the cosmological parameters). Using the data set previously built through numerical simulations (see section 7.2) we studied 3 machine learning techniques belonging to the `caret` [8] library of `R`: *Random Forest*, *Support Vector Machine* and *K-Nearest Neigh-*

bour. It is worth mentioning that because these algorithms are designed to predict a single variable using n predictor variables, it was necessary to train a different model for each cosmological parameter.

In order to avoid overfitting we randomly divide the data set into 2 independent sets, a training set and a test set, and then we train the machine learning algorithms using only the training set. Once the algorithms are trained, we can estimate the cosmological parameters corresponding to the test set, and then compare them with the real parameters.

In figures 7.6 and 7.7 you can see the comparison between the predicted cosmological parameters and the real ones for the 3 machine learning methods previously described.

To quantify the error made in the predictions it is convenient to define the following statistics:

$$\chi^2 = \sum_i^n \frac{(\Omega_{i,real} - \Omega_{i,predicho})^2}{n-1} \tag{7.16}$$

where $\Omega_{i,real}$ and $\Omega_{i,predicted}$ are the real and predicted values of each parameter for each cosmological model of the test set.

As can be seen, all the algorithms predict with good precision all the cosmological parameters with the exception of the Hubble constant. In principle this could be due to the fact that we are not analysing the appropriate features, i.e. that the variables that we are using to predict do not have enough information to predict H_0. To verify this, we made the same analysis but adding new features, such as: the position (i.e. the multipole) of the maximums and minimums of each spectrum, the value (i.e. $C_l l(l+1)$) of the maximums and minimums, the ratio between the different maximums, the ratio between the different minimums and the ratio between maximums and minimums. The figures 7.8 and 7.9 show the comparison between the cosmological parameters predicted and the real ones for the testing set with the 3 supervised machine learning methods previously described using the extended set of features. As it can be seen, even in this new feature space the algorithms have a bad estimation of H_0. Although this does not rule out that there could be another feature space not studied by us in which one can get a good estimate of H_0, this analysis indicates that to estimate H_0 it is necessary to continue with the traditional methods or to add information from other cosmic tests.

The results of the χ statistics for the different algorithms are summarised in table 7.3. It can be seen that the algorithm that has the best performance is the *Support Vector Machine*.

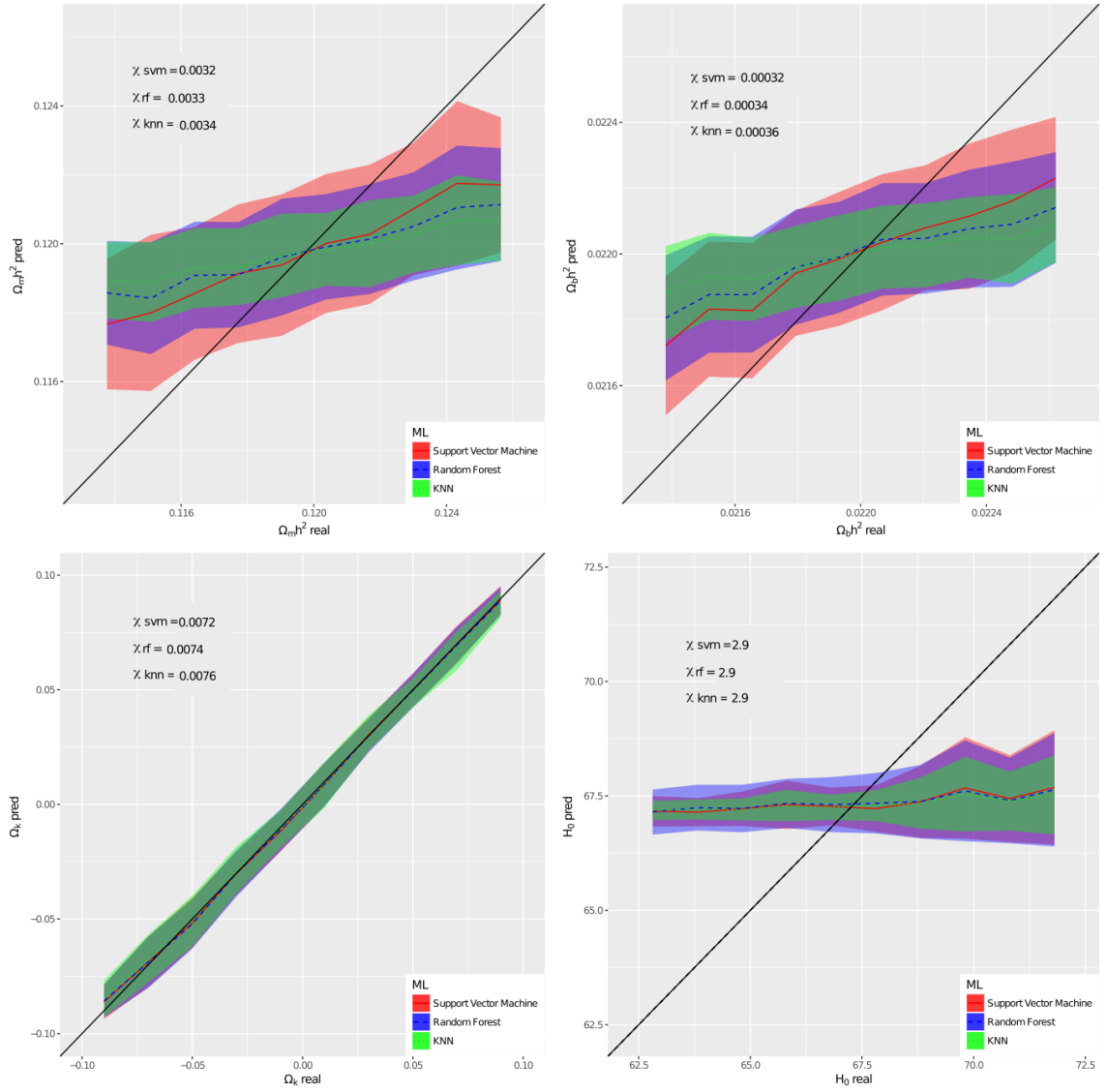

Figure 7.6: Comparison between cosmological parameters predicted with supervised learning methods and real ones for all cosmological models of the test set.

Finally, after having trained the different models, we can carry out an study on the relative importance of the features used to make the predictions. In figures 7.10 and 7.11 you can see the relative importance of the different features when estimating each of the cosmological parameters using the *Support Vector Machine* algorithm previously trained.

As can be seen, the estimation of $\Omega_m h^2$, $\Omega_b h^2$ and n depends mainly on the 5^{th} principal component, what matches the correlations plotted in figure 7.4. Note that although

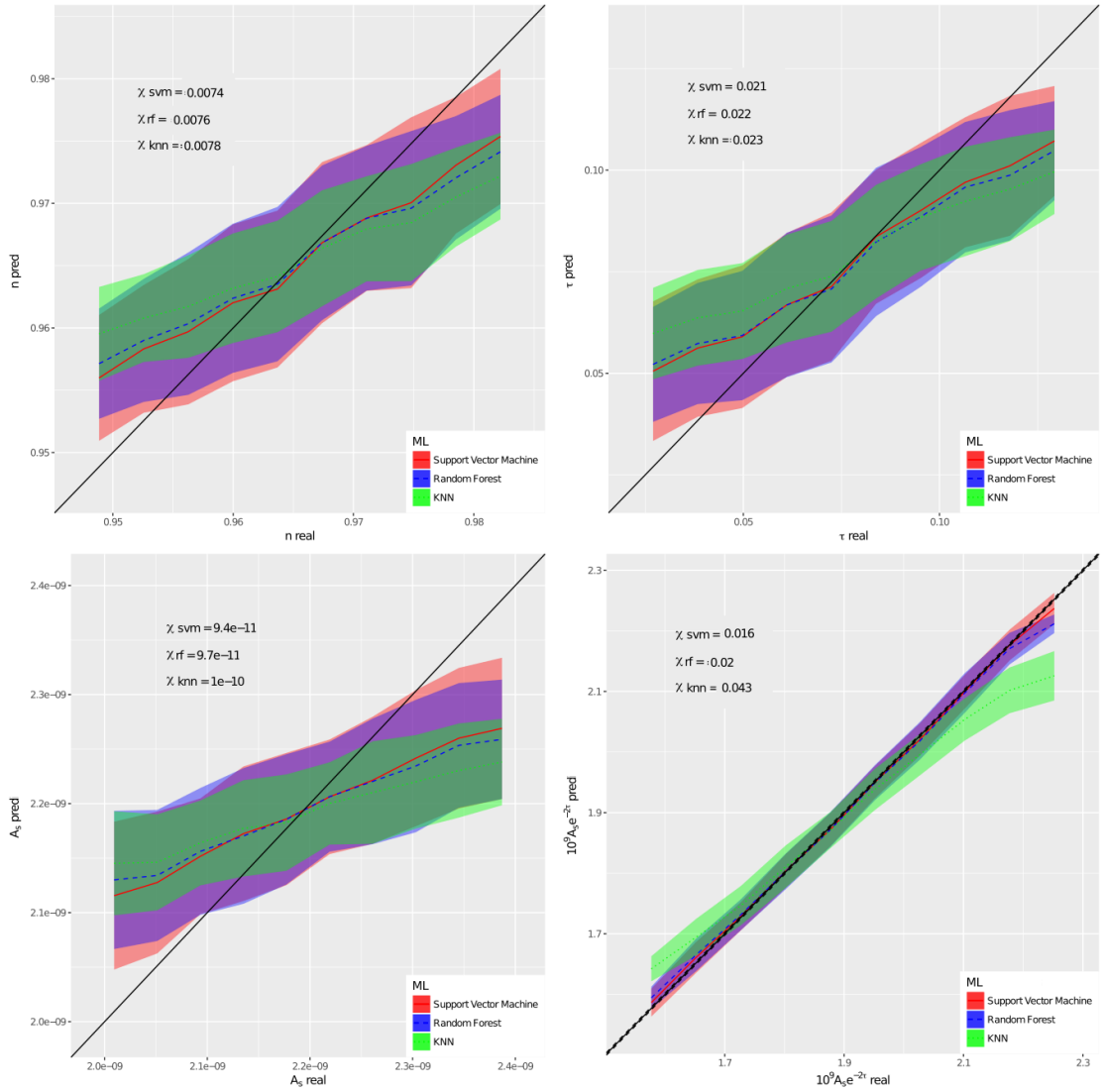

Figure 7.7: Comparison between cosmological parameters predicted with supervised learning methods and real ones for all cosmological models of the test set.

Model	$\Omega_m h^2$	Ω_k	$\Omega_b h^2$	H_0	τ	n	A_s $[10^{-11}]$	$10^9 A_s e^{-2\tau}$
RF	0.0032	0.0073	0.00030	2.9	0.021	0.007	9.1	0.016
KNN	0.0033	0.0083	0.00032	2.8	0.021	0.0073	9.4	0.031
SVM	0.0031	0.0071	0.00029	2.9	0.020	0.007	8.8	0.016

Table 7.3: Results of the performance (χ) of the different machine learning methods when estimating the cosmological parameters.

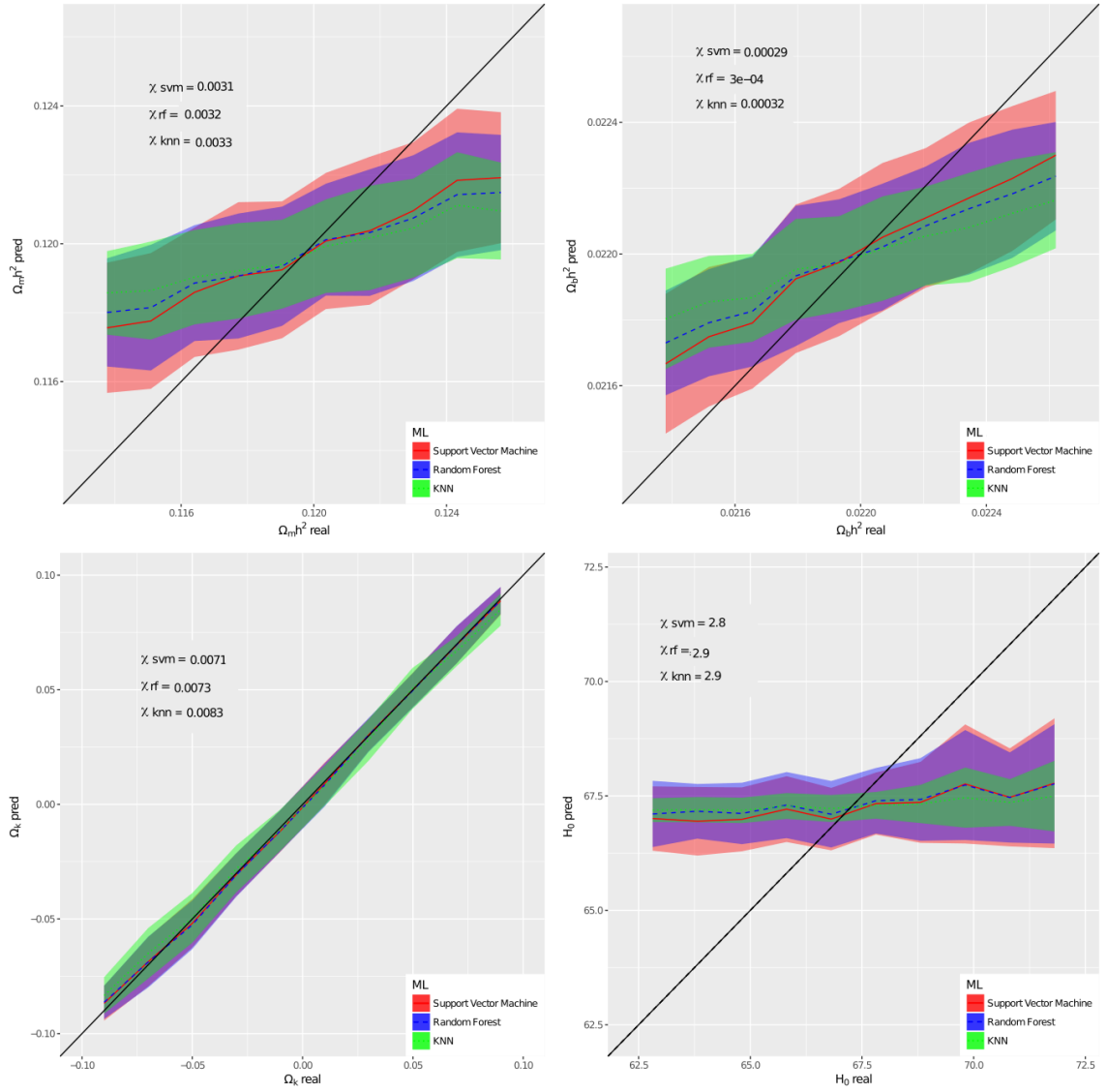

Figure 7.8: Comparison between the cosmological parameters predicted with the extended feature space and the real ones for all the cosmological models of the test set.

there is a correlation between these parameters and the 5^{th} component, this correlation has a certain scatter, which translates into uncertainties when estimating these cosmological parameters through supervised machine learning methods.

On the other hand, the estimation of τ, A_s and $10^9 A_s e^{-2\tau}$ depends mainly on the 2^{nd} component. Again this agrees with what is plotted in figure 7.4.

Finally, the estimation in Ω_k depends on the first component, with which it has a strong

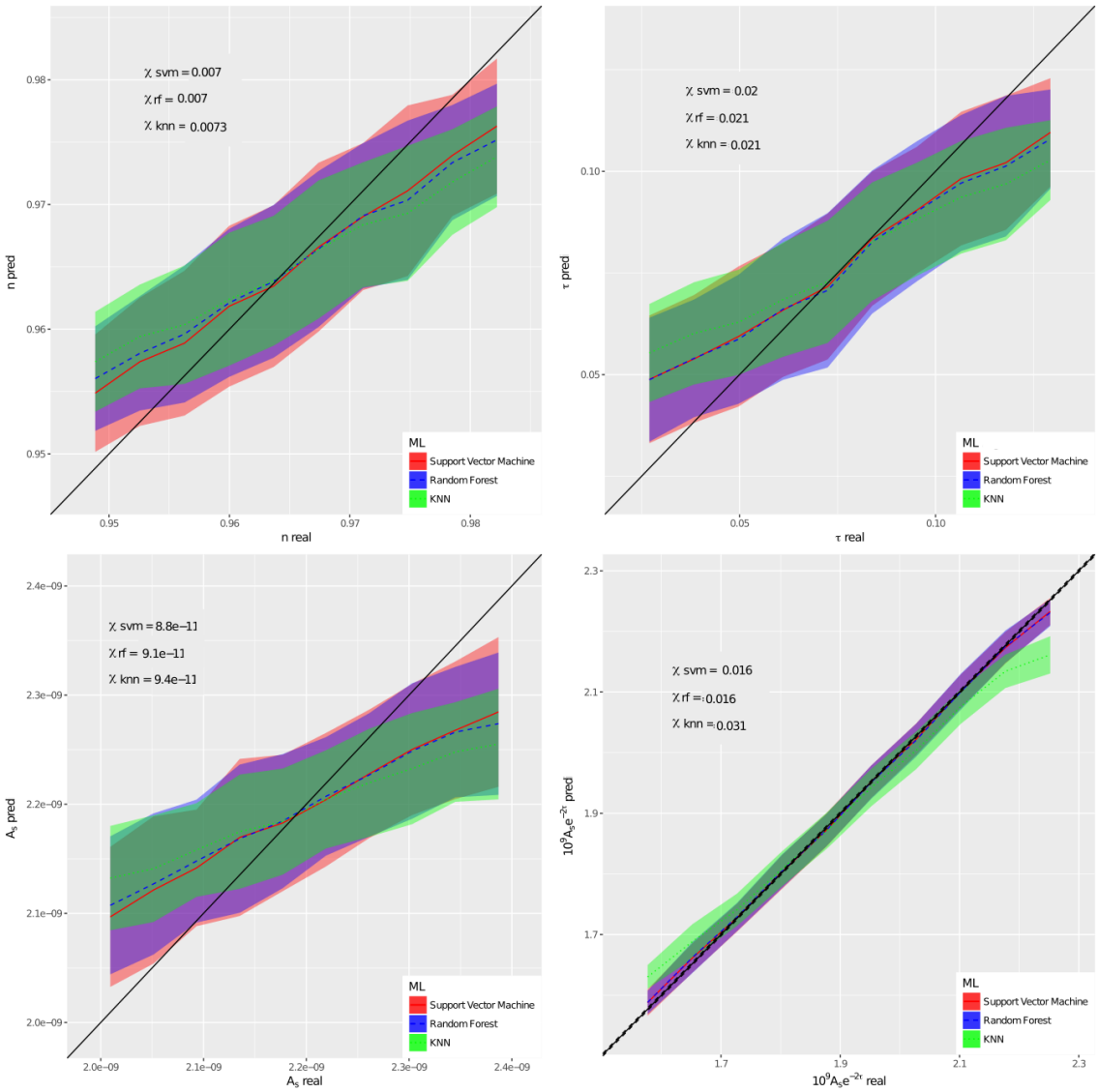

Figure 7.9: Comparison between the cosmological parameters predicted with the extended feature space and the real ones for all the cosmological models of the test set.

