

UNIVERSITÀ DEGLI STUDI DI MILANO

PHD SCHOOL IN PHYSICS, ASTROPHYSICS AND APPLIED PHYSICS
DEPARTMENT OF PHYSICS

CYCLE XXXII

Modeling massive galaxy clusters via strong gravitational lensing – the case for Abell 2163 –

Disciplinary Scientific Sector FIS/05

Director of the School: Professor Matteo Paris

Supervisor of the Thesis: Professor Marco Lombardi

Co-Supervisor of the Thesis: Professor Claudio Grillo

PhD Thesis of
Umberto Rescigno

Academic Year 2019-2020

Commission of the final examination:

External Member:
Professor Lauro Moscardini

External Member:
Professor Sherry Suyu

Internal Member:
Professor Marco Lombardi

Final examination:

Date 2019/12/20

Università degli Studi di Milano, Dipartimento di Fisica, Milano, Italy.

Contents

List of Figures **v**

List of Tables **xi**

Introduction **xiii**
Motivation xiii
Thesis objectives and relevance xiv
Organizational note, thesis overview and conventions xvi

1 Gravitational Lensing **1**
1.1 Theoretical basics 1
1.1.1 Light deflection and the *factor 2* 1
1.1.2 Thin lens approximation and the ray-tracing equation 3
1.1.3 Convergence and gravitational lensing potential 4
1.1.4 Magnification and distortion 5
1.2 Lens models of an extended mass distribution 9
1.2.1 Axisymmetric lenses 9
1.2.2 A benchmark example: the isothermal sphere 10
1.2.3 Non-circularly symmetric lenses: the elliptical isothermal model 11
1.3 Lensing by galaxy clusters: a cosmological perspective 13
1.3.1 Cosmological parameter estimate using weak lensing 14
1.3.2 Dark matter profiles of clusters using strong lensing 15

2 Galaxy clusters **17**
2.1 Generality 17
2.1.1 Galaxy cluster formation 17
2.1.2 Definitions 18
2.1.3 Properties 20
2.1.4 Cluster classification(s) 21
2.2 Dark matter 24
2.2.1 Characterization and distribution 24
2.2.2 Comparing simulations and observations 26
2.3 Intracluster medium 28
2.3.1 Discovery and applications 28
2.3.2 ICM features 29

2.3.3 Cluster mass estimate from X-ray data 32
2.4 Galaxies 33
2.4.1 Galactic content in cluster cores 33
2.4.2 Overall galactic component 33
2.5 Minor components 35

3 Our strong gravitational lens: Abell 2163 39
3.1 Multi-band cluster overview 39
3.2 Merging history 42

4 Observations and data 47
4.1 Large galaxy cluster surveys with the *HST* 47
4.1.1 CLASH and CLASH-VLT 47
4.1.2 HFF 49
4.1.3 RELICS 49
4.2 Imaging 53
4.2.1 *HST* 53
4.2.2 Abell 2163 imaging data 56
4.3 Spectroscopy 60
4.3.1 MUSE 60
4.3.2 Abell 2163 spectroscopic data 64

5 The redshift catalogs 69
5.1 Extraction of source spectra and redshift measurements 69
5.2 Selection of cluster members 72
5.2.1 Spectroscopic sample 72
5.2.2 Photometric sample 75
5.3 Multiple images 77

6 Strong lensing modeling of Abell 2163 81
6.1 Gravitational lens modeling software 81
6.2 `Lenstool` 82
6.3 Method description 84
6.3.1 Cluster mass decomposition 84
6.3.2 Run setup 87

7 Results and discussion 89
7.1 Computational runs 89
7.2 Comparing the fiducial model with the literature 95
7.3 Conclusions and future directions 104

A Catalog of foreground and background sources 109

B Notes on `Lenstool` 111

Bibliography 113

Acknowledgments 120

List of Figures

1.1 Representation of the geometrical configuration of a possible gravitational lensing event (see the text for the description); the figure is a reimagining of that taken from A. B. Congdon (2018). 3

1.2 Gravitational arcs in 6 images taken from the *Hubble Space Telescope*'s Advanced Camera for Surveys (credits: NASA, ESA, A. Bolton, Harvard-Smithsonian CfA and the SLACS Team). The reddish-white objects in the center of each image are massive galaxies and the gravitational arcs are the blue curved patterns around them. Notice the high degree of distortion, especially for the arcs in the column at the right. 6

1.3 Color-composite image of the core of the galaxy cluster MACS J0416.12403 from *HST* data; the yellow arrows indicate the position of some giant arcs. In contrast to the arcs in Fig. 1.2, they are extremely elongated across very large scales, reflecting the substantial mass difference between a cluster and a galaxy and the consequent grater *lensing power* of former compared to the latter. The characteristics of galaxy clusters producing giant arcs are described in Narayan & Bartelmann (1995). 8

1.4 Caustics, critical lines and the formation of multiple images for an elliptical lens. The right and left panels show, respectively, the lens and the source planes; the little symbols on the circle of each source allow one to understand the parity of the corresponding multiple images. 12

2.1 The Coma cluster (at a distance of $\sim$99 Mpc) is a classical example of a regular cluster, with practically no spiral galaxies and a core dominated by a pair of super-giant elliptical galaxies: NGC 4874 and NGC 4889 (picture taken at the *Mount Lemmon SkyCenter*, using the 0.8 m Schulman Telescope). 23

2.2 The Virgo cluster (at a distance of $\sim$16.5 Mpc), with its irregular galaxy distribution, is the largest cluster and the nearest to our Local Group; its core hosts at least three sub-clusters, centered around the galaxies M87, M86, and M49 (picture by Dr. F. Neyer, Antares Observatory). 24

2.3 Tuning fork classification of galaxy clusters by Rood & Sastry (1971). Image taken from Schneider (2015). 25

2.4 Three-panel image of Abell 2125. Left: a *Chandra*'s X-ray view of the entire galaxy cluster, where red, green, and blue represent, respectively, low, medium, and high-energy X-ray emissions. Middle: Abell 2125 core, where the X-ray emitting gas distribution is in green and blue, while the red color refers to the optical emission. Right: Galaxy C153 in visible light (red) from *HST* and X-ray emission (green and blue) observed by *Chandra* (Credit: X-ray: NASA/CXC/UMass; Optical: NASA/STScI and NOAO/Kitt Peak). 30

2.5 The spiral galaxy ESO 137-001, in the Norma cluster located at a distance of about 68 Mpc. This image combines NASA/ESA *HST* observations with data from the *Chandra* X-ray Observatory (in blue). The gas trail of this spiral (also known as the *jellyfish galaxy*) is a remarkable example of a process known as *ram pressure stripping*: high-velocity galaxies traveling through the hot ICM lose their gas, preserving their stellar content, which instead remains intact due to the binding force of their higher gravity. 31