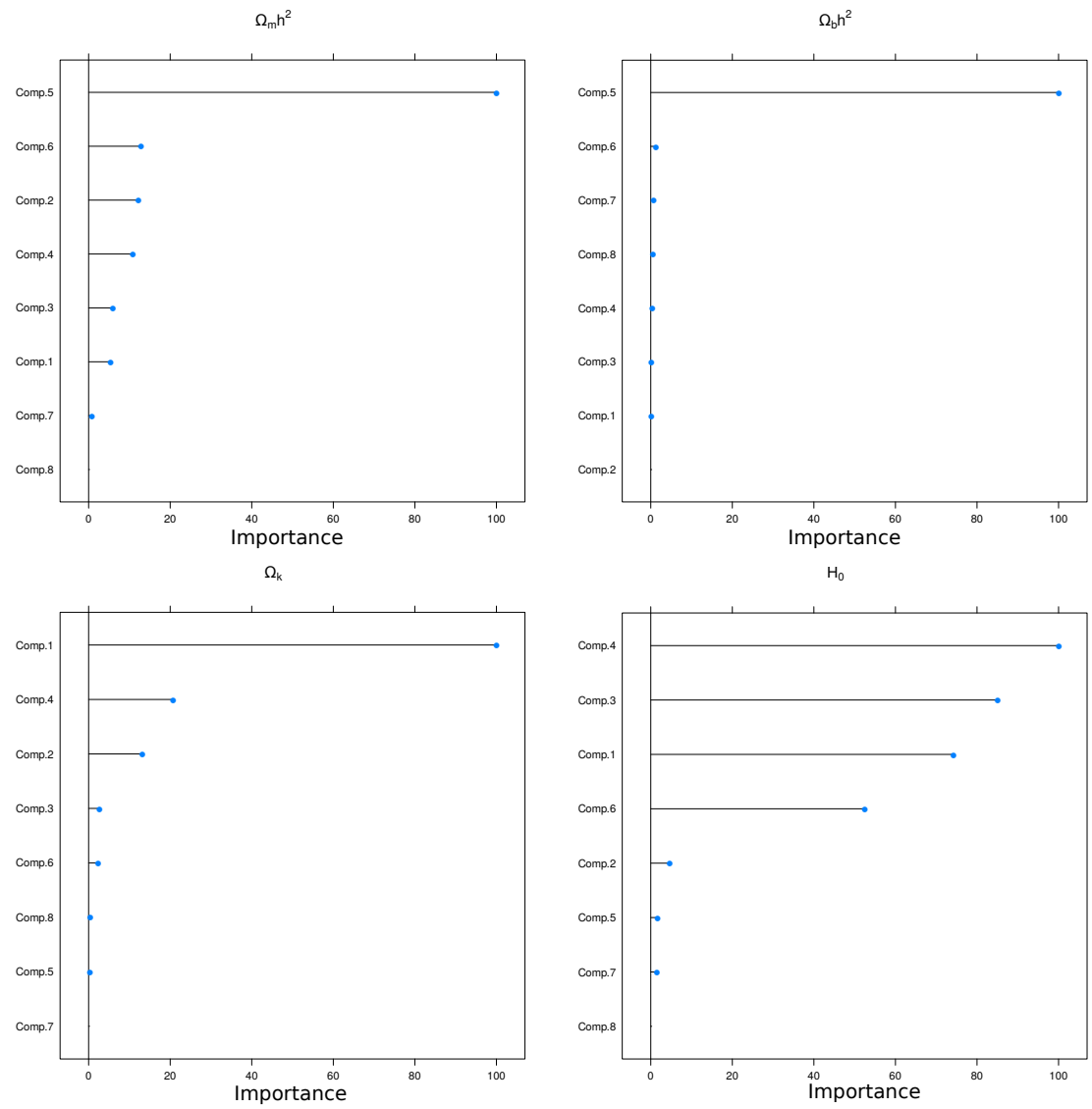

Figure 7.10: Relative importance of the different features when predicting the cosmological parameters.

correlation (see figure 7.4) with a low scatter. In addition, this parameter also has a great correlation with other components, which leads to the fact that its estimation is the one with the least uncertainty.

As can be seen in figure 7.6 the estimation of the Hubble constant H_0 through machine learning techniques has a great uncertainty. This can be understood due to the lack of correlation (see figure 7.4) between this parameter and the principal components (see figure

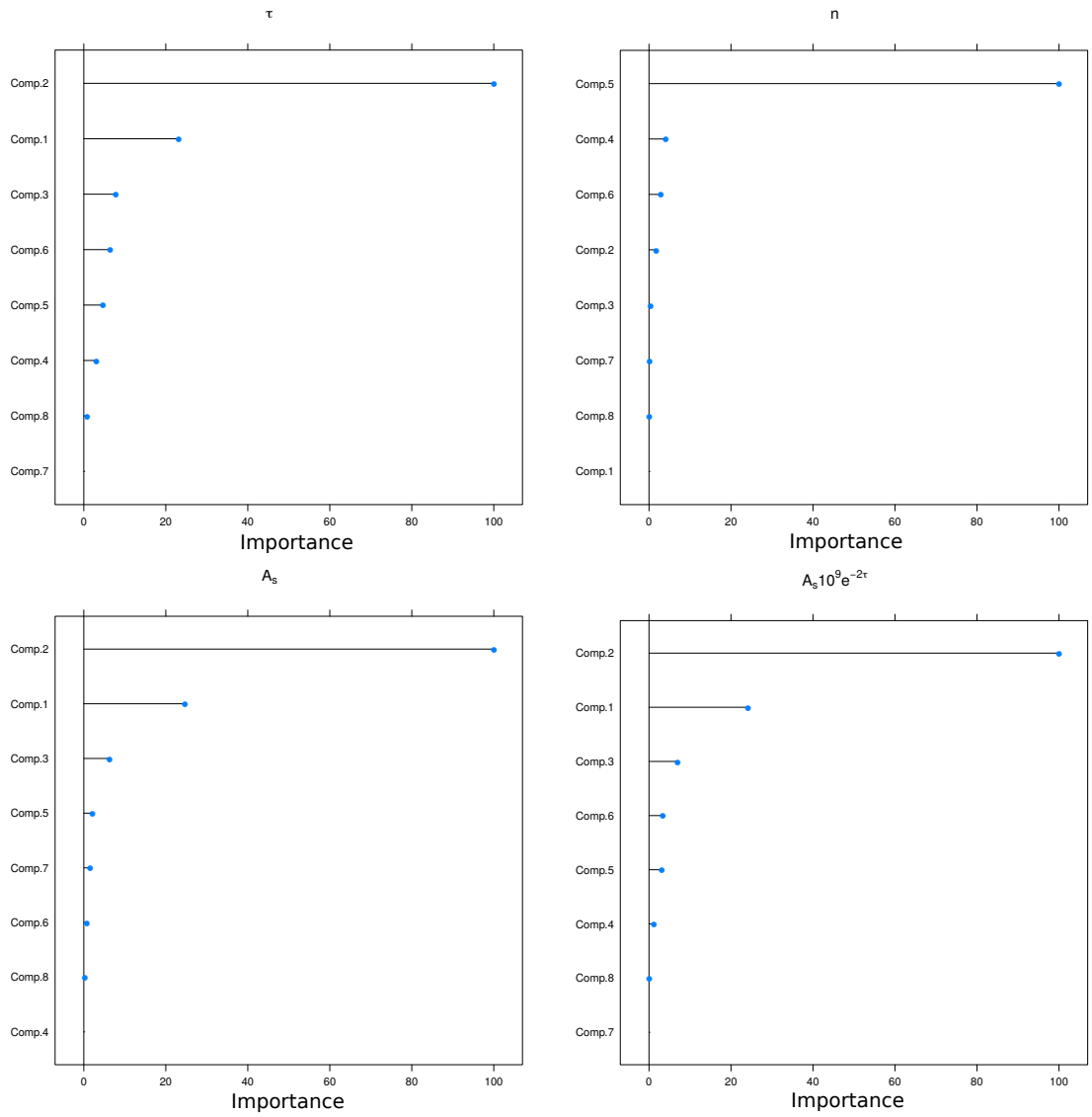

Figure 7.11: Relative importance of the different features when predicting the cosmological parameters..

7.10).

7.3.1 Cosmological parameters estimation using CMB + Supernovas.

As is known, there are different degenerancies between the cosmological parameters that produce that different combinations of parameters give rise to very similar observables.

This is why it is essential to combine different cosmological tests that break these degenerations and so, get more precise constraints in the different parameters. With this in mind, we decided to make a 'toy model' to study how our predictions would be affected if we add features that contain information about the Hubble diagram of supernovas. The Hubble diagram is a relation between the distance module (DM), estimated from the comparison between the absolute and the apparent magnitude of an object, and the redshift. This relationship is given by the following equation [5]:

$$DM = 5log\left(\frac{(1+z)D_M}{10Pc}\right) \tag{7.17}$$

where D_M is the comoving distance that depends on the cosmological parameters.

For this study we generated 5000 simulated Hubble diagrams using the same cosmological models used previous to generate the CMB power spectra. With this goal in mind we construct, for each set of cosmological parameters, a set of simulated galaxies that have the same distribution of redshift as the JLA supernova sample [1]. Then, using the equation 7.17, we estimate the DM corresponding to each galaxy according to its redshift. Finally we added a Gaussian error with $\sigma = 0.05 + 0.004 * DM$ (using the function `generate_mu_z` of the software `astroml`) in the distance module to simulate the real uncertainties produced by observational effects. As features we used the average of the distance module of the galaxies in 10 bins of redshift.

In figure 7.12 you can see the binned Hubble diagram of the supernovae of the JLA sample, superimposed on the simulated Hubble diagrams.

Finally our new training set consists of 5000 cosmological models with their respective CMB power spectrum and a simulated Hubble diagram. Using as features the 8 principal components of the power spectra, plus the 10 bins from the Hubble diagram, we train 3 supervised machine learning algorithms (*Random Forest*, *Support Vector Machine* and *K-Nearest Neighbour*) using the training set.

In figures 7.13 and 7.14 we show the comparisons between the predicted cosmological parameters and the real ones for the cosmological models of the test set.

You can see that all the algorithms have a good performance when estimating all the cosmological parameters. It is interesting to note the improvement in the prediction of the Hubble constant H_0 if we compare with the predictions made without information from the Hubble diagram.

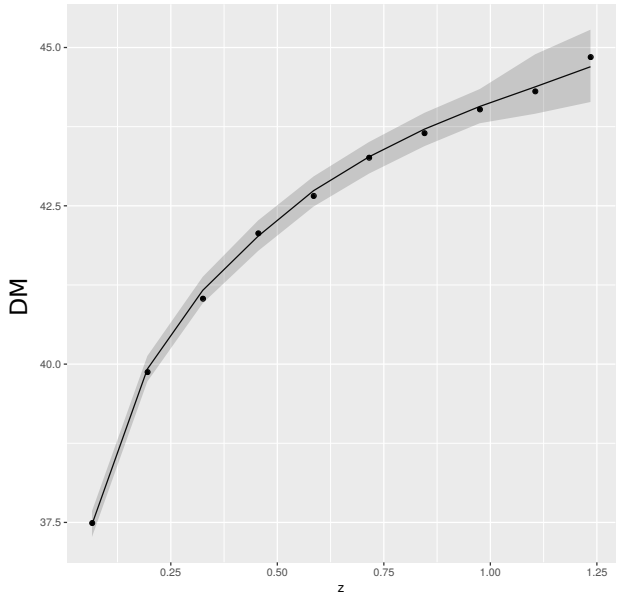

Figure 7.12: Binned Hubble diagram (Modulus distance in function of the redshift). In gray lines are shown the supernovas of the simulated samples and in the black line is highlighted the mean of DM for each bin of z. While in black points are plotted the average of the DM in bins of z for the JLA supernova sample.

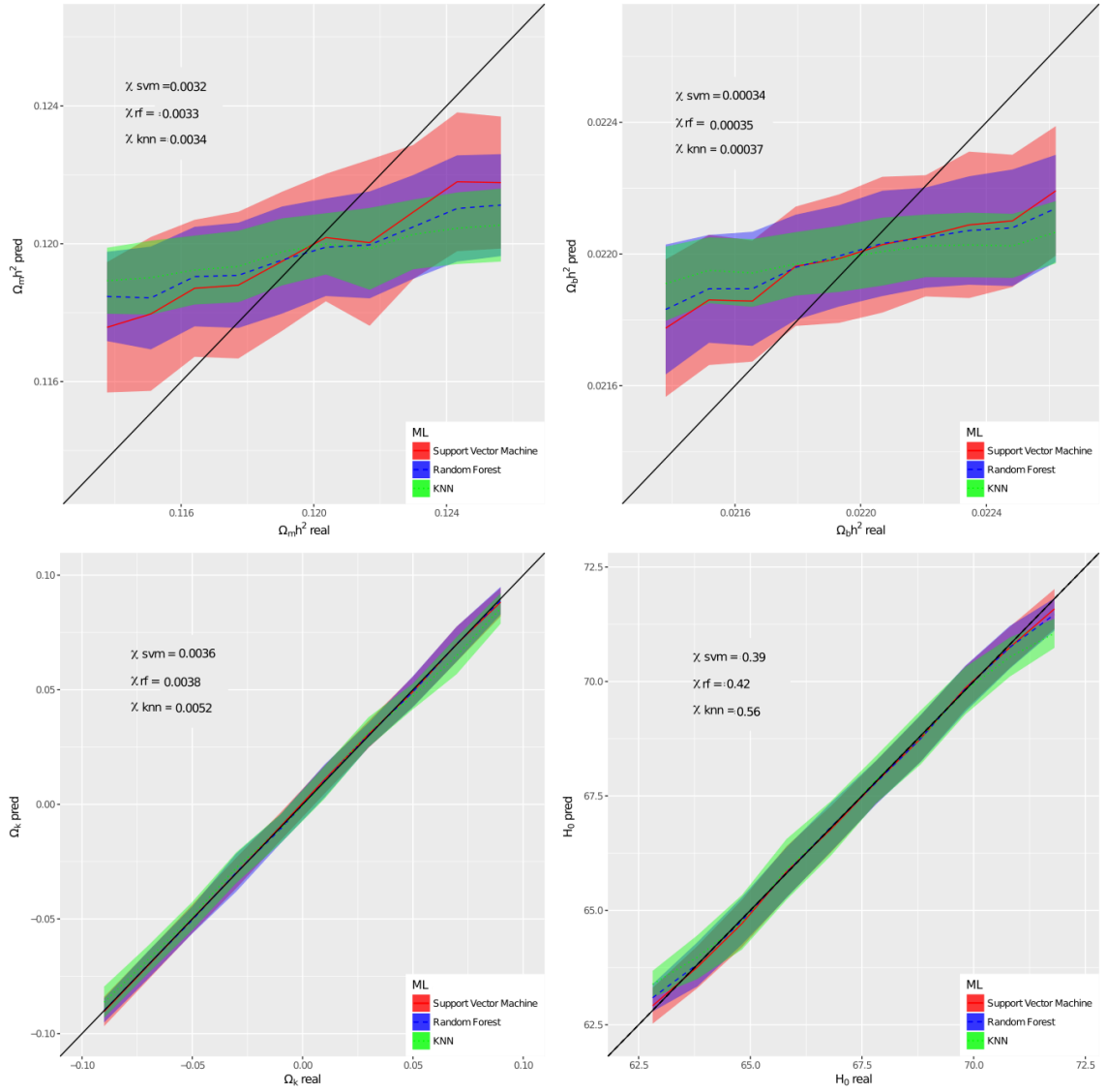

Figure 7.13: Comparison between the cosmological parameters predicted with CMB+Hubble-diagram features and the real ones for all cosmological models of the test set.

This is due to the fact that when using 2 different cosmological probes (CMB + Hubble Diagram), the degenerancies present in the estimation of the parameters can be broken and thus, achieve a better prediction.

The figure 7.15 shows the relative importance of the features when predicting the Hubble constant. It is interesting to note that the more relevant features are those that have in-

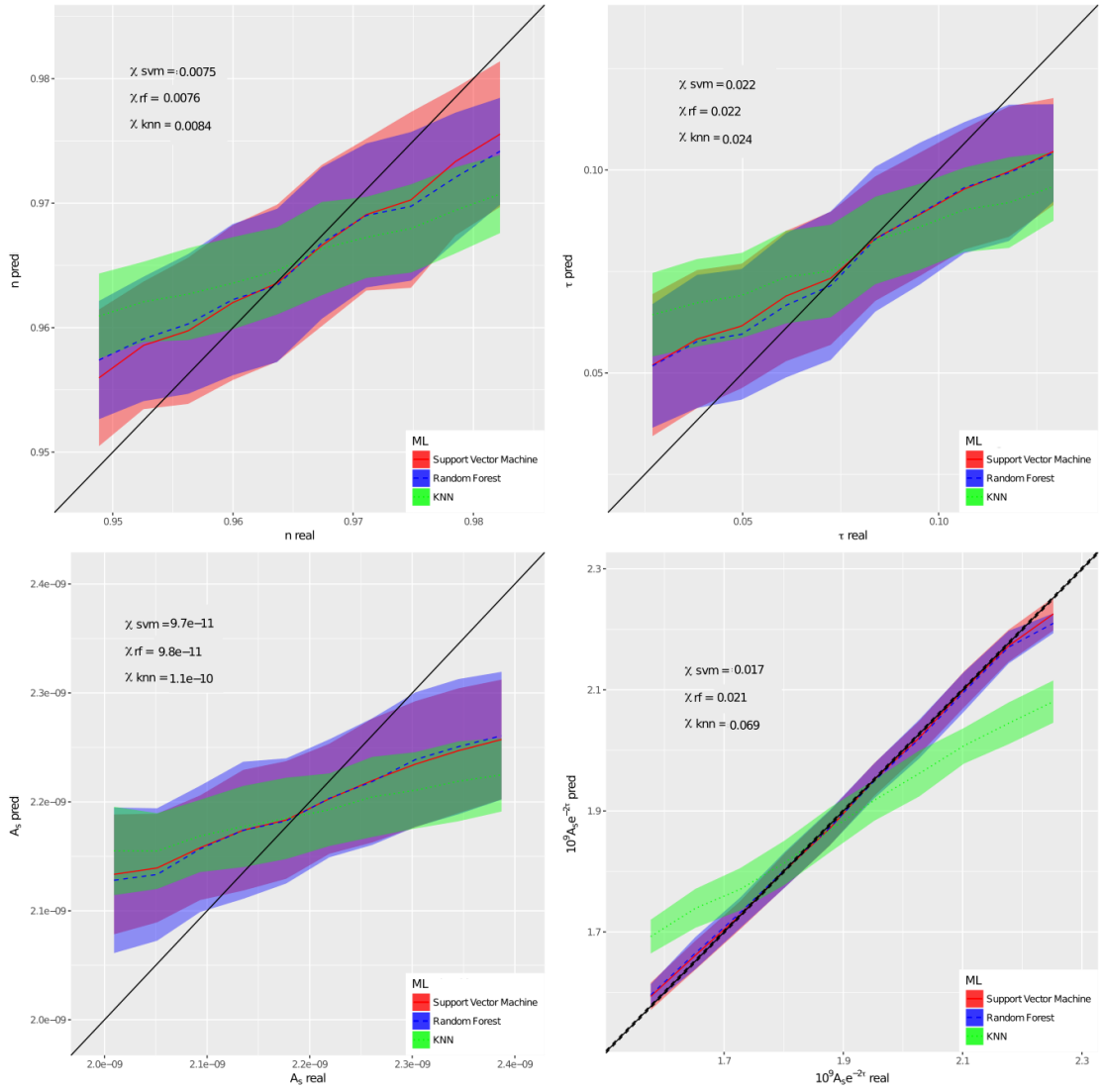

Figure 7.14: Comparison between the cosmological parameters predicted with CMB+Hubble-diagram features and the real ones for all cosmological models of the test set.

formation about the Hubble diagram, demonstrating the importance of adding this features in order to predict this parameter.

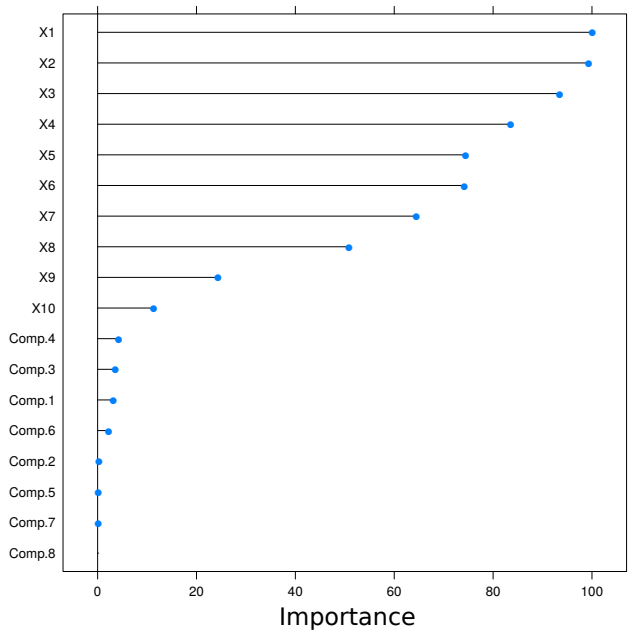

Figure 7.15: Relative importance of the features when estimating the Hubble constant.

7.4 Angular distribution of the cosmological parameters.

The standard cosmological model described in chapter 1 is based on 2 fundamental assumptions: homogeneity and spatial isotropy. Thanks to these 2 assumptions the Einstein equations are simplified and can be solved arriving at the FRW metric. As previously specified, although numerous observations support this model, still remain important details that are not fully understood. This is why it is necessary to test the different assumptions in which this model is supported. In particular, we know that the distribution of matter in the Universe is highly inhomogeneous, and that only reaches homogeneity when it is averaged on large scales ($\approx$ 100 Mpc) [10]. However, it is not clear which is the correct way to average tensors in the framework of general relativity [2, 6, 15], and if such inhomogeneities can affect the different cosmological observations.

While, according to the standard model, we should not find significant anisotropies in the values of the cosmological parameters, it is not clear if the inhomogeneities present in the Universe can affect these distributions. In previous works *Carvalho & Marques 2015* [3] studied the distribution of cosmological parameters using Hubble diagrams constructed with supernovae in different zones of the sky. Although they do find some anisotropies, these will not be large enough to generate a *backreaction* effect as the one proposed by *Kolb et al. 2006* [7] as an alternative to dark energy. Also, it is important to note that the sample of supernovas used in such analysis has a highly anisotropic angular distribution, which leads to the fact that the estimation of the parameters in some areas of the sky is through the fitting of very few supernovas and, therefore, have a big uncertainty. Bearing in mind that this problem is not present in the CMB, since we have homogeneous information (coming from the same instrument and calibrated in the same way) throughout the sky (with the possible exception of the contamination mask provided by Planck), we decided to study the angular distribution of the cosmological parameters using the temperature power spectra of different zones of the sky.

To be able to estimate the angular distribution of the cosmological parameters, we divide the sky into 192 frames corresponding to the pixels estimated by the software `Healpix` with a pixelization schema with $NSIDE = 4$. The figure 7.16 shows the areas of the sky that we will study, corresponding to the pixelization scheme of `Healpix` with $NSIDE = 4$. Then we estimate the power spectrum of each zone using the function `anafast` of the `Healpix` code. The figure 7.17 shows, as an example, the spectrum of an area of the sky compared to the *full-sky* spectrum . As it can be seen, there are notable differences between both spectra, this is mainly because the angular mask of the studied area (in this case the pixel shape given by `Healpix`) add a false signal that must be taken into account when making cosmological analysis with the spectra.

7.4.1 Reconstruction of the power spectra.

With the aim of disaffecting the spectrum of each zone from the signal introduced by the mask and so, reconstruct the original power spectrum that correspond to a given set of cosmological parameters, we use a machine learning technique called *Denoising AutoEnconders* (DAEs) that was previously implemented with success in the reconstruction of temporary signals with low signal-noise in other areas of science [16].

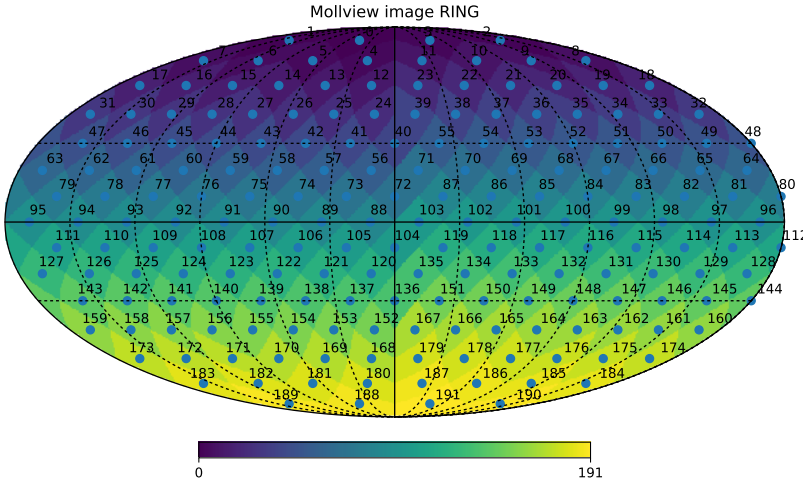

Figure 7.16: Analysed sky areas corresponding to the pixelization scheme of *Healpix* with $NSIDE = 4$.

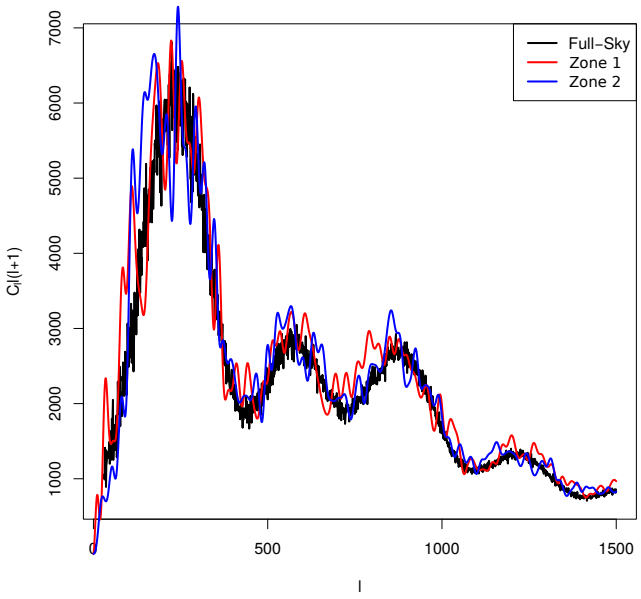

Figure 7.17: Comparison between the power spectrum of 2 randomly selected areas of the sky and the correspondent *full-sky* power spectrum.

This technique consists of a neural network (see chapter 2) whose input data is the

signal with the noise incorporated (in our case the value of each multipole of the spectrum of each area of the sky) and whose output is the signal without the noise (in our case, the value of each multipole of the original spectrum corresponding to a given cosmological model). Finally, we add a hidden layer with a number of neurons smaller than the number of input and output data. In figure 7.18 it can be seen an outline of the architecture used.

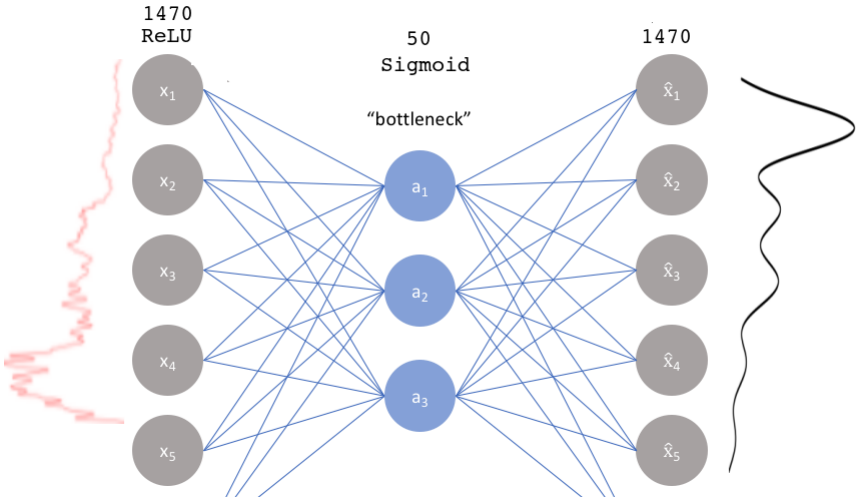

Figure 7.18: Architecture Scheme used in the neural network.

To build the data set necessary to train the neural network, we will use the previously simulated sky maps. We divide each map of the sky into 192 frames following the same pixelization scheme with $NSIDE = 4$. Then, we calculate the power spectrum for 20 randomly selected areas of the sky. Finally, our data set will consist of 20 spectra from different areas of the sky for each of the 5000 cosmologies, related to the corresponding theoretical spectra of each cosmological model. We again divide the data set into 2 independent sets, train the DAEs with one of this set and test it with the other one.

The figure 7.19 shows, as an example, the full-sky spectrum of a given model of the test set, superimposed to the spectrum of a randomly chosen area corresponding to the same model and to the reconstructed spectrum of that frame.

To quantify the reconstruction obtained through this method, we define the following statistic:

$$\Delta = \sum_{l=50}^{1500} (C_{l,real} - C_{l,rec})^2 \tag{7.18}$$

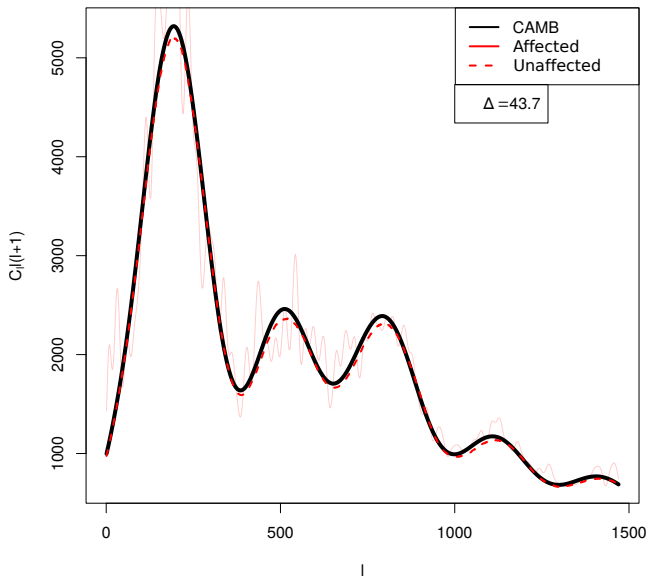

Figure 7.19: Example of the reconstruction of an spectrum of a zone of the sky corresponding to a pixel of the pixelization scheme of *Healpix* with $NSIDE = 4$.

where $C_{l,real}$ is the real value of the multipole in the theoretical spectrum of a given cosmological model and $C_{l,rec}$ is the value of the multipole in the reconstructed spectrum.

In figure 7.20 we present the histogram of Δ for all the cosmological models of the test set. It can be seen that, with the exception of a few spectra in the distribution tail, this technique reconstructs the power spectrum with a good signal-noise.

7.4.2 Estimation of the cosmological parameters.

Once we reconstruct the spectra of each area of the sky, we can apply the methods described in section 7.3 to estimate the cosmological parameters in each frame and found the angular distribution of those parameters.

Considering that the reconstruction of the spectra has associated errors, it is necessary to study how these errors can affect the estimation of the parameters. For this we estimate the cosmological parameters from the reconstructed spectra for the test set, and then we compare them with the real values. In figures 7.21 and 7.22 we plot the values of the cosmological parameters predicted with the reconstructed spectra of the test set vs the real values. It can be seen that the predictions get considerably worse if we compare them with the predictions made with the full-sky spectra (see figures 7.6 and 7.7). This means that

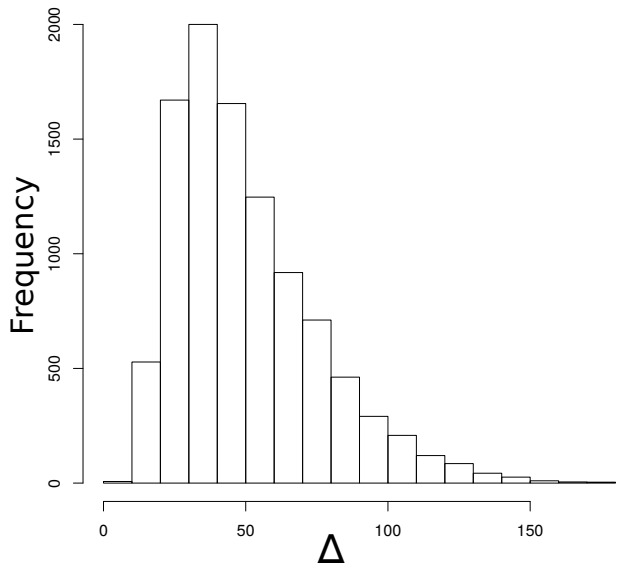

Figure 7.20: Histogram of Δ for all the cosmological models of the test set

Model	$\Omega_m h^2$	Ω_k	$\Omega_b h^2$	H_0	τ	n	A_s $[10^{-11}]$	$10^9 A_s e^{-2\tau}$
SVM	0.0031	0.0071	0.00029	2.9	0.020	0.007	8.8	0.016

Table 7.4: Results on the performance (χ) of the *Support Vector Machine* method when estimating the cosmological parameters in the different frames of the sky corresponding to the pixelization scheme of `Healpix` with $NSIDE = 4$.

small errors in the reconstruction of the spectra, lead to large errors in the prediction of the parameters. As you can see we only have a good estimation in Ω_k, τ, A_s and $10^9 A_s e^{-2\tau}$. In table 7.4 we show the values of the χ statistic for the estimation of the parameters in different areas of the sky using the *Support Vector Machines* algorithm.