2.6 Image of the galaxy cluster MACS J0416.1–2403 (observed by the HFF team), in which the ICL is highlighted in blue. Credit: ESA/Hubble, NASA, *HST* Frontier Fields team (STScI), and M. Montes and I. Trujillo. 36

3.1 Spatial distribution of the A2163 sub-structures, as observed in the three bands reported in the upper-right legend. The optical mass clumps (A1, A2 and B) and the radio relic (R) are described in the text. 40

3.2 Optical image of A2163, taken by *HST* ACS and WFC3 cameras, partially for the RELICS observing program. The image FoV is 6.26×3.10 arcminutes and the North is $56^{\circ}.1$ left of horizontal. The list of the 7 filters used and further information can be found on the *HST* website, at the link `https://www.spacetelescope.org/images/potw1750a/`. 41

3.3 Radio maps of A2163 at 20 cm from Thölken et al. (2018), where the letter R marks the position of the radio relic, as identified by Feretti et al. (2001), and the contour levels are 0.2, 0.3, 0.5, 0.7, 1, 2, 3, 5, 7 mJy/beam. The right part of the figure shows the extended emission after the removal of compact sources which instead are present in the image on the left. 42

3.4 Formation of shock fronts in a merger event from Thölken et al. (2018). In an early stage of the merger (left), shock fronts develop after the shock heating of the gas from the two clusters; successively, the cluster core merger can create further shock fronts, while the initial shock waves travel outwards. 43

3.5 Temperature profiles of A2163 core clusters from Thölken et al. (2018). Top panel: projected profiles of the sub-clusters at NE (red diamonds) and SW (blue diamonds); the horizontal red and blue dotted lines are from the best-fit models, while vertical dashed lines indicate the discontinuity positions (S stands for "Schock"). Bottom panel: deprojected temperature profile of A2163-A (solid line) around the shock position; the gray shaded area is the 68% confidence interval from 1000 Markov Chain realizations and the dashed line shows the median of all of them. 44

3.6 Regions of the shock fronts in A2163 (blue), whose widths correspond to the position uncertainties. The black diamond and the magenta ellipse are, respectively, the X-ray peak and the approximate position of the cool core (remnants of the merging constituents). Green contour levels (the same as in the right panel of Fig. 3.3) show the extended radio emission. 45

4.1 Variation of the systematic error of the magnification as a function of the number of spectroscopically confirmed multiple images. The data refer to several gravitational lensing models (shown in different colors) and the statistical errors are not included (Credits: T. L. Johnson, C. Cerny, American Astronomical Society, AAS Meeting 229, 01/2017). 51

4.2 The figure shows the cluster mass from the *Planck* SZ catalog (*Planck* Collaboration: Adam et al. 2016) as a function of the redshift, for different observations and astronomical programs (taken from the RELICS project website). The RELICS clusters (squared red boxes) are very massive; most of the boxes are in fact concentrated in the high mass area. In particular, A2163, the most massive and at $z \sim 0.2$, is at the top-left corner. 52

4.3 Sequence of images taken with the *HST* (columns 1 and 3) and the Palomar observatory telescope (columns 2 and 4). The first and third couple illustrate the spherical aberration effects on the same *HST* pointing, before (upper panels) and after (lower panels) the telescope mirror fixing (Credits: New Eyes and Messier 100, ESA *HST*/NASA). The second and fourth couple of images are described in the text and show, respectively, the globular cluster M13 and the Cat's-Eye Nebula (Credits: Caltech/JPL/Cambridge/Palomar Observatory). 54

4.4 Representation of the *HST* and its principal instruments, labeled. On the right part, the vertical numerical scale, in nanometers, indicates which wavelengths each component studies. All these instruments are described in the text, except for the Near Infrared Camera and Multi-Object Spectrometer (NICMOS), since it is currently inoperative: it was only active during about 8 years (1997-2008) for IR imaging and spectroscopy. The figure is a modified and composite version of different NASA images (Credits: NASA's Goddard Space Flight Center). 55

4.5 Response curves of the ACS and WFC3/IR filters used to observe A2163, as a function of the wavelength. A vertical offset for F105W and F140W was applied for clarity and the black dots indicate the filter effective "pivot" wavelengths (Tokunaga & Vacca 2005), listed in Table 4.2. 58

4.6 Color composite image of A2163. This is a combination of ACS and WFC3 data, which were collected thanks to two different *HST* Proposals, respectively, a 32-orbit Treasury program (ID GO 12253, cycle 18, P.I. D. Clowe) and a 190-orbit program (ID GO 14096, cycle 23, P.I. D. Coe). The yellow region marks the FoV of one of the MUSE pointing, from which our spectroscopic data were taken (see Sect. 4.3.2). 59

4.7 Representation of a MUSE data-cube. The figure shows the different cube components and the information that can be extracted from them: the side of the cube at the origin wavelength λ_0 and one cube slice at a certain $\lambda \neq \lambda_0$ give, respectively, the source image at all or at a *single* wavelength. Navigating the data-cube, from a pixel at $\lambda = \lambda_0$ until the end of the wavelength axis, one can also obtain the pixel spectrum. 60

4.8 Examples of IFS capabilities from Bacon et al. (2015). Upper panel: [O II] and $Ly\alpha$ sources, blended in *HST* images (filters F606W and F814W), were separated looking at two different data-cube slices, corresponding to the wavelength interval of each of those two emission lines. Lower panel: a $(Ly\alpha)$ source discovered with IFS observations, navigating through the data-cube, otherwise invisible in *HST* data or even in the white-light image of the cube. 61

4.9 *HST* image cutouts, with the circular aperture (green circle) we use to extract the image spectra, and the relevant spectral details for the 7 secure multiple images in A2163 core. The panels of each column refer to the same family and (from the top to the bottom) to the images a, b and c, respectively. QFs for all multiple images are reported in Table 5.2 of Sect. 5.3. The spectral line doublets (in gray) are C III] ($\lambda\lambda$ 1907, 1909 from a semi-forbidden transition) and [O II] ($\lambda\lambda$ 3727, 3729), respectively, for the first two columns and for the last one. 66

5.1 Left panel: examples of *HST* catalog sources (in green) which are wrongly segmented. Right panel: some sources (in yellow) discarded because of the small number of readable pixels in the data-cube. The red line is the west edge of the MUSE FoV (its inner region is on the left) in an ideal, conservative representation: the real edge is more jagged, hence including (partially) the pixels within the yellow apertures. 70

5.2 Spectrum of A2163-A1 BCG (in black) as visualized in the EZ spectral viewer, after the fit procedure. The vertical axis corresponds to the counts, while the horizontal one to the wavelength, in Å. The orange and the blue curves are, respectively, the galaxy template and the continuum adopted during the fit step; spectral line wavelengths, shifted at the galaxy redshift, are marked in green. 71

5.3 Color composite image of Abell 2163 from the *HST* data. The overlaid MUSE pointing (yellow box) is about one arc minute across and is centered on the yellow plus sign position. Circles mark the location of the 29 spectroscopically confirmed galaxy members and they are colored according to the galaxy velocities (relative to the cluster mean redshift). The range of the values is shown on the bar on the top. 74

5.4 Magnitude distribution of the cluster members, having $m_{\mathrm{F814W}} < 24$ and a member probability $P > 95\%$; the whole sample is shown as a gray histogram, while spectroscopically confirmed galaxies (29) and photometrically selected members (82) are in red and green, respectively. 76

5.5 Comparison between the *HST* image (right panel) and the MUSE FoV (left panel) region around the galaxy member generating the multiple images of family F4. The new image (4d) position in marked by the orange circle, while the two arcs of F4 are within the red curves. 78

5.6 Color composite image illustrating the distribution of the multiple images included in our fiducial model. 79

7.1 Variation of the number of iterations over the total one, as a function of the run time for the burn-in (left panel) and the sampling (right panel) phases. 90

7.2 Evolution of the Evidence in function of the parameter $Rate$, regulating the convergence speed (from Jullo et al. 2007). 91

7.3 Multiple images predicted by RUN 1. In both panels, solid and dashed curves mark, respectively, the positions of the observed and model-predicted images of family F1 (in magenta) and F3 (in green). The left panel figure is a zoom on the total projected mass map of A2163-A1, derived from RUN 1 (with denser regions in bright colors). The two model-predicted images which are very close to the center of the BCG are not visible in the *HST* data, while the remaining ones (1d and 3d) have visible counterparts. The latter are reported in the two figures of the right panel. 93

7.4 Left panel: collage of *HST* cutouts illustrating the four images of F1 and F2. Right panel: distribution of the model-predicted F2 redshift values, z_{F2}, in RUN 3. The solid red line corresponds to the median value of z_{F2}, while the blue one marks the measured value of F1 redshift, z_{F1}. 94

7.5 Posterior distributions of the parameter values of the cluster-scale mass component. Blue contours correspond to the 1, 2 and 3 σ confidence levels of a Gaussian distribution, while vertical blue dashed lines in the histograms are the 16th, 50th and 84th percentiles. Corner plots were created using the `corner.py` module by Foreman-Mackey (2016). 96

7.6 As in Fig. 7.5. Left panel: BCG parameters, which are linked to the small-scale sub-halo mass component, through the member galaxy scaling relations (6.6)-(6.7). Right panel: free parameters of the galaxy which contributes to the formation of F4 images. 97

7.7 Cumulative projected mass profiles from our fiducial model; black, blue and orange curves represent the total, the smooth and the clumpy components, respectively. Solid and dashed lines trace the median and 16th - 84th percentiles, while the light ones complete the sub-sample extracted from the final MCMC chains. 98

7.8 Projected total mass profiles of a sample of clusters described in Caminha et al. (2019) and references therein (colored curves); the blue circles mark some points of the projected total mass profiles of A2163 as derived by our fiducial model. All the curves are rescaled by the quantities M_{200} and R_{200} from independent weak lensing measurements. 100

7.9 Composite image of A2163-A1, with overlays of the total surface mass density distribution. Green and red contours refer, respectively, to this work and C18. Contour line values are $[0.75\,, 1.00\,, 1.50\,, 2.00\,, 2.50\,, 3.50] \times 10^9\,\mathrm{M}_\odot\,\mathrm{kpc}^{-2}$. 101

7.10 Critical lines of the fiducial model, for sources at redshifts 1.164 (gold), 2.389 (orange) and 2.723 (red), superposed on a color composite image of A2163-A1. 102

7.11 Critical lines (in white) for some galaxies (in yellow), having redshifts of about $z \sim 0.665$. This value is compatible with that of the arc-shape galaxy (4147).. 103

7.12 MUSE pointings in the core of A2163 (white boxes), in the north-est and south-west regions, respectively, on the left and right part of the figure; the length of each FoV side is $\sim 1'$. 106

7.13 Spatial configuration of the three mass components of two merging sub-clumps (A and B) in a SIDM scenario: the stellar distribution (in green), usually centered at the BCG position, and those of the SIDM (in violet) and of the gas (in red) halos. The two arrows point to the motion direction of the sub-clumps after the collision. 107

List of Tables

2.1 Abell's richness classes. 22
2.2 Comparison between clusters in different classification schemes. 34

4.1 List of some general characteristics of the *HST*. The different kinds of information are organized in groups of rows, separated by an horizontal line; they refer, respectively, to the mission, size, orbit, power needs/storage and mirrors. 57
4.2 Information on A2163 observations: *HST* camera used, its filter and "pivot" wavelength (Tokunaga & Vacca 2005), with exposure times related to the two days in which data were collected. The project ID is also listed. ACS camera data were taken from C18. 58
4.3 VLT Observations of the cluster A2163, using MUSE. Data were obtained in the MUSE commissioning phase, under the program 60.A-9100(C) (PIs: MUSE TEAM, Commissioning). 65

5.1 Catalog of the galaxy cluster members, with *HST* IDs, celestial coordinates and spectroscopic redshifts with QFs. 73
5.2 Catalog of multiple images, with ID, celestial coordinates and spectroscopic redshifts with the related QFs; we also report the ID of the reference work in the literature (C18). From the top to the bottom, the horizontal lines separate, respectively, the multiple images of families F1, F2, F3, and F4. 77

7.1 Summary of run characteristics and results. For each run, we report, respectively, the ID, the list of images considered, the four quantities to calculate the model *dof* of Eq. (6.5) and two relevant statistical quantities: the initial value of the minimum χ^2 (the final one equals the number of *dof*, column 8) and the rms error, as defined in Sect. 6.3.2 (column 9). 92
7.2 Median values and confidence level (CL) uncertainties of the parameters for the lens model RUN 5 of A2163. Centers are relative to the BCG (R.A. = 243.9539405° and Decl. = −6.1448406°). 95

Introduction

Motivation

The composition of the present Universe has proven to be much more complex than was thought even just a few decades ago. High-precision observations in the last years (Ade et al. 2016a) suggest that ordinary matter, namely the entire set of fundamental constituents of the *Standard Model of Particle Physics* and present in different forms (plasma, atoms and molecules, which make the stars and galaxies we see), contributes just a few percentages (5%) of the present cosmic content. The remaining 95% is made up of two "dark" and still unknown components: 27% of yet undetected, weakly- or non-interacting material, called Dark Matter (hereafter DM), and 68% of a physically unexplained form of energy, called dark energy, possibly responsible for accelerating the expansion of the present Universe (Peebles & Ratra 2003). The former cannot be directly observed in astrophysical systems since it does not absorb or emit in the electromagnetic spectrum.