7.4.3 Estimation of the angular distribution of the cosmological parameters in PLANCK data.

As was previously specified, the Planck data presents an homogeneous distribution throughout the sky with the exception of the contamination mask (see figure 7.23). While most of the cosmological analysis are made with the Planck maps excluding the areas of the

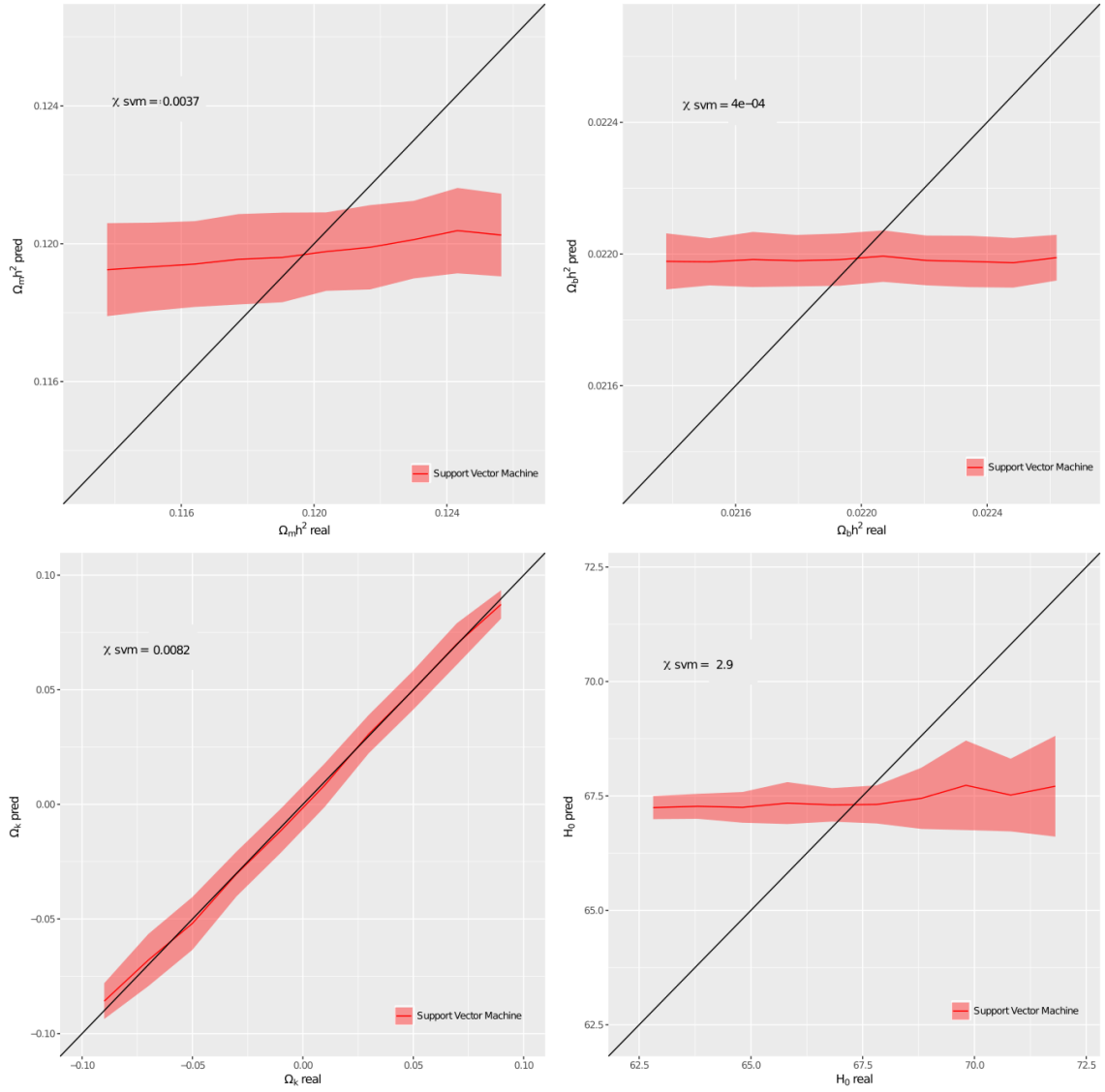

Figure 7.21: Comparison between the cosmological parameters predicted from the reconstructed spectra and the real ones for all the cosmological models of the test set.

contamination mask, in our work we decided to use the map of the whole sky provided by *Rogers et al.*[3] [12, 13] (see figure 7.24). These map was constructed through statistical methods that reconstruct the primary anisotropies in those areas of the sky that are on top of the contamination mask. It is important to note that this reconstruction is not perfect so

[3] http://www.silc-cmb.org/

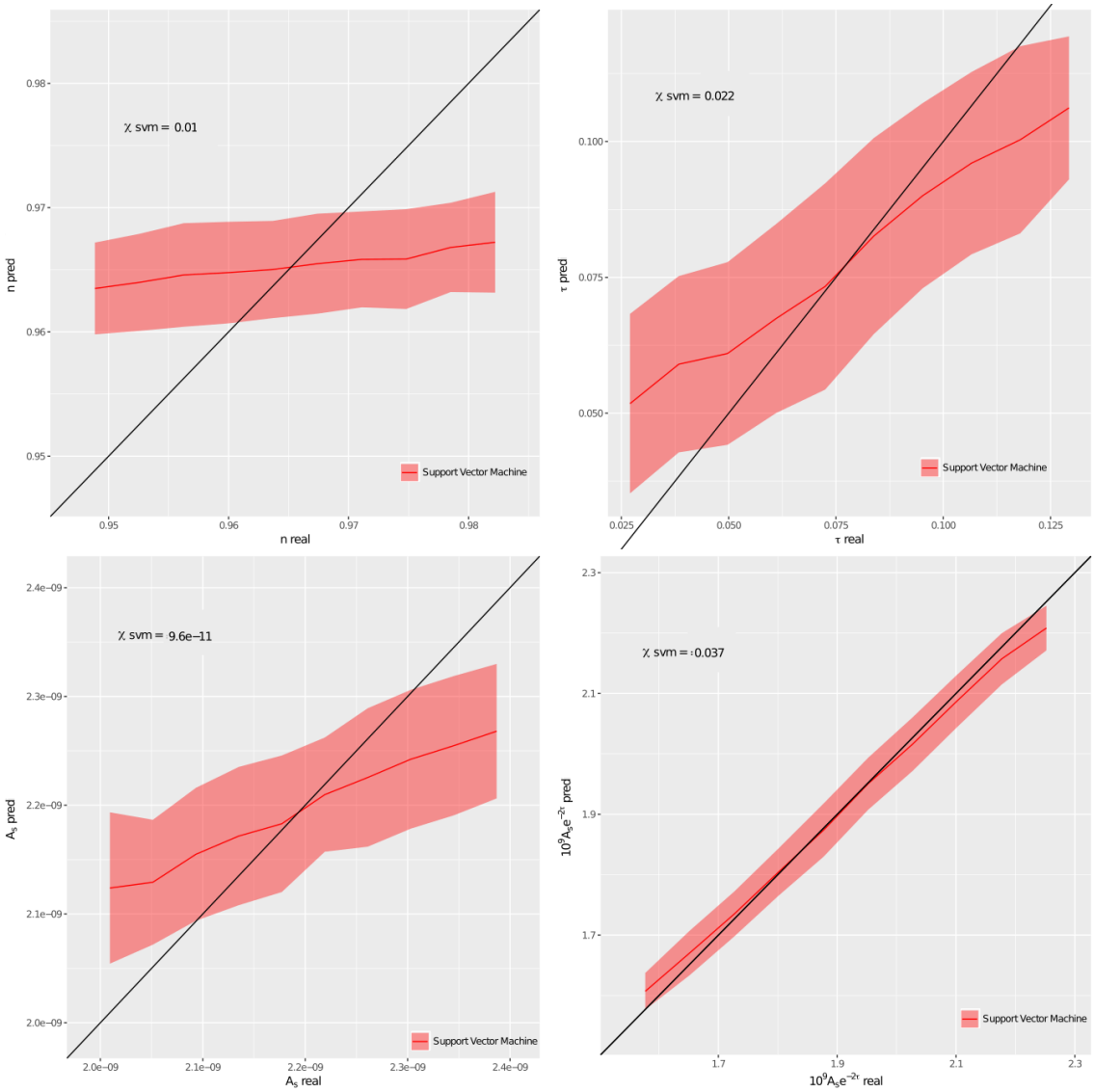

Figure 7.22: Comparison between the cosmological parameters predicted from the reconstructed spectra and the real ones for all the cosmological models of the test set.

the results obtained in sky areas that have a large percentage of contamination mask will be unreliable.

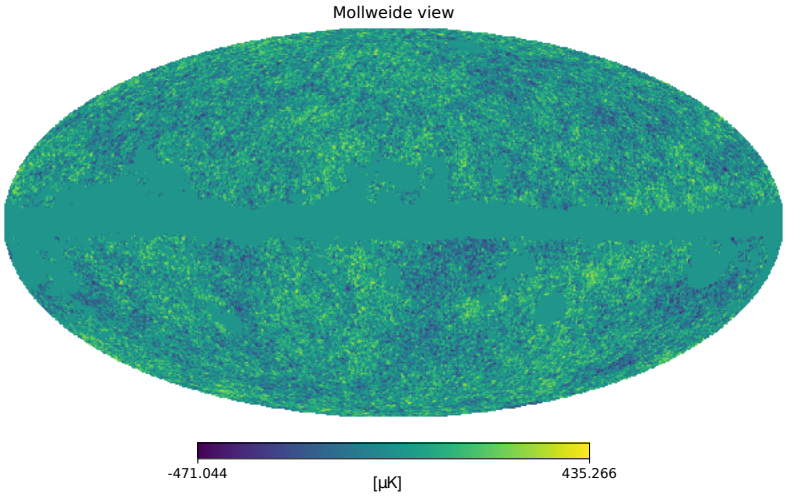

Figure 7.23: Angular map of the contamination mask.

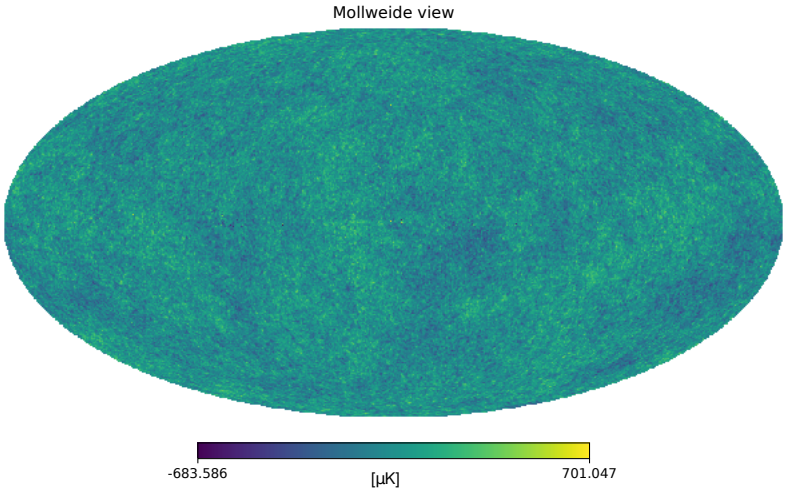

Figure 7.24: SILC angular map, provided by *Rogers et al.* [12, 13].

Using this map of the whole sky, we divide it into zones corresponding to the pixels of `Healpix` with $NSIDE = 4$ and then we estimate the power spectrum in each of these frames. In order to reconstruct the original power spectrum and disaffect it from the

signal introduced by the shape of each area of the sky, we applied the DAEs previously trained according to what was described in section 7.4.1. Once the spectra have been reconstructed, we estimate the corresponding cosmological parameters for each zone. Taking into account the results of the previous section, we will only estimate Ω_k, τ, A_s and $10^9 A_s e^{-2\tau}$.

In figure 7.25 we show the histograms of the 4 parameters for the 192 sky areas.

As it can be seen, the distributions have an approximately Gaussian form, however there are some extreme values that need to be studied in more detail to determine if they are significant.

To study this, we made 100 simulated maps of the CMB with the cosmological parameters found by Planck collaboration [11]. Then, we apply to each map the same procedure as to the real data, obtaining 100 distributions for each parameter corresponding to simulated maps of the standard cosmological model. In figure 7.25 you can see, in black lines, the distributions of each simulated map, while in red line is shown the distributions corresponding to the real data. In addition, in black vertical line we show the average of the estimated cosmological parameter with the full-sky map of each simulation, where the error band is given by the standard deviation of the estimated values. In red vertical line we shown the value estimated through the methods described in the section 7.3 for the full-sky map of Planck, where the error band corresponds to the χ statistic. For a better comparison we added, in the vertical green line, the value estimated by the Planck collaboration with its respective error [11].

As you can see, the results of the estimations with the full-sky map of the real data are totally consistent with the values found by the Planck collaboration and with the results obtained from the 100 simulation of the standard model, with the exception of the parameter $10^9 A_s e^{-2\tau}$ in which we found a slight inconsistency of $\approx 1.5\sigma$.

On the other hand we can see that the distributions found in the real data show notable differences with the distributions found in simulations. However, it is important to note that the estimations of the parameters in each area of the sky are affected by the uncertainties detailed in the table 7.4. Also as specified previously, the map used does not contain the corresponding contamination mask, so many pixels can be affected by various secondary anisotropies. To study if the differences between the distributions can come from these anisotropies, we estimate the percentage of the map of sky that is not affected

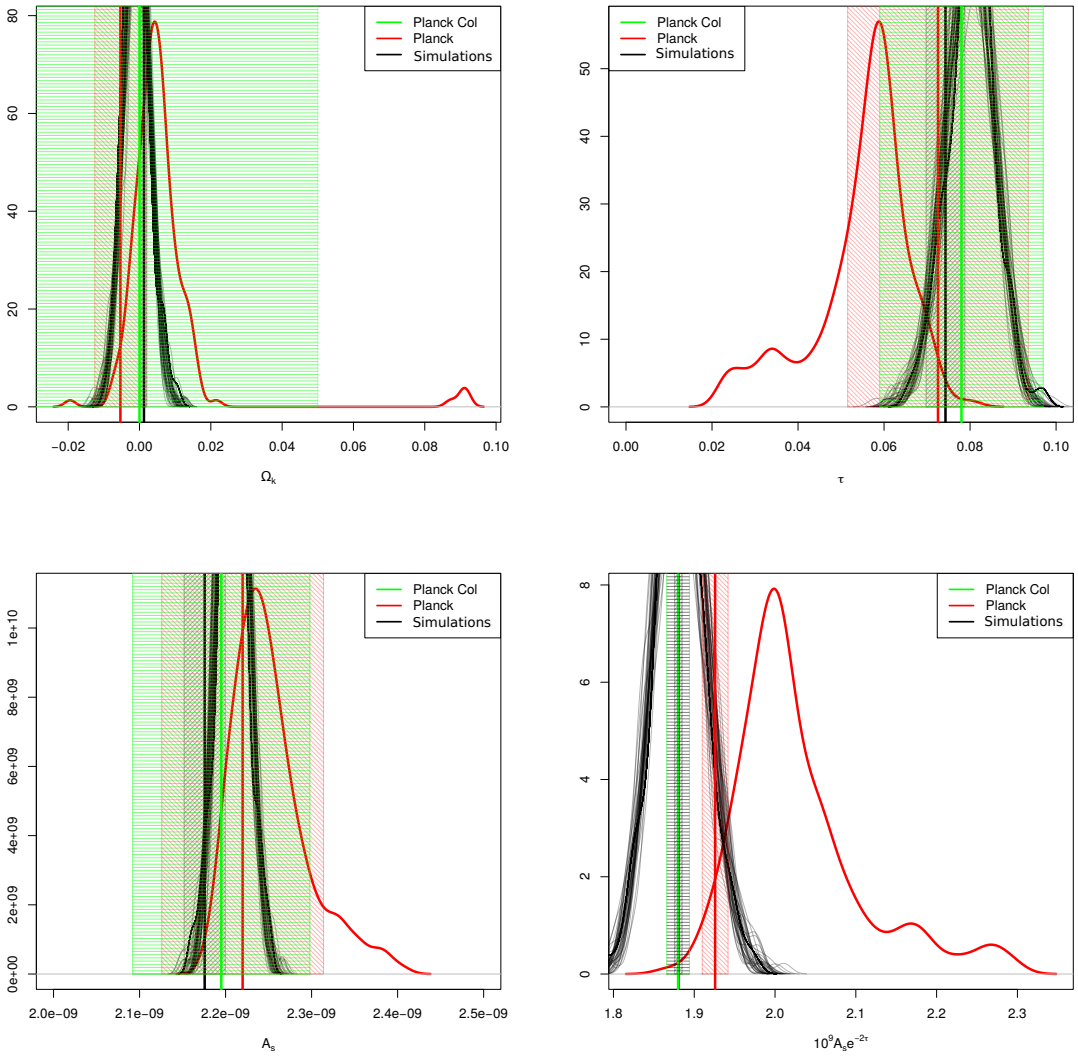

Figure 7.25: Histograms of the estimated cosmological parameters for the 192 areas of the sky. In black lines are shown the distributions of each simulated map, while in red lines are shown the distributions corresponding to the real data. Also, in vertical line is plotted the average of the cosmological parameter estimated with the full-sky map of each simulation, where the error band is given by the standard deviation of the estimated values. In red vertical line is shown the value estimated by the methods described in section 7.3 for the full-sky map of Planck, where the error band corresponds to the χ statistic. For a better comparison we added, in vertical green line, the value estimated by the Planck collaboration with its respective error [11].

by the Planck's contamination mask in each area of the sky. The figure 7.26 shows the estimations of the cosmological parameters as a function of the percentage of sky that is not affected by the contamination mask in each area of the sky. You can see that extreme values occur in areas of the sky that have a low percentage of sky unaffected. This shows that while the reconstruction of primary anisotropies carried out by *Rogers et al.* [12, 13] is good, small errors can be translated into bad estimations in the cosmological parameters.

With this in mind, we studied the distribution of the cosmological parameters on the SILC map with the Planck mask (see figure 7.23). The figure 7.27 shows the distributions of the parameters for the SILC maps with the Planck contamination mask (red lines). In black lines we plot the results for the simulated maps. For a better comparison we also added the contamination mask to these maps. Finally, the green distribution shows the distribution found in the Planck data adding the value of the χ statistic corresponding to the error detailed in the table 7.4.

It can be seen that once taken into account the uncertainty in the estimation of the parameters introduced by the machine learning method and the effects produced by the contamination mask, the results found in the actual Planck data are totally consistent with what is found in the standard model simulated map. This result shows that the anisotropies found in the different cosmological parameters are not significant and are totally consistent with the assumptions of homogeneity and isotropy.

7.5 Conclusions.

In this chapter, we present the results of the study of the CMB power spectrum through machine learning methods.

As was specified in previous chapters, the first step in the application of machine learning techniques is the construction of a data set representative of the problem we want to address. In our case, using the software `CAMB` that resolves numerically the Boltzmann-Einstein equations (Ec. 7.13), we build a data set that includes the power spectrum of 5000 cosmological models around Planck cosmology.

Using this data set, we first studied 2 unsupervised learning techniques, PCA and mixture of Gaussian. It was found that the spectra can be represented accurately by the value of the fist 8 principal components, which are related to the cosmological parameters

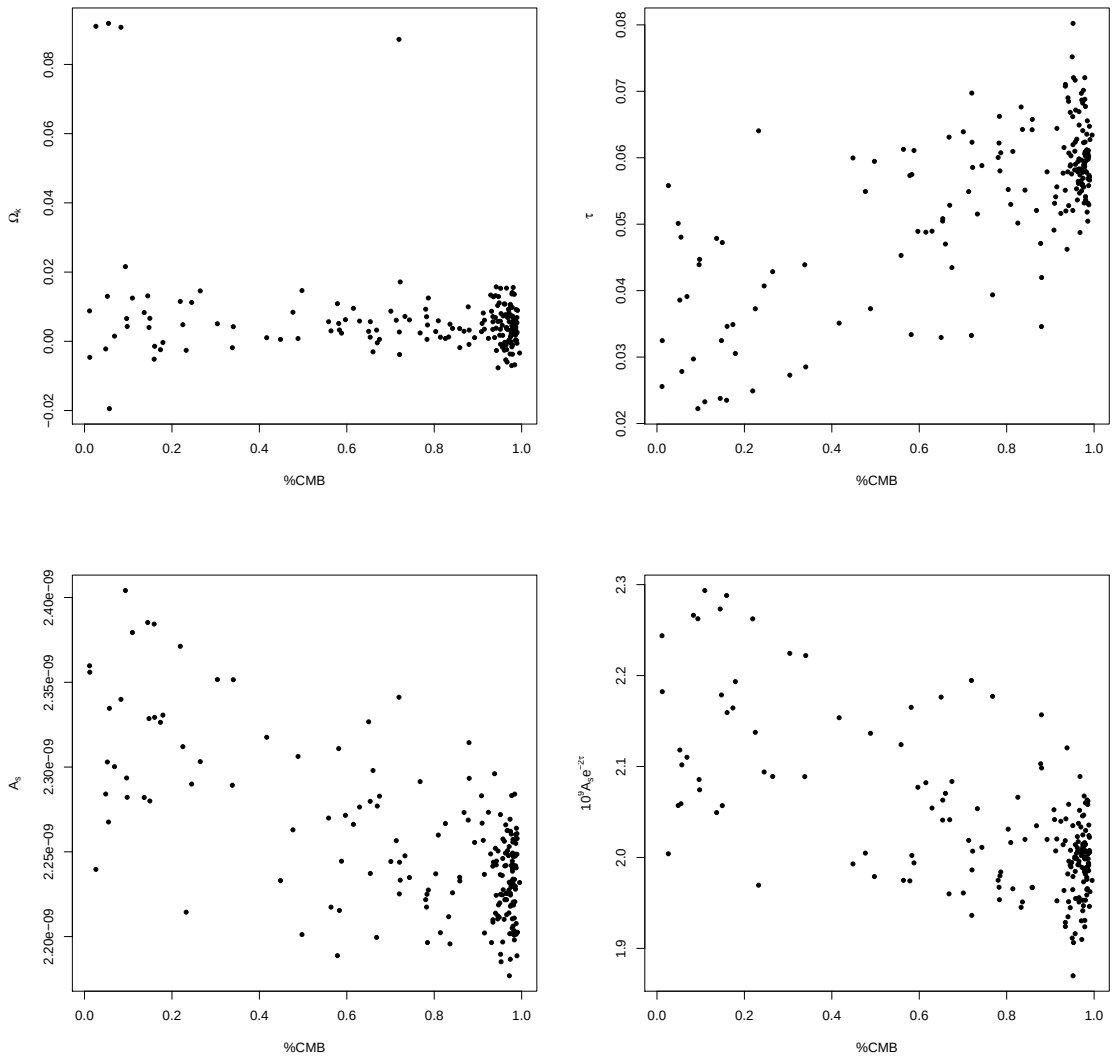

Figure 7.26: Value of the estimated cosmological parameter as a function of the percentage of the sky that is unaffected by the Planck contamination mask in each zone.

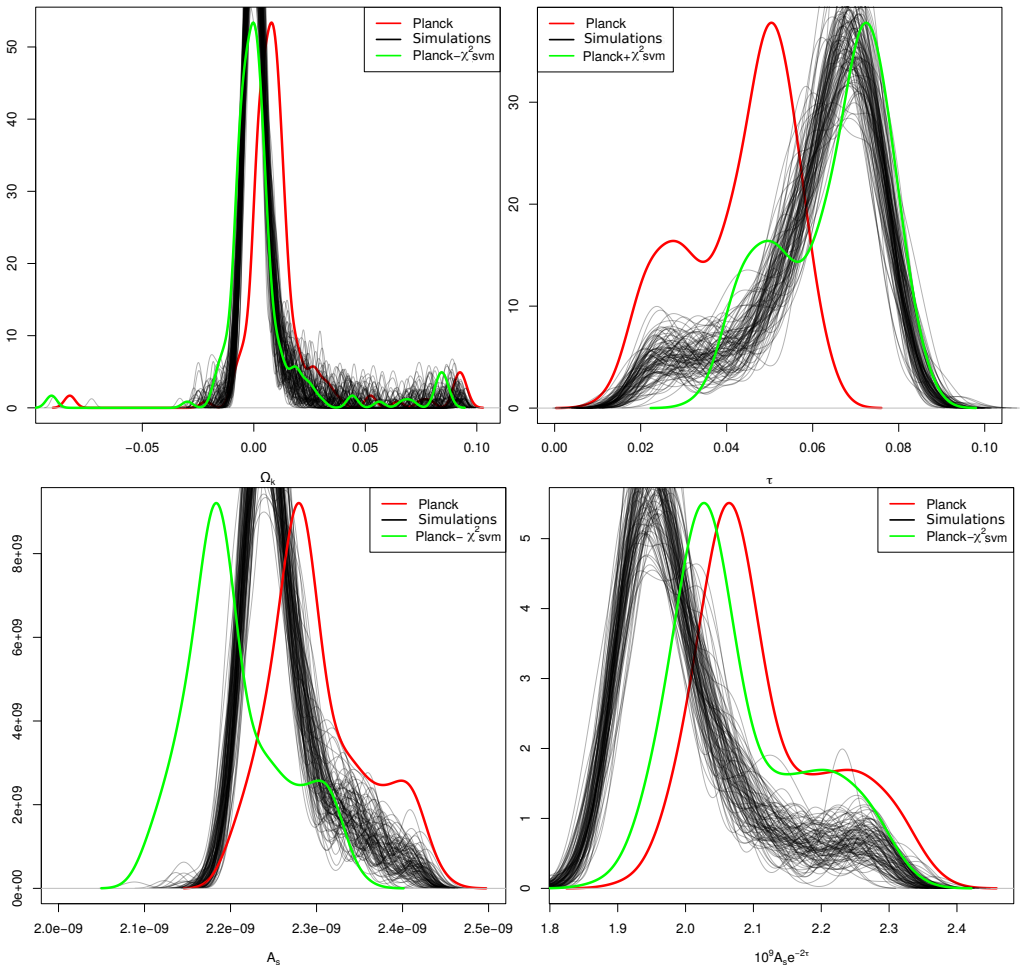

Figure 7.27: Histograms of the estimated cosmological parameters for the 192 areas of the sky with the contamination mask. In black lines are shown the distributions of each simulated map, while in red lines are shown the distributions corresponding to the real data. Finally, the green line shows the distribution found in Planck data adding the value of the χ statistical corresponding to the error detailed in the table 7.4.

(see figure 7.4). When making a mixture of Gaussian in the space of the 8 principal components, we discovered that those models that have the same curvature are grouped (see figure 7.5).

Starting with the supervised learning techniques, we studied the ability of these techniques to estimate the cosmological parameters using as features the first 8 principal components previously calculated. We find that these methods can estimate reliably all the parameters with the exception of the Hubble constant H_0 (see figures 7.6 and 7.7).

Taking into account the degenerancies present in the estimation of the cosmological parameters, we build a 'toy model' in which we analyse the improvement in the estimation when adding features with information about the Hubble diagram of supernovas (section 7.3.1). We found that adding this information considerably improves the estimations of the parameters, in particular the estimation of the Hubble constant (see figures 7.13 and 7.14). This demonstrates the ability of the machine learning methods to analyse features of different nature and break the present degenerancies.

In section 7.4 we studied the possibility of estimating the cosmological parameters in different areas of the sky. For this was necessary to build an automatic machine learning method that reconstructs the power spectra of each zone of the sky and eliminate the signal produced by the shape of the studied area. Although this reconstruction method has a good performance, small errors in the reconstruction lead to very bad estimations of some of the cosmological parameters. In particular, we could only estimate with good precision Ω_k, τ, A_s and $10^9 A_s e^{-2\tau}$ (see figure 7.8 and 7.9).

Finally, in section 7.4.3 we estimate the full-sky parameters and their angular distribution using Planck data. In particular, we work with the SILC map built by *Rogers et al.* [12, 13]. The values found by our methods for the full-sky map coincide completely with the values found by the Planck collaboration [11], with the exception of the parameter $10^9 A_s e^{-2\tau}$ where we found a slight inconsistency of $\approx 1.5\sigma$. On the other hand, we found differences between the angular distributions of the parameters in the real data compared with the 100 simulated maps. However, we were able to demonstrate that these differences can be explained completely due to the contamination present in some pixels. In particular, if we analyse the Planck data with the contamination mask we do not found differences with which is found in the simulated map (see figure 7.27).

From the results presented in this chapter we can conclude that there are no evidences

of significant anisotropies in the angular distribution of the cosmological parameters.

Bibliography

[1] M. Betoule, R. Kessler, J. Guy, and J. Mosher. Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples. *A&A*, 568:A22, August 2014. doi: 10.1051/0004-6361/201423413.

[2] Thomas Buchert, Martin Kerscher, and Christian Sicka. Back reaction of inhomogeneities on the expansion: The evolution of cosmological parameters. *PhRvD*, 62: 043525, August 2000. doi: 10.1103/PhysRevD.62.043525.

[3] C. S. Carvalho and K. Marques. Angular distribution of cosmological parameters as a probe of space-time inhomogeneities. *ArXiv e-prints*, December 2015.

[4] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. *ApJ*, 622:759–771, April 2005. doi: 10.1086/427976.

[5] David W. Hogg. Distance measures in cosmology. *arXiv e-prints*, art. astro-ph/9905116, May 1999.

[6] Akihiro Ishibashi and Robert M. Wald. Can the acceleration of our universe be explained by the effects of inhomogeneities? *Classical and Quantum Gravity*, 23: 235–250, January 2006. doi: 10.1088/0264-9381/23/1/012.

[7] E. W. Kolb, S. Matarrese, and A. Riotto. On cosmic acceleration without dark energy. *New Journal of Physics*, 8:322, December 2006. doi: 10.1088/1367-2630/8/12/322.

[8] Max Kuhn. Building predictive models in r using the caret package. *Journal of*

Statistical Software, Articles, 28(5):1–26, 2008. ISSN 1548-7660. doi: 10.18637/jss.v028.i05. URL https://www.jstatsoft.org/v028/i05.

[9] C.-P. Ma and E. Bertschinger. Cosmological Perturbation Theory in the Synchronous and Conformal Newtonian Gauges. *ApJ*, 455:7, December 1995. doi: 10.1086/176550.

[10] P. Ntelis. The Homogeneity Scale of the universe. *ArXiv e-prints*, July 2016.

[11] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2015 results. XIII. Cosmological parameters. *A&A*, 594:A13, September 2016. doi: 10.1051/0004-6361/201525830.

[12] Keir K. Rogers, Hiranya V. Peiris, Boris Leistedt, Jason D. McEwen, and Andrew Pontzen. SILC: a new Planck internal linear combination CMB temperature map using directional wavelets. *MNRAS*, 460:3014–3028, August 2016. doi: 10.1093/mnras/stw1121.

[13] Keir K. Rogers, Hiranya V. Peiris, Boris Leistedt, Jason D. McEwen, and Andrew Pontzen. Spin-SILC: CMB polarization component separation with spin wavelets. *MNRAS*, 463:2310–2322, December 2016. doi: 10.1093/mnras/stw2128.

[14] Luca Scrucca, Michael Fop, Thomas Brendan Murphy, and Adrian E. Raftery. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal*, 8(1):205–233, 2016.

[15] David L. Wiltshire. Average observational quantities in the timescape cosmology. *PhRvD*, 80:123512, December 2009. doi: 10.1103/PhysRevD.80.123512.

[16] Peng Xiong, Hongrui Wang, Ming Liu, Suiping Zhou, Zengguang Hou, and Xiuling Liu. Ecg signal enhancement based on improved denoising auto-encoder. *Engineering Applications of Artificial Intelligence*, 52:194 – 202, 2016. ISSN 0952-1976. doi: https://doi.org/10.1016/j.engappai.2016.02.015. URL http://www.sciencedirect.com/science/article/pii/S0952197616300331.

Chapter 8

Study on the interaction between dark matter and the CMB photons.

As was introduced in chapter 1 one of the main properties of dark matter is that it only interacts with the other standard model particles through gravity, and maybe through weak interactions. Current observations support this assumption by putting upper limits in the magnitude of other possible interactions. However, numerous studies show that even an small interaction within the allowed limits, can have observational consequences in different cosmological scales [4, 12, 13], and even relieve some problems of the standard cosmological model [2, 8, 9].

In this chapter we present the preliminary results of a study on the impact of an interaction between dark matter and CMB photons. In particular we developed a code that implements the Boltzmann formalism [3] to study the variation on the energy distribution of the CMB photons when interacting with any massive particle, generating a similar effect to the Sunyaev-Zeldovich [10, 11].

It is worth to remark that this work was done in collaboration with Dra. Celine Bœhm and Dr. Mariano Dominguez.

8.1 Redistribution function for the generalised Sunyaev-Zeldovich effect.

The Sunyaev-Zeldovich effect is the change in the CMB black-body spectrum due to the interactions between CMB photons and the electrons observed in the ICM. These small spectral variations are presented as secondary anisotropies of the CMB, and so it is necessary to model them before doing cosmological analysis with the CMB. On the other hand, considering that this effect is independent of the redshift, it is a powerful method to search for high-redshift galaxy clusters [7].

Traditionally the Sunyaev-Zeldovich effect is calculated by the radiative transfer method of Chandrasekar [1], however it was shown by *Lavalle & Bœhm* [3] that the estimation of the Sunyaev-Zeldovich effect using the the collisional method, taking into account the Boltzmann equation, is totally equivalent to the traditional one. Nevertheless, it is important to note that the collisional method has the advantage of being easily generalised to other interactions between photons and massive particles.

Considering Boltzmann's collisional formalism, the change in the black-body spectrum of the CMB photons due to the interactions with the electrons present in the ICM is (see equation 5 of *Bœhm & Lavalle 2009*[3]):

$$\begin{aligned}\Delta I_\gamma(E_{k_1}) &= -\int dl \int \frac{d^3\vec{p_1}}{(2\pi)^3}\frac{d^3\vec{p_2}}{(2\pi)^3}\frac{d^3\vec{k_2}}{(2\pi)^3}\frac{(2\pi)^4\delta^4(p_1+k_1-p_2-k_2)}{4E_{p_2}E_{k_2}}\frac{M^2}{4E_{k_1}E_{p_1}} \\ &* \quad \{I_\gamma(E_{k_1})f(E_{p_1},\vec{x}) - \frac{E_{k_1}^3}{E_{k_2}^3}I_\gamma(E_{k_2})f(E_{p_2},\vec{x})\} \end{aligned} \tag{8.1}$$

where $k_{1,2}$ are the quadri-vectors corresponding to the incoming and outgoing photons and $p_{1,2}$ are the quadri-vectors corresponding to the incoming and outgoing electrons. On the other hand, I_γ^0 is the black-body distribution, M^2 is the amplitude of the interaction (in this case the Compton scattering) and $f(E_{p_1},\vec{x})$ is the energy distribution of the electrons.