The currently accepted cosmological model assumes DM as a collisionless and cold particle (CDM) and explains dark energy and the increasing expansion rate of the present Universe by introducing in the gravitational equations of *General Relativity* a constant term (the cosmological constant, Λ), identified as the vacuum energy of the contemporary *Quantum Field Theory*: for these reasons it is indicated with the acronym ΛCDM, stressing the chosen assumptions on the model dark components. Despite the enigmatic nature of the two essential, dark constituents of the Universe mentioned above, the information extracted from the *Planck* satellite data provides excellent confirmation of the standard cosmological model, with unprecedented accuracy (Ade et al. 2016a). Along the lines of the most recent observational evidence, the standard paradigm has been remarkably effective in providing physical interpretations for a wide range of phenomena and has provided a compelling backbone to galaxy formation theory, a field that is becoming increasingly successful in reproducing the detailed properties of galaxies, including their counts, clustering, colors, morphology, and evolution over time. In spite of the celebrated success of today's standard predictions on cosmological scales, a certain number of (tantalising) issues emerges on the scale of collapsed structures ($\leq$ 1 Mpc). In the following outline, we report some relevant discrepancies with available observations, from smaller to larger astrophysical scales.

- **The core-cusp problem** ($\lesssim$ 1 kpc, dwarf galaxy scale; de Blok 2009). Standard simulations imply a high density (a "cusp", namely a divergent density) at the center of DM-dominated systems, in contrast to some dynamical analyses of observed dwarf and low surface brightness galaxies, which seem to favor an approximately constant density (a "core") for their DM halos.

- **The missing satellites problem** ($\sim$ 30 kpc, galaxy scale; Klypin et al. 1999; Moore et al. 1999). It deals with satellite galaxies (i.e., dwarf systems orbiting a common-size galaxy) and arises from their different numbers in the expectations of cosmological simulations and the observed counts: the former exhibit more satellites than the latter.

- **The overabundance of massive cluster galaxies** ($\lesssim$ 200 kpc, galaxy cluster inner core; Grillo et al. 2015; Munari et al. 2016). The number of observed sub-halos, associated with cluster galaxies, is greater than that in simulations and this mismatch is more severe in the cluster innermost regions. More specifically, the observed number of low-mass cluster members is in good agreement with the predicted number from the N-body simulations, whereas simulated clusters have a statistically significant deficiency in more massive substructures.

- **The inner density profile of galaxy clusters** ($\sim$ 350 kpc, galaxy cluster core radius; Bartelmann & Meneghetti 2004). The precise value of the inner slope of cluster mass density profiles is intensely debated and, to date, significant progress has been made towards the measurement of this value, albeit with intriguing and somewhat controversial results by different groups for the very same clusters. This parameter contains valuable information about the nature of DM: cosmological models based on different DM particle candidates predict different values.

The above-introduced deviations from cosmological standard predictions have stimulated several specific questions, mostly converging to an interesting debate on whether the physics of ordinary matter can alter the DM distribution, whether this tension is simply the result of selection effects and/or projection biases, and whether the characteristics of DM assumed in standard cosmology are incorrect. Currently, the prospects for making a giant step forward in our understanding of the nature of DM and in obtaining precise maps of its distribution are very promising. Within this general framework, the main motivation of the present research is contributing to disentangle the DM distribution in an unbiased sample of massive galaxy clusters in order to quantify the amount of DM in their cores. The juxtaposition of the results of the studies proposed here for strong lens galaxy clusters and of the outcomes of recent cosmological simulations has the potential, among other possibilities, for defining better the dynamics of complex clusters and for distinguishing among different DM characterizations (standard massive candidates, self-interacting particles, etc.): this in turn can solve or alleviate the most challenging issues of the current ΛCDM paradigm.

Thesis objectives and relevance

Addressing the fundamental research questions arising from the standard cosmological model can only be made by obtaining homogeneous, high quality data on a sizable and unbiased sample of astrophysical objects. Clusters of galaxies, the largest and the most massive collapsed cosmic structures, are at the boundary of the critical scale at which cosmological issues manifest themselves, so that, by virtue of their features, they offer unique tests of any viable cosmology and structure formation scenario, suggesting possible solutions to ΛCDM problems and helping in discriminating among the most common possibilities. However, one fundamental, preliminary step in this direction is to exploit independent/complementary mass diagnostics of galaxy clusters and, most importantly, these estimates must be necessarily as accurate as possible: this is where

gravitational lensing can help. In fact, this discipline could provide the most precise total mass estimates, since lensing uncertainties do not depend on common assumptions about galaxy cluster composition or dynamical state (on which other mass diagnostics are based). This is why applying GL techniques to galaxy clusters is crucial.

Our understanding of galaxy clusters has significantly improved in recent years, thanks to high-quality datasets from multi-band surveys with the *Hubble Space Telescope* (*HST*) and spectroscopic follow-up programs with ground-based telescopes. In particular, the VIsible Multi-Object Spectrograph (VIMOS) and the Multi Unit Spectroscopic Explorer (MUSE), both mounted on the Very Large Telescope (VLT), have been successfully used to study the mass distribution of galaxy clusters. They have also helped to characterize some of the most distant galaxies known to date, as in the Reionization Lensing Cluster Survey (RELICS, Coe et al. 2019), whose main goal is to discover and study hundreds of galaxies at $z > 6$ to better understand the epoch of reionization (Salmon et al. 2017, 2018). In this program, 46 fields were selected among the most massive *Planck* clusters ($M_{500} > 4 \times 10^{14}\,\mathrm{M}_{\odot}$), showing exceptional strong-lensing features. Two other relevant examples are the Cluster Lensing And Supernova survey with *Hubble* (Postman et al. 2012), a 524-orbit *HST* Multi-Cycle Treasury program targeting 25 high-mass clusters, and the *Hubble* Frontier Fields program (Lotz et al. 2017), an initiative aimed to obtain the deepest *HST* and *Spitzer Space Telescope* observations of six clusters and their lensed galaxies. In this thesis, we use *HST* data products from the RELICS survey, supplemented with MUSE spectroscopy, to derive the total mass distribution in the core of the cluster Abell 2163 (hereafter A2163) via a strong gravitational lensing analysis.