It is clear from equation 8.1 that to study the generalised Sunyaev-Zeldovich effect (SZG) produced by the interaction between photons and any massive particle with a given energy distribution, we need just to put in equation 8.1 the correct M^2 amplitude of the interaction and the correct energy distribution $f(E_{p_1},\vec{x})$ of the particle in study.

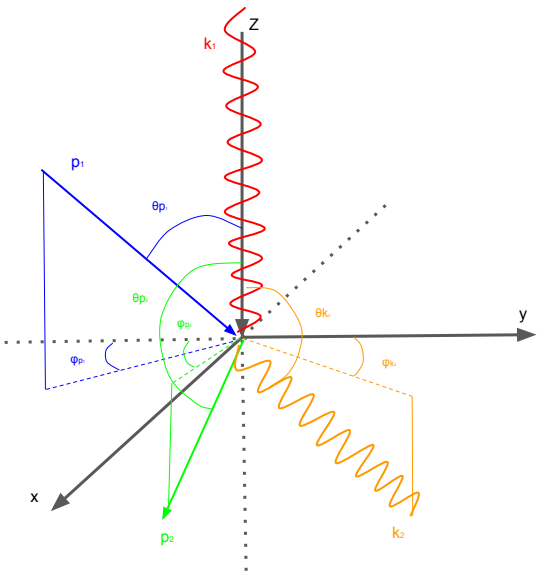

Figure 8.1: Coordinate reference system for the elastic two-body interaction.

From now on we will work in a general way assuming a given amplitude M^2 and a given energy distribution $f(E_{p1}, \vec{x})$. To simplify equation 8.1 it is convenient to work in a coordinate system like the one shown in figure 8.1

In this reference system the quadri-vectors are:

$$\begin{aligned} p_1^\mu &= (E_{p_1}, p_1 sin\theta_{p_1} cos\phi_{p_1}, p_1 sin\theta_{p_1} sin\phi_{p_1}, p_1 cos\theta_{p_1}) \\ p_2^\mu &= (E_{p_2}, p_2 sin\theta_{p_2} cos\phi_{p_2}, p_2 sin\theta_{p_2} sin\phi_{p_2}, p_2 cos\theta_{p_2}) \\ k_1^\mu &= (E_{k_1}, 0, 0, k_1) \\ k_2^\mu &= (E_{k_2}, k_2 sin\theta_{k_2} cos\phi_{k_2}, k_2 sin\theta_{k_2} sin\phi_{k_2}, k_2 cos\theta_{k_1}) \end{aligned} \tag{8.2}$$

In addition, taking into account that for a massive particle $p^\mu p_\mu = m^2$ and that for photons $k^\mu k_\mu = 0$, we get (see equations 31-36 of *Bœhm & Lavalle 2009*[3]):

$$\begin{aligned} p_{1,2} &= \sqrt{E_{p_{1,2}}^2 - m^2} \\ k_{1,2} &= E_{k_{1,2}} \end{aligned} \tag{8.3}$$

If we restrict ourselves to studying low density environments, such as the ICM, we can assume that the photons only interact once with the massive particle and so, we can approximate $I_\gamma \approx I^0_\gamma$. If we also take into account the restrictions imposed by the δ^4 factor on the energy of the outgoing particles, we get:

$$\begin{aligned}
\Delta I_\gamma(E_{k_1}) &= -\frac{1}{4(2\pi)^3 E_{k_1}} \int dl \int p_1^2 dp_1 d\Omega_{p_1} \left[\frac{p_2^2 dp_2 d\Omega_{p_2}}{(2\pi)^3} \frac{k_2^2 dk_2}{(2\pi)^3} \frac{(2\pi)^4 \delta^4(p_1 + k_1 - p_2 - k_2)}{4E_{p_2} E_{k_2}} \right] d\Omega_{k_2} \frac{M^2}{E_{p_1}} \\
&* \quad \{ I^0_\gamma(E_{k_1}) f(E_{p_1}, \vec{x}) - \frac{E^3_{k_1}}{E^3_{k_2}} I^0_\gamma(E_{k_2}) f(E_{p_2}, \vec{x}) \} \\
&= -\frac{1}{4(2\pi)^3 E_{k_1}} \int dl \int p_1^2 dp_1 d\Omega_{p_1} \left[\frac{\alpha t^2}{4(2\pi)^2 (1 - \beta\mu)} \right] d\Omega_{k_2} \frac{M^2}{E_{p_1}} \\
&* \quad \{ I^0_\gamma(E_{k_1}) f(E_{p_1}, \vec{x}) - \frac{I^0_\gamma(t E_{k_1}) f(E_{p_2}, \vec{x})}{t^3} \}
\end{aligned} \tag{8.4}$$

where,

$$\begin{aligned}
t &= \frac{E_{k_2}}{E_{k_1}} = \frac{(1 - \beta\mu)}{(1 - \beta\mu') + \alpha(1 - \Delta)} \\
\alpha &= \frac{E_{k_1}}{E_{p_1}} \\
\mu &= \vec{p_1}.\vec{k_1}/|\vec{p_1}||\vec{k_1}| = cos\theta_{p_1} \\
\mu' &= \vec{p_1}.\vec{k_2}/|\vec{p_1}||\vec{k_2}| = cos\theta_{p_1} cos\theta_{k_2} + sin\theta_{p_1} sin\theta_{k_2} cos(\phi_{p_1} - \phi_{k_2}) \\
\Delta &= \vec{k_1}.\vec{k_2} = cos\theta_{k_2} \\
E_{p_2} &= E_{p_1} + E_{k_1} - tE_{k_1} \\
p_2 &= \sqrt{E^2_{p_2} - m^2} \\
sin\theta_{p_2} &= \frac{\sqrt{(p_1 sin\theta_{p_1} cos\phi_{p_1} - k_2 sin\theta_{k_2} cos\phi_{k_2})^2 + (p_1 sin\theta_{p_1} sin\phi_{p_1} - k_2 sin\theta_{k_2} sin\phi_{k_2})^2}}{p_2}
\end{aligned} \tag{8.5}$$

Making a change of variables $p_1 \mapsto E_{p_1}$ and taking into account that $p_1^2 = E^2_{p_1} - m^2 \Rightarrow 2p_1 dp_1 = 2E_{p_1} dE_{p_1} \Rightarrow \frac{p_1^2 dp_1}{E_{p_1}} = p_1 dE_{p_1}$ we obtain:

$$\begin{aligned}\Delta I_\gamma(E_{k_1}) &= -\frac{1}{16(2\pi)^5}\int dl \int p_1 dE_{p_1} d\Omega_{p_1} t^2 d\Omega_{k_2} \frac{M^2}{(1-\beta\mu)E_{p_1}} \\ &* \quad \{I^0_\gamma(E_{k_1}) f(E_{p_1},\vec{x}) - \frac{I^0_\gamma(tE_{k_1}) f(E_{p_2},\vec{x})}{t^3}\} \end{aligned} \tag{8.6}$$

Finally, to estimate the change in the black-body spectrum produced by a given interaction, it is enough to compute the amplitude M^2, the corresponding energy distribution and to perform the integral described in the equation 8.6.

If we also assume that we can decompose the distribution function of the massive particles into a spatial component and a component related with the energy, and that the spatial component will not be affected by the interactions with low energy photons (as in the case of the CMB photons), we get:

$$\begin{aligned}\Delta I_\gamma(E_{k_1}) &= -\frac{1}{16(2\pi)^5}\int f(\vec{x}) dl \int p_1 dE_{p_1} d\Omega_{p_1} t^2 d\Omega_{k_2} \frac{M^2}{(1-\beta\mu)E_{p_1}} \\ &* \quad \{I^0_\gamma(E_{k_1}) f(E_{p_1}) - \frac{I^0_\gamma(tE_{k_1}) f(E_{p_2})}{t^3}\} \\ \Delta I_\gamma(E_{k_1}) &= \tilde{\tau}\delta I_\gamma(E_{k_1}) \end{aligned} \tag{8.7}$$

It is important to remark that, depending on the particle mass, the interaction amplitude and the photons energy, the assumptions used previously may not be correct, and so, it will be necessary a detailed analysis of those cases. For example, if we work with a light dark matter particle, we will need a high numeric density in order to reach the mass observed in galaxy clusters ($\approx 10^{14} M_\odot$), and so it may affect the one-scattering assumption. On the other hand, if we have a strong interaction (high M^2), it may affect the spatial distribution of the dark matter particles [2, 8, 9].

In figure 8.2 it can be seen the energy redistribution function for the interaction between the CMB photons and the free electrons in the ICM that follows a Maxwell-Boltzmann energy distribution, estimated with the collisional method. As a comparison we show the same energy redistribution function, but calculated using the traditional radiative transfer method. It can be seen the the redistribution function estimated by both method are similar, demonstrating that these methods are completely equivalent.

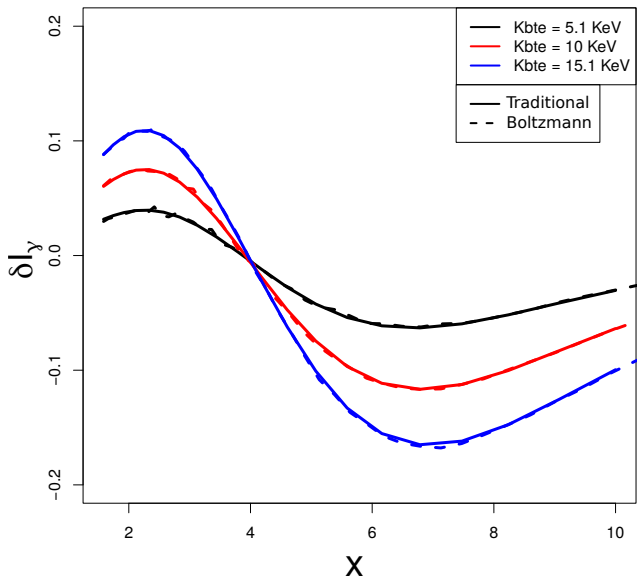

Figure 8.2: Energy redistribution function for the interaction between CMB photons and electrons in the ICM that follow a Maxwell-Boltzmann energy distribution with different temperatures. In dotted lines are plotted the results found using the Boltzmann method, while in lines are shown the results obtained through the traditional radiative transfer method.

From equation 8.6 it can be seen that if we change the energy distribution of the electrons, the effect produced on the CMB photons will change. In particular, in figure 8.3 is shown the energy redistribution of the CMB photons if we take into account the results found by *Lavalle et al. 2010* [5]. They study the SZ effect that will produce a population of electrons with a non-thermal energy distribution that comes from the annihilation between dark matter particles (See equation 4.9 of *Lavalle et al. 2010* [5]).

As it can be seen, although the redistribution function estimated through the collisional method follows the same trend as the one estimated through the traditional method, it have some instabilities. It is important to remark that because of computational time the integrals were done numerically using the function `mcquad` from the `scikit-monaco` library of `Python`. This function uses Monte-Carlo methods in order to perform the integrals, and so, it could introduce the observed instabilities.

Taking into account the definitions of the equation 8.5, we can define the cross-section for an elastic two-body interaction as (see equation 51 of *Bœhm & Lavalle 2009*[3]):

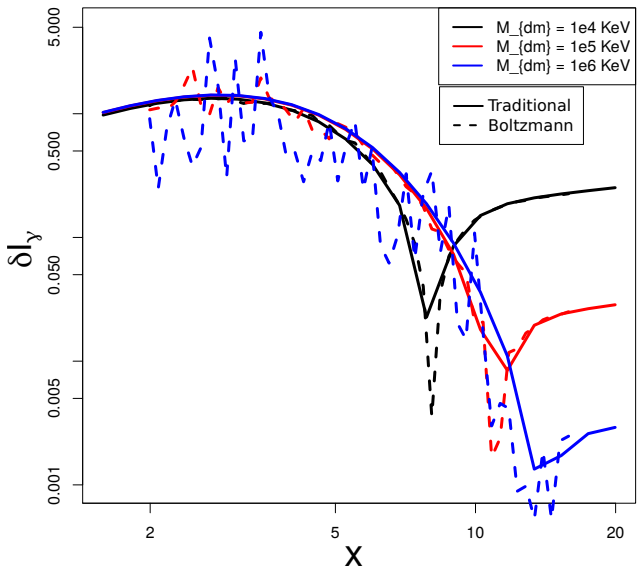

Figure 8.3: Energy redistribution function for the interaction between CMB photons and electrons coming from the annihilation between dark matter particles that follow a non-thermal energy distribution with different temperatures. In dotted lines are plotted the results found using the Boltzmann method, while the results obtained with the traditional radiative transfer method are plotted in solid lines.

$$d^2\sigma = \frac{(2\pi)^4}{4p_1k_1}\left[\frac{d^3\vec{p_2}}{(2\pi)^3 2E_{p_2}}\frac{d^3\vec{k_2}}{(2\pi)^3 2E_{k_2}}\right]\delta^4(p_1+k_1-p_2-k_2)|M^2| \tag{8.8}$$

this cross-section depends on the energy of the interacting particles (p_1 and K_1) and on the angle between their velocities (θ_{p_1}) and quantifies the strength of the interaction. In figure 8.4 it can be seen the cross-section for the Compton scattering between photons and electrons normalised to the theoretical value of $\sigma_{th} = 6.652458\ 10^{-29} m^2$.

It can be seen that, through the collisional method we recover the correct value for the cross-section of this interaction for all the values of p_1, k_1 and θ_{p_1}.

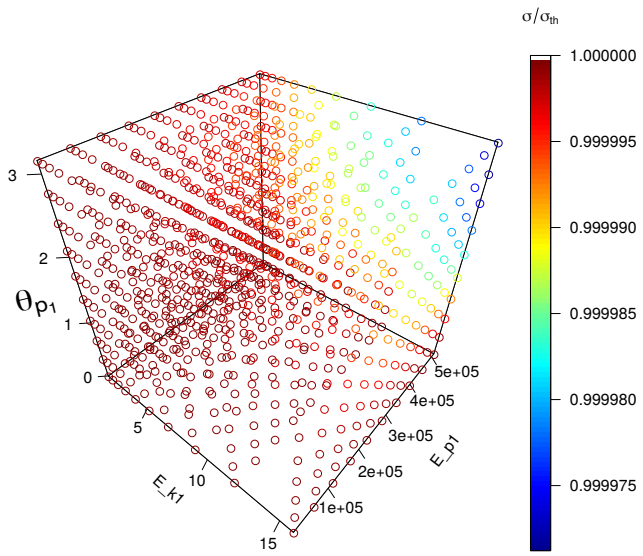

Figure 8.4: Thomson cross-section σ_{th} estimated through the collisional method and normalised to the theoretical value.

8.2 Study on the interaction between dark matter and photons.

8.2.1 Dipolar dark matter model.

As was described in chapter 1 a dark matter candidate must be neutral. The dipole dark matter model, give us a neutral candidate that have an interaction with the photons through their magnetic and electric dipoles. In this section we will study the consequences that these interaction will have on the energy function of the CMB photons.

The Lagrangian for the dipolar dark matter model is:

$$\mathcal{L}_{\gamma,\chi} = -\frac{i}{2}\bar{\chi}\sigma_{\mu,\nu}(\mathcal{M} + \gamma_5\mathcal{D})F^{\mu,\nu} \tag{8.9}$$

where $\sigma_{\mu,\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu]$ and $F^{\mu,\nu} = \partial^\nu A^\mu - \partial^\mu A^\nu$. On the other hand, $\mathcal{M}$ and $\mathcal{D}$ are parameters that quantify the magnetic and electric dipole respectively and $\bar{\chi}$ is the corresponding dark matter particle spinnor.

To study the interactions between dark matter and photons, it is necessary to take into account the Feynman diagrams plotted in figure 8.5

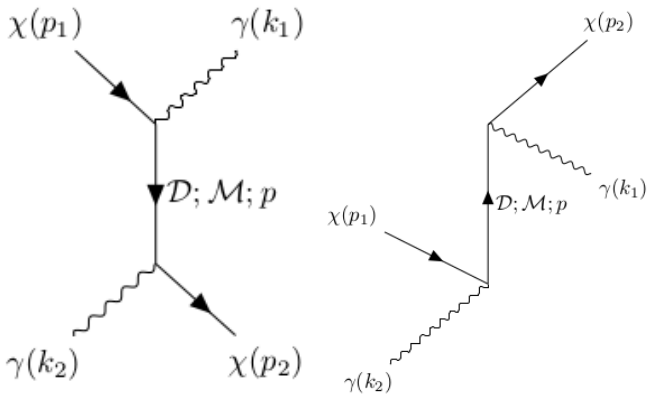

Figure 8.5: Feymann diagrams representative of the interactions between dark matter and photons.

Interaction with the magnetic dipole.