A2163 is a challenging system and a goldmine for contemporary astrophysics: it is one of the richest Abell clusters, with remarkable features, a complex structure and a variety of interacting sub-systems manifesting their presence and activity all over the electromagnetic spectrum. As pointed out in Soucail (2012), the relations among its mass components are not yet well explored, thus making hard to define a clear picture of the cluster physical state. In particular, the relations between its different components have never been confronted via detailed strong lensing studies based on spectroscopic measurements. With this thesis we want to bridge this gap, providing the redshift measurements of all sources in the A2163 core within our MUSE field of view (FoV): the multiple images previously detected, the new ones we have identified, the cluster member and the background and foreground objects. Moreover, focusing on this particular massive galaxy cluster allows us to set the stage for interesting prospects: our final results can improve a number of scientific programs with different goals, such as the RELICS one, and can contribute significantly to different lines of research, e.g., the definition of the A2163 merging history and an initial test on a possible self-interaction of DM.

These investigations are also addressed throughout this thesis and are related to its main objective as described in what follows. To reconstruct a realistic merging scenario for A2163, it is necessary to use all information at our disposal and combine all mass measurements. In this cluster, X-ray, optical and weak lensing studies are all present in the literature, with mass estimates somehow discordant (Okabe et al. 2011; Bourdin et al. 2011), while strong lensing mass models rely on photometric redshifts only (Cerny et al. 2018). In this regard, the present thesis is intended to complete the results of the other mass diagnostics, providing the most precise strong lensing mass estimate of the core of A2163. Our work is not aimed at confirming any merging scenario of this galaxy cluster, although many relevant considerations in this respect can be inferred based on

our results, as explained in Sec. 7.3. There, we also focus on a further intriguing research opportunity, which can be investigated starting from a high-resolution cluster mass map, like the one provided here: for massive clusters, having sub-clumps where the first merging passage has already occurred, it should be possible to test the presence of the DM self-interaction (Spergel & Steinhardt 2000) by analyzing the displacements of the three mass components of each cluster merging sub-structure (Markevitch et al. 2004; Harvey et al. 2015): the DM halo, the galaxies and the hot gas.

Organizational note, thesis overview and conventions

The present thesis consists of seven Chapters. They begin with an abstract which outlines the subjects of each section. An overview of its contents are detailed below. For a summary of the thesis and a schematic description of our main results, the reader can refer to Sec. 7.3.

Chapter 1. Gravitational Lensing. We pave the way for the interpretation of the proposed mass model of A2163 by introducing the basics of Gravitational Lensing theory. In particular, we present the axisymmetric profiles routinely used in lens modeling of extended mass distributions, with emphasis on the elliptical isothermal model we adopt.

Chapter 2. Galaxy clusters. The definition, the classification(s) and the general properties of galaxy clusters are treated. We discuss the reliability of the most common assumptions at the base of cluster mass diagnostics and describe thoroughly the mass components of these systems, mainly dark matter, intracluster medium and galaxies.

Chapter 3. Our strong gravitational lens: Abell 2163. We focus on this massive cluster overall, on its multi-wavelength characteristics (in X-ray, optical and radio bands) and on its sub-components. Then, we trace its possible merging history, explaining the difficulties why it is not conclusively well defined.

Chapter 4. Observations and data. We address the large galaxy cluster surveys employing the two instruments which supply the imaging and spectroscopic data of A2163, respectively, the *HST* and the MUSE, mounted on the Very Large Telescope, in Chile. The observations of the cluster, the data collection and reduction are also illustrated, together with the essential instrument equipment.

Chapter 5. The redshift catalogs. We sketch our procedure to measure the redshifts of the sources and to assign them quality flags, as well as the selection criteria for both cluster members and multiple images. All the sources, including the background and foreground objects, are then organized in accurate catalogs, collecting all the reliable redshift measurements.

Chapter 6. Strong lensing modeling of Abell 2163. We give an overview of the kinds of software for gravitational lens modeling, concentrating on that adopted in this thesis, `Lenstool`. The method to model A2163 via strong gravitational lensing and then to derive the total, projected mass profile of its core is detailed. In particular, we explain how to differentiate the mass components of this galaxy cluster and how to implement them in the phase of the run setup.

Chapter 7. Results and discussion. We explore a number of computational runs to evaluate our theoretical models of the cluster mass distribution and rank them to infer the fiducial model. Its results are then discussed and compared with those from the literature, before drawing some conclusions and commenting on the future directions that might be investigated on the basis of our findings.

Throughout this thesis, we adopt a ΛCDM cosmology with $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70$ km s^{-1} Mpc^{-1}, so that, in this cosmology, $1''$ corresponds to a physical scale of 3.31 kpc at the cluster redshift ($z_{\rm Lens} = 0.201$). Moreover, all magnitudes are measured in the AB system ($AB := 31.4 - 2.5 \log \langle f_\nu / \mathrm{nJy} \rangle$) and astronomical images are oriented north-east, with north at top and east to the left, with angles measured counterclockwise, from the west direction.

CHAPTER 1

Gravitational Lensing

The aim of this chapter is to introduce the basics of Gravitational Lensing (Sect. 1.1), namely the theoretical background used to design the mass model of galaxy clusters and to interpret the corresponding results. In particular, we focus on the strong regime on cluster scales and on lensing events involving cosmological distances. We present the axisymmetric mass profiles which are commonly adopted in lens modeling (Sect. 1.2) and, in particular, the elliptical isothermal model we assume to describe the diffuse mass component of the cluster A2163 (Sect. 1.2.3). Finally, the last section (Sect. 1.3) is intended to provide some interesting gravitational lensing applications in galaxy cluster studies from a cosmological perspective.

1.1 Theoretical basics

1.1.1 Light deflection and the *factor 2*

The gravitational lensing (hereafter, GL) effect is the phenomenon for which the electromagnetic radiation, departing from a luminous object (the *lensed source*), is deflected in the proximity of a mass distribution (the *gravitational lens*), placed between the source and an observer. In the relativistic interpretation, massive objects curve the space-time continuum and light rays, which follow this curvature, are bent passing close to the deflectors. Thus, a gravitational lens essentially acts on light paths as a normal lens does in geometrical optics. Depending on the mass of the lens, the nature of the source, and the geometric configuration, different phenomena can be observed:

- the lens can produce one image of the source, with a slight displacement and deformation (weak GL regime);
- it can split a source into (very) distorted and magnified multiple images (strong GL regime);
- the lens may generate an apparent brightness change and a tiny displacement of the light centroid, when the separation between multiple images is not resolved (microlensing regime).