Taking into account the contribution of the magnetic dipole ($\mathcal{M}$) present in the Lagrangian described in equation 8.9, we can estimate the amplitude of the interaction represented in the Feymann diagram (see figure 8.5 *left*) as:

$$\begin{aligned} M^2_{mag} &= Tr[M_{mag}M^*_{mag}] \\ M^2_{mag} &= Tr[\bar{u}(p_2)\gamma_\mu\gamma_\nu k_2^\mu \epsilon_2^\nu(\not{p}+m)\gamma_\rho\gamma_\sigma k_1^\rho \epsilon_1^\sigma u(p_1)\bar{u}(p_1)\epsilon_1^{\sigma'} k_1^{\rho'}\gamma_{\sigma'}\gamma_{\rho'}(\not{p}+m)\epsilon_2^{\nu'} k_2^{\mu'}\gamma_{\nu'}\gamma_{\mu'}u(p_2)] \\ M^2_{mag} &= [(\not{p}_2+m_2)\not{k}_2\gamma_\nu(\not{p}+m)\not{k}_1\gamma_\sigma(\not{p}_1+m_1)\gamma_{\sigma'}\not{k}_1(\not{p}+m)\gamma_{\nu'}\not{k}_2]\eta^{\nu\nu'}\eta^{\sigma\sigma'} \end{aligned}$$

Taking into account that the trace of an odd number of γ matrix is 0 and that contracting the dumb indexes ν' and σ', the resulting amplitude is the sum of 8 terms, $M^2 = I + II + III + IV + V + VI + VII + VIII$, where:

$$
\begin{aligned}
I &= Tr[\not p_2 \not k_2 \gamma_\nu \not p \not k_1 \gamma_\sigma m_1 \gamma^\sigma \not k_1 m \gamma^\nu \not k_2] \\
II &= Tr[\not p_2 \not k_2 \gamma_\nu m \not k_1 \gamma_\sigma \not p_1 \gamma^\sigma \not k_1 m \gamma^\nu \not k_2] \\
III &= Tr[\not p_2 \not k_2 \gamma_\nu m \not k_1 \gamma_\sigma m_1 \gamma^\sigma \not k_1 \not p \gamma^\nu \not k_2] \\
IV &= Tr[m_2 \not k_2 \gamma_\nu \not p \not k_1 \gamma_\sigma \not p_1 \gamma^\sigma \not k_1 m \gamma^\nu \not k_2] \\
V &= Tr[m_2 \not k_2 \gamma_\nu \not p \not k_1 \gamma_\sigma m_1 \gamma^\sigma \not k_1 \not p \gamma^\nu \not k_2] \\
VI &= Tr[m_2 \not k_2 \gamma_\nu m \not k_1 \gamma_\sigma \not p_1 \gamma^\sigma \not k_1 \not p \gamma^\nu \not k_2] \\
VII &= Tr[\not p_2 \not k_2 \gamma_\nu \not p \not k_1 \gamma_\sigma \not p_1 \gamma^\sigma \not k_1 \not p \gamma^\nu \not k_2] \\
VIII &= Tr[m_2 \not k_2 \gamma_\nu m \not k_1 \gamma_\sigma m_1 \gamma^\sigma \not k_1 m \gamma^\nu \not k_2]
\end{aligned}
$$

As it can be noticed, the terms I, III, IV, V, VI and $VIII$ are identically 0 because $\not k_{1,2} \not k_{1,2} = k_{1,2}^2 = 0$. Then we just have to calculate the II and VII variables.

- II:

$$
\begin{aligned}
II &= Tr[\not p_2 \not k_2 \gamma_\nu m \not k_1 \gamma_\sigma \not p_1 \gamma^\sigma \not k_1 m \gamma^\nu \not k_2] \\
II &= -2m^2 Tr[\not p_2 \not k_2 \gamma_\nu \not k_1 \not p_1 \not k_1 \gamma^\nu \not k_2] \\
II &= 4m^2 Tr[\not p_2 \not k_2 \not k_1 \not p_1 \not k_1 \not k_2] \\
II &= 4m^2 [p_2.k_2 Tr(\not k_1 \not p_1 \not k_1 \not k_2) \\
&\quad - p_2.k_1 Tr(\not k_2 \not p_1 \not k_1 \not k_2) \\
&\quad + p_2.p_1 Tr(\not k_2 \not k_1 \not k_1 \not k_2) \\
&\quad - p_2.k_1 Tr(\not k_2 \not k_1 \not p_1 \not k_2) \\
&\quad + p_2.k_2 Tr(\not k_2 \not k_1 \not p_1 \not k_1)] \\
II &= 4m^2 [2p_2.k_2 Tr(\not k_1 \not p_1 \not k_1 \not k_2)] \\
II &= 2\mathcal{M}^4 32 m^2 p_2.k_2 [(k_1.p_1)(k_1.k_2) - (k_1^2)(p_1.k_2) + (k_1.k_2)(p_1.k_1)] \\
II &= 2\mathcal{M}^4 \left[64 m^2 (p_2.k_2)(k_1.p_1)(k_1.k_2)\right]
\end{aligned}
$$

- VII:

$$
\begin{aligned}
VII &= Tr[\not p_2 \not k_2 \gamma_\nu \not p \not k_1 \gamma_\sigma \not p_1 \gamma^\sigma \not k_1 \not p \gamma^\nu \not k_2] \\
VII &= 4Tr[\not p_2 \not k_2 \not p \not k_1 \not p_1 \not k_1 \not p \not k_2] \\
VII &= 4[pk_1 Tr(\not p_1 \not k_1 \not p \not k_2 \not p_2 \not k_2) \\
&\quad -pp_1 Tr(\not k_1 \not k_1 \not p \not k_2 \not p_2 \not k_2) \\
&\quad +pk_1 Tr(\not k_1 \not p_1 \not p \not k_2 \not p_2 \not k_2) \\
&\quad -ppTr(\not k_1 \not p_1 \not k_1 \not k_2 \not p_2 \not k_2) \\
&\quad +pk_2 Tr(\not k_1 \not p_1 \not k_1 \not p \not p_2 \not k_2) \\
&\quad -pp_2 Tr(\not k_1 \not p_1 \not k_1 \not p \not k_2 \not k_2) \\
&\quad +pk_2 Tr(\not k_1 \not p_1 \not k_1 \not p \not k_2 \not p_2)] \\
VII &= 128(p.k_1)(p_1.k_1)(p.k_2)(p_2.k_2) - 64(p^2)(k_1.p_1)(k_1.k_2)(p_2.k_2) \\
VII &= 2\mathcal{M}^4 \left[128((p_1-k_1).k_1)(p_1.k_1)((p_1-k_1).k_2)(p_2.k_2) - 64(p^2)(k_1.p_1)(k_1.k_2)(p_2.k_2)\right]
\end{aligned}
$$

Finally adding II and VII, and taking into account that $p^2 = m^2$, we obtain:

$$M^2_{mag} = 2\mathcal{M}^4 128(p_1.k_1)^2[(p_1.k_2)(p_2.k_2) - (k_1.k_2)(p_2.k_2)] \tag{8.10}$$

To estimate the contribution of the interaction represented in the Feymann diagram 8.5 *right*, it is enough to replace k_1 with k_2 and put $p = p_1 + k_2$. Finally, the amplitude of the interactions coming from the magnetic dipole is:

$$M^2_{mag} = 2\mathcal{M}^4 128(p_1.k_1)^2[(p_1.k_2)(p_2.k_2) - (k_1.k_2)(p_2.k_2)] + 128(p_1.k_2)^2(p_2.k_1)[(p_1.k_1) + (k_1.k_2)] \tag{8.11}$$

Interaction with the electric dipole.

Taking into account the contribution of the magnetic dipole ($\mathcal{D}$) present in the Lagrangian described in the equation 8.9, we can estimate the amplitude of the interaction represented in the Feymann diagram (See figure 8.5 *left*) as:

$$
\begin{aligned}
M^2_{ele} &= Tr[M_{ele}M^*_{ele}] \\
M^2_{ele} &= Tr[\bar{u}(p_2)\gamma_\mu\gamma_\nu\gamma_5 k_2^\mu \epsilon_2^\nu(\not{p}+m)\gamma_\rho\gamma_\sigma\gamma_5 k_1^\rho \epsilon_1^\sigma u(p_1)\bar{u}(p_1)\epsilon_1^{\sigma'} k_1^{\rho'}\gamma_5\gamma_{\sigma'}\gamma_{\rho'}(\not{p}+m)\epsilon_2^{\nu'} k_2^{\mu'}\gamma_5\gamma_{\nu'}\gamma_{\mu'}u(p_2)] \\
M^2_{ele} &= [(\not{p}_2+m_2)\not{k}_2\gamma_\nu\gamma_5(\not{p}+m)\not{k}_1\gamma_\sigma\gamma_5(\not{p}_1+m_1)\gamma_5\gamma_{\sigma'}\not{k}_1(\not{p}+m)\gamma_5\gamma_{\nu'}\not{k}_2]\eta^{\nu\nu'}\eta^{\sigma\sigma'}
\end{aligned}
$$

In a similar way to the amplitude of the interaction of the magnetic dipole, we can decompose this interaction into 8 terms. In addition, if we consider that $k^2_{1,2}=0$, the only non-null terms are:

$$
\begin{aligned}
II &= Tr[\not{p}_2\not{k}_2\gamma_\nu\gamma_5 m\not{k}_1\gamma_\sigma\gamma_5\not{p}_1\gamma_5\gamma^\sigma\not{k}_1 m\gamma_5\gamma^\nu\not{k}_2] \\
VII &= Tr[\not{p}_2\not{k}_2\gamma_\nu\gamma_5\not{p}\not{k}_1\gamma_\sigma\gamma_5\not{p}_1\gamma_5\gamma^\sigma\not{k}_1\not{p}\gamma_5\gamma^\nu\not{k}_2]
\end{aligned}
$$

It is easy to see that the calculation is identical to the magnetic dipole so that finally we get that $M^2_{ele}=M^2_{mag}$.

Then the total amplitude of the interaction between dipolar dark matter and photons is:

$$
\begin{aligned}
M^2 = &\big\{128(p_1.k_1)^2[(p_1.k_2)(p_2.k_2)-(k_1.k_2)(p_2.k_2)]+ \\
&128(p_1.k_2)^2(p_2.k_1)[(p_1.k_1)+(k_1.k_2)]\big\}*2(\mathcal{M}^4+\mathcal{D}^4) \qquad (8.12)
\end{aligned}
$$

Taking into account this amplitude we can estimate the cross-section for this interaction for different values of p_1, k_1 and θ_{p_1}. In figure 8.6 it can be seen the results obtained through the collisional method.

Taking into account this amplitude, we can estimate the energy redistribution produced by a population of dark matter particles that follow a thermal energy distribution. In figure 8.7 are shown the results obtained for energy distributions of different temperatures. While in figure 8.8 are shown the results for a population of dark matter particles with $m_{DM}=10^6 KeV$ that follows a thermal distribution with $k_BT=5*10^4 KeV$ and have different values of M and D. It is important to clarify that, due to its nature, it is not expected that

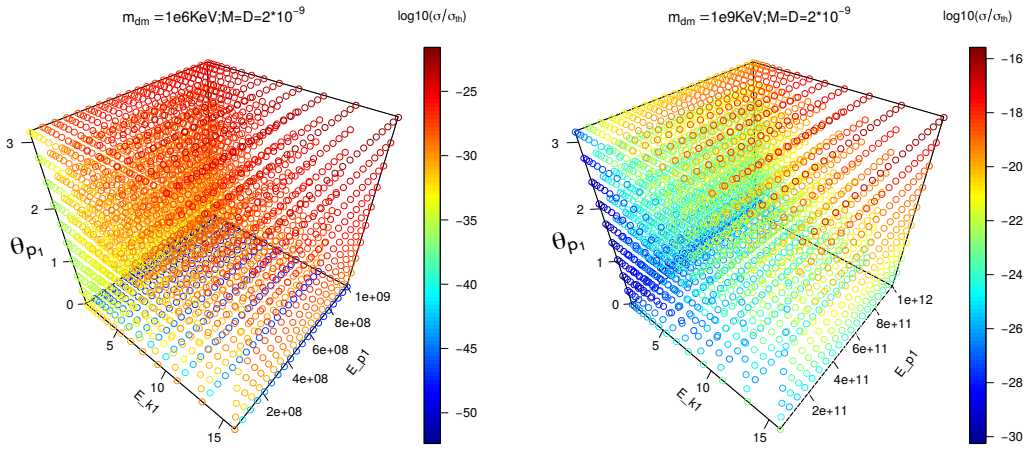

Figure 8.6: Cross-section for the interaction between photons and dipolar dark matter particles of different masses, estimated through the collisional method. *Left.* Dipolar dark matter with $M = D = 2 * 10^{-16} KeV^{-1}$ and $m_{DM} = 10^6 KeV$. *Right* Dipolar dark matter with $M = D = 2 * 10^{-16} KeV^{-1}$ and $m_{DM} = 10^9 KeV$.

dark matter particles follow a thermal energy distribution, however this analysis serves to demonstrate that just by changing the amplitude M^2, the effect is significantly less than the one produced by the electrons through Compton scattering.

8.2.2 Velocity distribution of dark matter particles.

If we assume that dark matter particles are in virial equilibrium, we can estimate their velocity distribution from the density profile. As was introduced in previous chapters, the density profile of dark matter haloes follows an NFW profile that can be described in a general way as:

$$\rho(r) = \rho_0 r_0^3 r^{-1} x^{-2} \tag{8.13}$$

where $x := r + r_0$.

Taking into account that the density distribution has spherical symmetry, we can decompose the distribution function in the phase space, in a spatial component and a component related to the energy $f_{DM}(E_p, \vec{x}) = p_{DM}(E_p)\rho(r)$. Then, if we assume that the

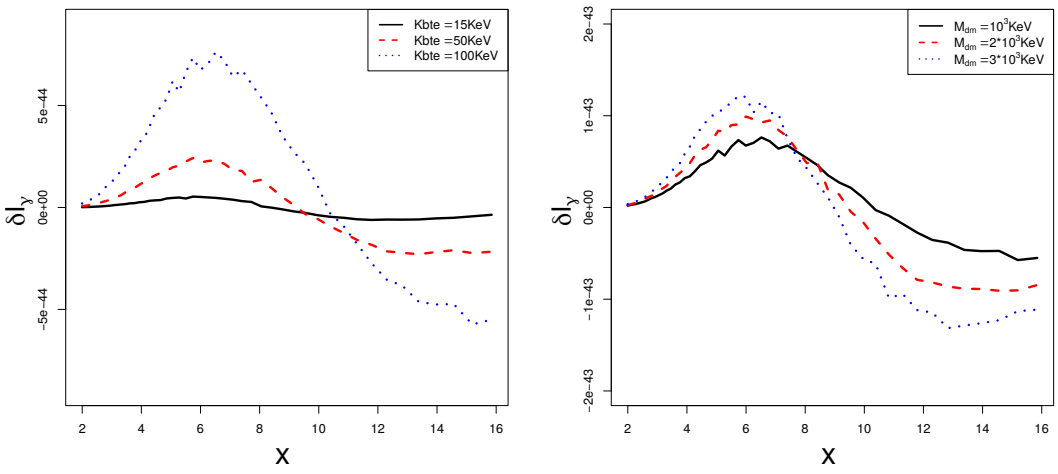

Figure 8.7: Energy redistribution function for the interaction between photons and dipolar dark matter particles with $M = D = 10^{-9} KeV^{-1}$. *Left* Particles with $m_{dm} = 1e3 KeV$ and with a thermal energy distribution with different temperatures. *Right* Particles with a thermal energy distribution with $K_b T = 100 KeV$ and different masses.

interactions between dark matter and photons do not produce significant changes in the dark matter distribution, we can solve the Vlasov equation (eq. 8.14) to find the velocity distribution of the dark matter particles.

$$\frac{d(p_{DM}\rho)}{dt} = 0 \tag{8.14}$$

It is important to clarify that the velocity distribution must fulfil the normalisation condition $4\pi \int_0^\infty v^2 p_{DM}(v) dv$ which ensures that each dark matter particle has a velocity between 0 and ∞.

Finally, the velocity distribution is:

$$P_{DM}(v) = \frac{4\pi}{\sqrt{8}\pi^2 \rho(r)} \int_\infty^{\Psi^{-1}(\mathcal{E}(r,v))} \frac{d\bar{r}}{\sqrt{\mathcal{E}(r,v) - \Psi(\bar{r})}} \mathcal{F}(\bar{r}) \tag{8.15}$$

where $\Psi(r) \coloneqq \Phi(\infty) - \Phi(r)$, $\mathcal{E}(r,v) = (1/2)v^2 + \Psi(r)$, $\Phi(r)$ is the gravitational potential of dark matter and,

$$\mathcal{F}(r) = \left[(\frac{d\Psi}{dr})^{-1} \frac{d^2\rho}{dr^2} - (\frac{d\psi}{dr})^{-2} \frac{d\rho}{dr} \frac{d^2\Psi}{dr^2} \right] \tag{8.16}$$

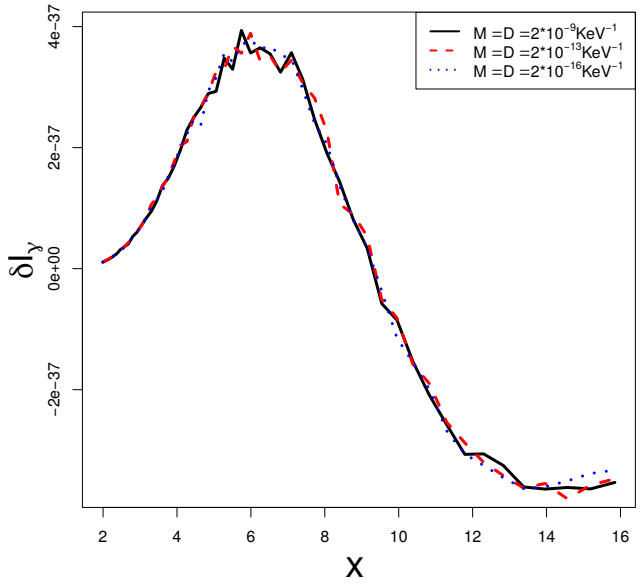

Figure 8.8: Energy redistribution function for the interaction between photons and dipolar dark matter particles with $m_{DM} = 10^6 KeV$, a thermal energy distribution with $k_B T = 5 * 10^4 KeV$ and different values of M and D. For a better comparison the value of δI_γ was multiply by 10^{16} and 10^{28} for the dark matter with $M = D = 2 * 10^{-13} KeV^{-1}$ y $M = D = 2 * 10^{-16} KeV^{-1}$ respectively.