The bending of light rays can be described by a *deflection angle*, $\hat{\alpha}$, which represents the total deflection undergone by light during its propagation, as calculated within the context of General Relativity (GR, Einstein 1915). In the case of a point mass, M, and under the assumption of *weak field* (see below), the deflection angle is

$$\hat{\alpha} = \frac{4GM}{bc^2} \quad , \tag{1.1}$$

where b is the impact parameter (i.e., the distance of closest approach to the lens) and G and c are the Newtonian gravitational constant and the speed of light in vacuum, respectively. Thus, the deflection angle is larger for light rays passing closer to a gravitational lens and for more massive lenses.

The correct determination of the numerical factor in Eq. (1.1) can be regarded as a watershed in the theoretical predictions of the last few centuries about the deflection angle value produced by a point mass: the value predicted by Einstein's theory of GR is twice that obtained in the limit of Newtonian mechanics and officially calculated by Soldner (1804). The factor 2 was tested for the first time during a total solar eclipse in 1919 (Dyson et al. 1920). To measure the change in the apparent positions of stars close in projection to the solar limb, two expeditions were organized: to Brazil and to Principe, an island off the Western coast of Africa. In the latter, Einstein's prediction was confirmed by Eddington, who measured a deflection angle of $\hat{\alpha} = 1''.75$. Together with the explanation of the precession of Mercury perihelion, this episode convinced the scientific community to widely accept the theory of GR. In the years between 1911 and 1915, there were several experimental efforts to measure the amount of light deflection during solar eclipses, but none of them succeeded in collecting useful data, due to cloudy skies, logistical problems and, in particular, the outbreak of World War I. Before November 1915, the value of the deflection angle predicted by Einstein (in 1911) was consistent with that obtained from Newtonian physics and it was doubled only few years later (Einstein 1915), as he completed GR. This means that, ironically, if those early tests to Einstein's light-bending prediction had been successful, they would have proven Einstein wrong.

The factor 2 in Eq. (1.1) is thus a distinctive result of Einstein's gravitation. Despite this, the computation of the same (correct) value of the deflection angle via GR can be replaced by a simpler dissertation, which is not based on the study of the GR field equations (see Schneider et al. 1992). In fact, the same expression for $\hat{\alpha}$ can be equivalently obtained by the Fermat's principle applied to light propagation in geometrical optics (where light's wave aspects are neglected). To this end, the gravitational lens potential must be related to a (variable) refraction index, under certain assumptions: the region where light is deflected is small compared to the overall propagation length (see Sect. 1.1.2) and the Born approximation can be used[1]. However, the essential working hypothesis is the weak field approximation: the Newtonian gravitational potential, ϕ, has to be

$$\phi \ll c^2 \quad . \tag{1.2}$$

In geometrical optics, the speed of the light, c', is reduced in a medium of refraction index n, such as $c' = c/n$. On the other hand, based on the aforementioned assumptions, the light slowdown in a gravitational field can be written as $c' \simeq c(1 + \frac{2\phi}{c^2})$. For condition (1.2), the refraction index in a weak gravitational field is thus

$$n \simeq 1 - \frac{2\phi}{c^2} \quad . \tag{1.3}$$

Hence, this quantity depends only on the matter distribution of the lens (via ϕ) and does not vary with the radiation wavelength, as it does in geometrical optics, meaning that gravitational lenses are completely achromatic.

[1] It is a scattering theory assumption, which allows to extend the interval of the integration along the light path from the effective light travel region to $[-\infty, \infty]$.

1.1.2 Thin lens approximation and the ray-tracing equation

Consider a three-dimensional mass distribution, acting as a gravitational lens, L, and placed between an observer, O, and a light source, S. Assume, then, that the lens has an extension along the line of sight (hereafter, *los*) which is small compared to the mutual distances between the observer, the lens and the source. In this way, light rays departing from the background object are deflected in a small region in proximity of the lens, compared to their paths before and after the lensing event. For this reason, the deflector can be approximated to a *thin* mass distribution causing an instantaneous change in the ray direction across a single plane, which contains the lens center of mass and is perpendicular to the *los*. We refer to the latter as the *lens plane* (or the *image plane*), π_L, and for the background source an analogous plane can be conceived, the *source plane*, π_S. The situation sketched above represents the so-called *thin lens approximation* (or *thin screen approximation*) and applies very often in astrophysics: even when the gravitational lenses are galaxy clusters, the biggest and most massive structures in the universe, their *los* extension is usually smaller than the distance between the lens and the source, D_{ls}, and that between the observer and the lens, D_l, which are typically on the order of Gpc. The distance between the source and the observer is denoted with D_s and is an angular-diameter distance, as well as D_l and D_{ls}. In general, all gravitational lensing distances are not additive and the relation $D_s = D_l + D_{ls}$ does not hold.

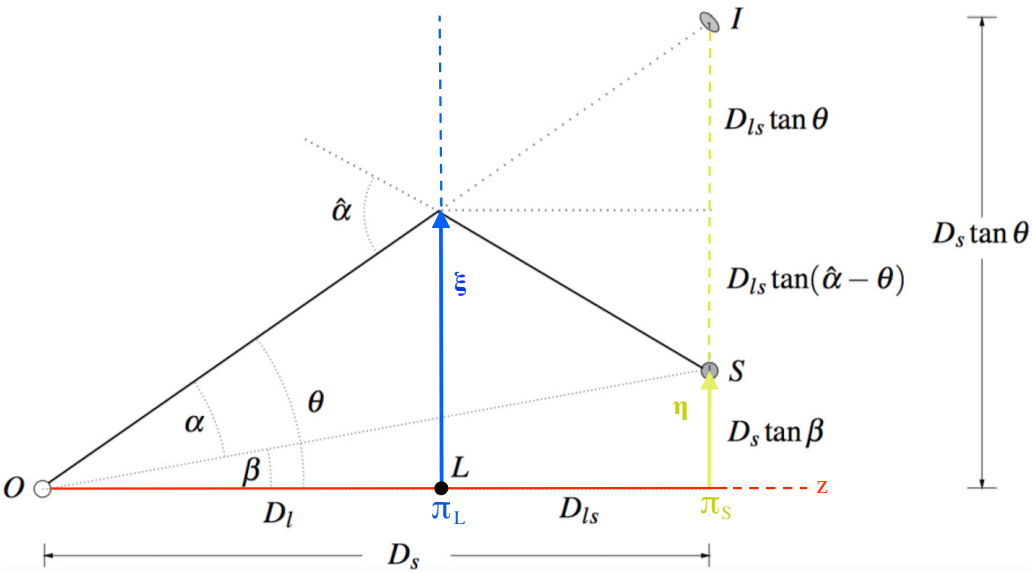


Figure 1.1: Representation of the geometrical configuration of a possible gravitational lensing event (see the text for the description); the figure is a reimagining of that taken from A. B. Congdon (2018).

A possible gravitational lensing event is illustrated in Fig. 1.1. Here, for simplicity, the plane of the lens and that of the source are shown sideways as vertical, dashed lines (in blue and green, respectively). Each of these planes contains one vector: $\boldsymbol{\xi}$ (on π_L and in blue), which is the impact parameter with the origin in the center of mass of the lens matter distribution, and $\boldsymbol{\eta}$ (on π_S and in green), which is the position of the source from the origin of π_S; the red half-line is the *optical axis*, z, defined as the axis perpendicular

to both the source and the lens planes and passing through the observer and the lens center of mass. The vectors $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ can be expressed in terms of the angular-diameter distances through the relations

$$\boldsymbol{\xi} = D_l \boldsymbol{\theta} \ , \qquad \boldsymbol{\eta} = D_s \boldsymbol{\beta} \ , \tag{1.4}$$

where the vectors $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ have, respectively, a magnitude of θ, the amplitude of the angle subtended on the sky by the source image (I), and of β, the amplitude of the angle subtended by the source, both measured from the optical axis. Eqs. (1.4) are only valid in the limit of small angles. With this assumption, each tangent function on the right part of Fig. 1.1 approaches to the value of its angle and we can write

$$D_s \boldsymbol{\theta} = D_s \boldsymbol{\beta} + D_{ls} \hat{\boldsymbol{\alpha}} \ , \tag{1.5}$$

where $\hat{\boldsymbol{\alpha}}$ is the deflection angle[2]. Rearranging the terms and defining the *scaled deflection angle* as

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) \equiv \frac{D_{ls}}{D_s} \hat{\boldsymbol{\alpha}}(D_l \boldsymbol{\theta}) \ , \tag{1.6}$$

the previous equation becomes

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}) \ . \tag{1.7}$$

Hence, for a geometrically thin lens with the configuration illustrated in Fig. 1.1, there exists a specific condition, given by Eq. (1.7), in order for the light ray to be collected by an observer: this relation, known as the *ray-tracing equation*, links the true angular position of the source, $\boldsymbol{\beta}$, with its apparent angular position, $\boldsymbol{\theta}$. Also notice that the intensity of the (scaled) deflection angle vector is linear in M and, in simple cases (under the assumption of the thin lens approximation), the deflections of a set of point mass lenses can be linearly superposed. Indicating with $\boldsymbol{\xi_i}$ and M_i ($1 \leq i \leq N$), respectively, the positions and masses of N point deflectors distributed on the lens plane, the (scaled) deflection angle of a light ray crossing the plane at $\boldsymbol{\xi_i}$ will be:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \sum_i \hat{\boldsymbol{\alpha}}_{\boldsymbol{i}}(\boldsymbol{\xi}) = \frac{4G}{c^2} \sum_i M_i \frac{\boldsymbol{\xi} - \boldsymbol{\xi_i}}{|\boldsymbol{\xi} - \boldsymbol{\xi_i}|^2} \ . \tag{1.8}$$

1.1.3 Convergence and gravitational lensing potential

In general, the ray-tracing equation may admit more than one solution: for a fixed $\boldsymbol{\beta}$, the same background source can be seen at different positions in the sky, i.e., the lens can produce multiple images. Considering a source lensed by a mass distribution with density $\rho(\boldsymbol{\xi}, z)$ and surface mass density $\Sigma(\boldsymbol{\xi}) = \int_{\mathbb{R}} \rho(\boldsymbol{\xi}, z) dz$, a sufficient condition to fulfill this requirement (Schneider et al. 1992) is

$$\kappa(\boldsymbol{\xi}) \equiv \frac{\Sigma(\boldsymbol{\xi})}{\Sigma_{cr}} > 1 \ , \tag{1.9}$$

[2]Compared to Eq. (1.1), notice that now the deflection angle does not refer to the point mass case, but is the sum of the deflections produced by several individual mass components $\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \sum_i \hat{\boldsymbol{\alpha}}_i(\boldsymbol{\xi})$; the use of the superposition principle is justified by the weak field approximation, which allows the linearization of GR equations.

where $\kappa(\boldsymbol{\xi})$ is called *convergence* (or *dimensionless surface mass density*) and

$$\Sigma_{cr} \equiv \frac{c^2}{4\pi G}\frac{D_s}{D_l D_{ls}} \simeq 0.35 \left(\frac{D_s\ 1\mathrm{Gpc}}{D_l D_{ls}}\right)\frac{\mathrm{g}}{\mathrm{cm}^2} \simeq 1.66\times 10^3 \left(\frac{D_s\ 1\mathrm{Gpc}}{D_l D_{ls}}\right)\frac{\mathrm{M}_\odot}{\mathrm{pc}^2} \tag{1.10}$$

is the *critical surface mass density*, a quantity which discriminates between the *strong* and the *weak* lensing regimes and depends only on the geometrical configuration of the lens system: it is only a function of both the lens and the source redshifts, via the angular-diameter distances (once a cosmological model is chosen). A mass distribution having in some points $(\boldsymbol{\xi_0})$ a convergence $\kappa(\boldsymbol{\xi_0}) > 1$ is exceeding the critical density and will produce multiple images of the same background source.

One problem of GL theory is to find all the image positions for a given source position, i.e., to invert Eq. (1.7), which describes a mapping $\theta \mapsto \beta$ from π_L to π_S. The scaled deflection angle can be also expressed as

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \nabla_{\boldsymbol{\theta}}\psi(\boldsymbol{\theta}) \ , \tag{1.11}$$

which implies that $\boldsymbol{\alpha}$ has null curl since $\psi(\boldsymbol{\theta})$, the *lensing potential*, is a scalar field. The latter is proportional to the three-dimensional Newtonian potential projected onto the lens plane (scaled by some factors) and can be defined in terms of the convergence as

$$\psi(\boldsymbol{\theta}) \equiv \frac{1}{\pi}\int_{\mathbb{R}^2} \kappa(\boldsymbol{\theta'})\ \ln|\boldsymbol{\theta}-\boldsymbol{\theta'}|\mathrm{d}^2\boldsymbol{\theta'} \ . \tag{1.12}$$

Another way to relate these two quantities is via the *Poisson equation* in two dimensions:

$$\nabla^2_{\boldsymbol{\theta}}\psi(\boldsymbol{\theta}) = 2\kappa(\boldsymbol{\theta}) \ . \tag{1.13}$$

In general, it is possible to find an analytic expression to invert the lens equation only for the simplest lens mass models and usually the numerical inversion is non-trivial because the image multiplicity, for a given source, is not known a priori.