Taking into account the NFW profile described in equation 8.13 its derivatives are:

$$\frac{d\rho}{dr} = \rho_0 r_0^3 \left[-r^{-2}x^{-2} - 2r^{-1}x^{-3}\right] \quad (8.17)$$

$$\frac{d^2\rho}{dr^2} = \rho_0 r_0^3 \left[2r^{-3}x^{-2} + 4r^{-2}x^{-3} + 6r^{-1}x^{-4}\right] \quad (8.18)$$

On the other hand, the gravitational potential (and its corresponding derivatives) generated by this density distribution is:

$$
\begin{aligned}
\Phi(r) &= -4\pi G\rho_0 r_0^2[\frac{r_0}{r}\int_0^{\chi}\frac{t}{1-t}dt + \int_0^{1-\chi}dt \\
\Phi(r) &= 4\pi G\rho_0 r_0^3\frac{Ln(r_0)-Ln(x)}{r} \\
\lim_{r\to\infty}\Phi(r) &= 0 \\
\Psi(r) &= \Phi(\infty)-\Phi(r) = 4\pi G\rho_0 r_0^3 r^{-1}\left[Ln(x)-Ln(r_0)\right] \qquad (8.19)\\
\frac{d\Psi}{dr} &= 4\pi G\rho_0 r_0^3\left[x^{-1}r^{-1}-Ln(x)r^{-2}+Ln(r_0)r^{-2}\right] \qquad (8.20)\\
\frac{d^2\Psi}{dr^2} &= -4\pi G\rho_0 r_0^3\left[x^{-2}r^{-1}+2x^{-1}r^{-2}+2r^{-3}(Ln(r_0)-Ln(x))\right] \qquad (8.21)
\end{aligned}
$$

Finally introducing the equations 8.17-8.21 in the equation 8.15, we get:

$$
\begin{aligned}
P_{DM}(v,r) &= \frac{1}{\pi^2 G\sqrt{8}\rho(r)}\int_{\infty}^{\Psi^{-1}(\mathcal{E}(r,v))}\frac{4(\bar{r}+r_0)^{-5}+(\bar{r}+r_0)^{-4}\bar{r}^{-1}(6\tilde{x}-1)}{[(\bar{r}+r_0)^{-1}+\bar{r}^{-1}\tilde{x}]} \\
&\quad \frac{d\bar{r}}{\sqrt{\mathcal{E}(r,v)-\Psi(\bar{r})}} \qquad (8.22)
\end{aligned}
$$

where $\tilde{x} \coloneqq Ln(r_0) - Ln(\bar{r}+r_0)$.

In figure 8.9 it can be seen the velocity distribution for a dark matter population in equilibrium with an NFW halo ($\rho_0 = 10^{14} M_\odot/Mpc^3$, $r_0 = 1Mpc$) at a distance $r = 0.3Mpc$ from the centre. In solid lines are shown the results obtained by the resolution of the Vlasov equation (Eddington method), while in dashed lines are shown Maxwell-Boltzmann distributions with a velocity dispersion dependent on the distance to the centre of the cluster (See equation 2.6 and 2.7 of *Petač et al. 2018* [6]).

As it can be seen, the distributions estimated with both methods are very similar. However the estimation of the distributions through the Eddington method require great computational power and present instabilities that must be studied in more detail to get reliable results. This is why to estimate the generalised Sunyaev-Zeldovich effect produced by a population of dark matter particles we will use Maxwell-Boltzmann distributions with a velocity dispersion dependent on the radio.

Taking into account this velocity distribution we can estimate the correspondent energy distribution taking into account that $E^2 = m^2(1+v^2)$ for a given particle of mass m.

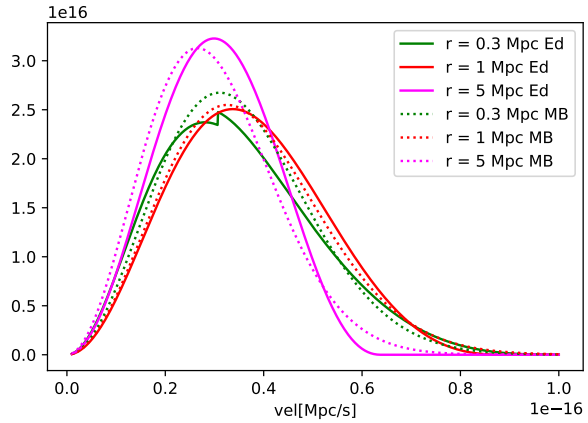

Figure 8.9: Velocity distribution of a dark matter population in equilibrium with an NFW halo ($\rho_0 = 10^{14} M_\odot / Mpc^3$, $r_0 = 1 Mpc$) at a distance of $r = 0.3 Mpc$ from the centre. In solid lines are shown the results obtained through the Eddington method, while in dashed lines are shown Maxwell-Boltzmann distributions.

Finally, taking into account this energy distribution and the amplitude of the interaction between the photons and the dipolar dark matter, we can estimate the effect produced by this interaction in the energy spectrum of the CMB photons (See figure 8.10).

As it can be seen although we are using a Maxwell-Boltzmann distribution, there are some instabilities in the final results that must be studied in more detail. However, from results shown in figures 8.6 and 8.10 we can conclude that the interaction between the CMB photons and a population of dipolar dark matter particles will not produce a significant signal in the power spectrum of the CMB.

8.3 Conclusions.

In this chapter we present the preliminary results of a study on the interaction between dipolar dark matter particles and the CMB photons. This study was done in collaboration with Dra. Celine Bœhm and Dr. Mariano Dominguez.

In the first part of this chapter we introduced the Boltzmann collisional method, which was demonstrated by *Bœhm et al. 2009* [3], that is completely equivalent to the radiative transfer method for the estimation of the Sunyaev-Zeldovich effect. However, the collisional method has the advantage of being easily generalised to study interactions between

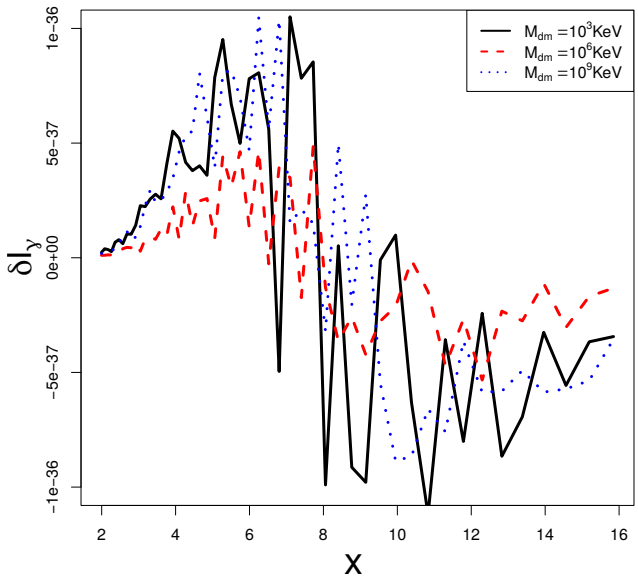

Figure 8.10: Energy redistribution function produced by the interaction of CMB photons with a population of dipolar dark matter particles that follows a Maxwell-Boltzmann distribution with a velocity dispersion dependent on the radio (See equation 2.6 and 2.7 of *Petač et al. 2018* [6]). For a better comparison the δI_γ value of the dark matter with $m_{dm} = 10^6 KeV$ and $m_{dm} = 10^9 KeV$ was multiply by $5 * 10^{11}$ and 10^4 respectively.

CMB photons and a population of any other massive particle that presents a given velocity distribution (see section 8.1).

In this direction, we developed a software to estimate the energy redistribution function of the CMB photons for a given interaction and a given energy distribution. As a check we estimate the redistribution function of the traditional Sunyaev-Zeldovich effect for electrons with different velocity distributions (see figures 8.3 and 8.2) and the Thomson cross-section (see figure 8.4).

In the second part of this work, we study the dipolar dark matter model. From the Lagrangian of this model we estimate the interactions that the CMB photons will have with dipolar dark matter particles. Using the previously described software we estimate the cross-section of this interaction if we vary the different free parameters of the dipolar dark matter model (M, D and m_{DM}). We find that the cross-section for all cases is several orders of magnitude (≈ 40) smaller than the Thomson cross-section σ_{th}.

Taking into account that the dark matter particles are clustered in haloes that follow a NFW spherical density profile, we estimate the velocity distribution that these particles would have in equilibrium with the potential of the halo. Finally, using the previously described software, we estimate the energy redistribution function that would cause the interaction of a population of dark matter particles in equilibrium. By way of comparison we also estimate the energy redistribution function if dark matter particles present other energy distributions. In all cases we found that the variations in the black-body spectrum of the CMB photons produced by these Interactions are completely negligible.

It is important to note that, depending on the dark matter model we are analysing, the assumptions used to estimate the energy redistribution of the CMB photons may not be adequate, and so, a more detailed analysis will be necessary. In addition the final results obtained by the numerical resolution of the equation 8.7 present some instabilities that must be analysed in future works. However from the results plotted in figures 8.6 y 8.10 we can concluded that the interaction between the photons of the CMB and a population of dipolar dark matter particles will not produce a significant signal in the energy spectrum of the CMB.

Bibliography

[1] M. Birkinshaw. The Sunyaev-Zel'dovich effect. *Physics Reports*, 310:97–195, March 1999. doi: 10.1016/S0370-1573(98)00080-5.

[2] C. Bœhm, J. A. Schewtschenko, R. J. Wilkinson, C. M. Baugh, and S. Pascoli. Using the Milky Way satellites to study interactions between cold dark matter and radiation. *MNRAS*, 445:L31–L35, November 2014. doi: 10.1093/mnrasl/slu115.

[3] Céline Bœhm and Julien Lavalle. Clarifying the covariant formalism for the Sunyaev-Zel'dovich effect due to relativistic nonthermal electrons. *PhRvD*, 79: 083505, April 2009. doi: 10.1103/PhysRevD.79.083505.

[4] Jonathan H. Davis and Joseph Silk. Glow in the Dark Matter: Observing Galactic Halos with Scattered Light. *Physical Review Letters*, 114:051303, February 2015. doi: 10.1103/PhysRevLett.114.051303.

[5] Julien Lavalle, Céline Bœhm, and Julien Barthès. On the Sunyaev-Zel'dovich effect from dark matter annihilation or decay in galaxy clusters. *Journal of Cosmology and Astro-Particle Physics*, 2010:005, February 2010. doi: 10.1088/1475-7516/2010/02/005.

[6] Mihael Petač, Piero Ullio, and Mauro Valli. On velocity-dependent dark matter annihilations in dwarf satellites. *Journal of Cosmology and Astro-Particle Physics*, 2018:039, December 2018. doi: 10.1088/1475-7516/2018/12/039.

[7] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, R. Barrena, J. G. Bartlett, N. Bartolo, E. Battaner, R. Battye, K. Benabed, A. Benoît, A. Benoit-Lévy, J. P. Bernard,

M. Bersanelli, P. Bielewicz, I. Bikmaev, H. Böhringer, A. Bonaldi, L. Bonavera, J. R. Bond, J. Borrill, F. R. Bouchet, M. Bucher, R. Burenin, C. Burigana, R. C. Butler, E. Calabrese, J. F. Cardoso, P. Carvalho, A. Catalano, A. Challinor, A. Chamballu, R. R. Chary, H. C. Chiang, G. Chon, P. R. Christensen, D. L. Clements, S. Colombi, L. P. L. Colombo, C. Combet, B. Comis, F. Couchot, A. Coulais, B. P. Crill, A. Curto, F. Cuttaia, H. Dahle, L. Danese, R. D. Davies, R. J. Davis, P. de Bernardis, A. de Rosa, G. de Zotti, J. Delabrouille, F. X. Désert, C. Dickinson, J. M. Diego, K. Dolag, H. Dole, S. Donzelli, O. Doré, M. Douspis, A. Ducout, X. Dupac, G. Efstathiou, P. R. M. Eisenhardt, F. Elsner, T. A. Enßlin, H. K. Eriksen, E. Falgarone, J. Fergusson, F. Feroz, A. Ferragamo, F. Finelli, O. Forni, M. Frailis, A. A. Fraisse, E. Franceschi, A. Frejsel, S. Galeotta, S. Galli, K. Ganga, R. T. Génova-Santos, M. Giard, Y. Giraud-Héraud, E. Gjerløw, J. González-Nuevo, K. M. Górski, K. J. B. Grainge, S. Gratton, A. Gregorio, A. Gruppuso, J. E. Gudmundsson, F. K. Hansen, D. Hanson, D. L. Harrison, A. Hempel, S. Henrot-Versillé, C. Hernández-Monteagudo, D. Herranz, S. R. Hildebrandt, E. Hivon, M. Hobson, W. A. Holmes, A. Hornstrup, W. Hovest, K. M. Huffenberger, G. Hurier, A. H. Jaffe, T. R. Jaffe, T. Jin, W. C. Jones, M. Juvela, E. Keihänen, R. Keskitalo, I. Khamitov, T. S. Kisner, R. Kneissl, J. Knoche, M. Kunz, H. Kurki-Suonio, G. Lagache, J. M. Lamarre, A. Lasenby, M. Lattanzi, C. R. Lawrence, R. Leonardi, J. Lesgourgues, F. Levrier, M. Liguori, P. B. Lilje, M. Linden-Vørnle, M. López-Caniego, P. M. Lubin, J. F. Macías-Pérez, G. Maggio, D. Maino, D. S. Y. Mak, N. Mandolesi, A. Mangilli, P. G. Martin, E. Martínez-González., S. Masi, S. Matarrese, P. Mazzotta, P. McGehee, S. Mei, A. Melchiorri, J. B. Melin, L. Mendes, A. Mennella, M. Migliaccio, S. Mitra, M. A. Miville-Deschênes, A. Moneti, L. Montier, G. Morgante, D. Mortlock, A. Moss, D. Munshi, J. A. Murphy, P. Naselsky, A. Nastasi, F. Nati, P. Natoli, C. B. Netterfield, H. U. Nørgaard-Nielsen, F. Noviello, D. Novikov, I. Novikov, M. Olamaie, C. A. Oxborrow, F. Paci, L. Pagano, F. Pajot, D. Paoletti, F. Pasian, G. Patanchon, T. J. Pearson, O. Perdereau, L. Perotto, Y. C. Perrott, F. Perrotta, V. Pettorino, F. Piacentini, M. Piat, E. Pierpaoli, D. Pietrobon, S. Plaszczynski, E. Pointecouteau, G. Polenta, G. W. Pratt, G. Prézeau, S. Prunet, J. L. Puget, J. P. Rachen, W. T. Reach, R. Rebolo, M. Reinecke, M. Remazeilles, C. Renault, A. Renzi, I. Ristorcelli, G. Rocha, C. Rosset, M. Rossetti, G. Roudier, E. Rozo, J. A. Rubiño-Martín,

C. Rumsey, B. Rusholme, E. S. Rykoff, M. Sandri, D. Santos, R. D. E. Saunders, M. Savelainen, G. Savini, M. P. Schammel, D. Scott, M. D. Seiffert, E. P. S. Shellard, T. W. Shimwell, L. D. Spencer, S. A. Stanford, D. Stern, V. Stolyarov, R. Stompor, A. Streblyanska, R. Sudiwala, R. Sunyaev, D. Sutton, A. S. Suur-Uski, J. F. Sygnet, J. A. Tauber, L. Terenzi, L. Toffolatti, M. Tomasi, D. Tramonte, M. Tristram, M. Tucci, J. Tuovinen, G. Umana, L. Valenziano, J. Valiviita, B. Van Tent, P. Vielva, F. Villa, L. A. Wade, B. D. Wandelt, I. K. Wehus, S. D. M. White, E. L. Wright, D. Yvon, A. Zacchei, and A. Zonca. Planck 2015 results. XXVII. The second Planck catalogue of Sunyaev- Zeldovich sources. *A&A*, 594:A27, September 2016. doi: 10.1051/0004-6361/201525823.

[8] J. A. Schewtschenko, R. J. Wilkinson, C. M. Baugh, C. Bœhm, and S. Pascoli. Dark matter-radiation interactions: the impact on dark matter haloes. *MNRAS*, 449:3587–3596, June 2015. doi: 10.1093/mnras/stv431.

[9] J. A. Schewtschenko, C. M. Baugh, R. J. Wilkinson, C. Boehm, S. Pascoli, and T. Sawala. Dark matter-radiation interactions: the structure of Milky Way satellite galaxies,. *MNRAS*, 461:2282–2287, September 2016. doi: 10.1093/mnras/stw1078.

[10] R. A. Sunyaev and Ia. B. Zeldovich. Microwave background radiation as a probe of the contemporary structure and history of the universe. *Annual Review of Astronomy and Astrophysics*, 18:537–560, January 1980. doi: 10.1146/annurev.aa.18.090180.002541.

[11] R. A. Sunyaev and Ya. B. Zeldovich. The Observations of Relic Radiation as a Test of the Nature of X-Ray Radiation from the Clusters of Galaxies. *Comments on Astrophysics and Space Physics*, 4:173, November 1972.

[12] Ryan J. Wilkinson, Céline Bœhm, and Julien Lesgourgues. Constraining dark matter-neutrino interactions using the CMB and large- scale structure. *Journal of Cosmology and Astro-Particle Physics*, 2014:011, May 2014. doi: 10.1088/1475-7516/2014/05/011.

[13] Ryan J. Wilkinson, Julien Lesgourgues, and C. Bœhm. Using the CMB angular power spectrum to study Dark Matter-photon interactions. *Journal of Cosmology*

and Astro-Particle Physics, 2014:026, April 2014. doi: 10.1088/1475-7516/2014/04/026.