1.1.4 Magnification and distortion

As highlighted above, a gravitational lens generally produces a non-isotropic deflection: a light ray coming from an extended source and passing closer to the lens is more deflected than another passing farther away. For this reason, the images are also distorted. Sometimes the distortion can be very pronounced resulting in curved images of a background source known as *gravitational arcs*, a collection of which is reported in Fig. 1.2.

The local properties of the lens mapping (like the deformation of images with respect to the source) are described by its Jacobian matrix

$$\begin{aligned} A_{i,j}(\boldsymbol{\theta}) &= \frac{\partial \boldsymbol{\beta}_i}{\partial \boldsymbol{\theta}_j} = \delta_{i,j} - \Psi_{i,j}(\boldsymbol{\theta}) \ , \\ A(\boldsymbol{\theta}) &= \begin{pmatrix} 1-\kappa(\boldsymbol{\theta})-\gamma_1(\boldsymbol{\theta}) & -\gamma_2(\boldsymbol{\theta}) \\ -\gamma_2(\boldsymbol{\theta}) & 1-\kappa(\boldsymbol{\theta})+\gamma_1(\boldsymbol{\theta}) \end{pmatrix} \ . \end{aligned} \tag{1.14}$$

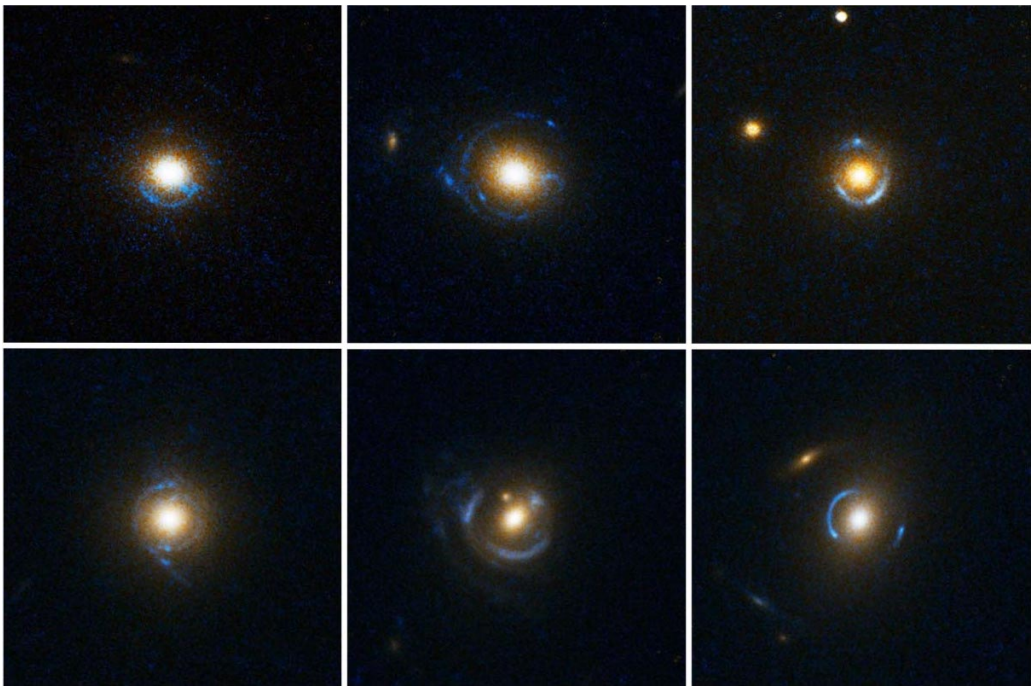


Figure 1.2: Gravitational arcs in 6 images taken from the *Hubble Space Telescope*'s Advanced Camera for Surveys (credits: NASA, ESA, A. Bolton, Harvard-Smithsonian CfA and the SLACS Team). The reddish-white objects in the center of each image are massive galaxies and the gravitational arcs are the blue curved patterns around them. Notice the high degree of distortion, especially for the arcs in the column at the right.

Here, we have used the lens equation, Eq. (1.11), and the following definitions

$$\begin{aligned}
\gamma_1(\boldsymbol{\theta}) &\equiv \frac{1}{2}\left(\Psi_{1,1} - \Psi_{2,2}\right) \ , \\
\gamma_2(\boldsymbol{\theta}) &\equiv \Psi_{1,2} = \Psi_{2,1} \ , \\
\kappa(\boldsymbol{\theta}) &\equiv \frac{1}{2}\left(\Psi_{1,1} + \Psi_{2,2}\right) \ , \\
\Psi_{i,j}(\boldsymbol{\theta}) &\equiv \frac{\partial^2 \psi(\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j} \ ,
\end{aligned} \tag{1.15}$$

where the first two quantities are the components of the *shear*, whose magnitute is

$$\gamma(\boldsymbol{\theta}) = \sqrt{\gamma_1^2(\boldsymbol{\theta}) + \gamma_2^2(\boldsymbol{\theta})} \ . \tag{1.16}$$

To understand how the Jacobian matrix could help in interpreting the image deforma-

tion, we rewrite it as the sum of two matrixes and in the diagonal form

$$\begin{aligned} A(\boldsymbol{\theta}) &= A^{iso}(\boldsymbol{\theta}) + A^{aniso}(\boldsymbol{\theta}) \ , \\ A^{iso}(\boldsymbol{\theta}) &= (1-\kappa(\boldsymbol{\theta}))\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \ , \\ A^{aniso}(\boldsymbol{\theta}) &= -\begin{pmatrix} \gamma_1(\boldsymbol{\theta}) & \gamma_2(\boldsymbol{\theta}) \\ \gamma_2(\boldsymbol{\theta}) & -\gamma_1(\boldsymbol{\theta}) \end{pmatrix} \ , \\ A^{diag}(\boldsymbol{\theta}) &= \begin{pmatrix} 1-\kappa(\boldsymbol{\theta})-\gamma(\boldsymbol{\theta}) & 0 \\ 0 & 1-\kappa(\boldsymbol{\theta})+\gamma(\boldsymbol{\theta}) \end{pmatrix} \ , \end{aligned} \tag{1.17}$$

where $A^{aniso}(\boldsymbol{\theta})$ is the so-called *shear matrix*; $\lambda_r(\boldsymbol{\theta}) = 1-\kappa(\boldsymbol{\theta})+\gamma(\boldsymbol{\theta})$ and $\lambda_t(\boldsymbol{\theta}) = 1-\kappa(\boldsymbol{\theta})-\gamma(\boldsymbol{\theta})$ are, respectively, the radial and the tangential eigenvalue of the Jacobian matrix (with respect to the lens center). We can thus conclude that the image deformation is due to an isotropic and an anisotropic contribution: the former is induced by the convergence, which rescales the source by a constant factor in all directions, and the latter is a consequence of the shear, which stretches the intrinsic shape of the source along one privileged direction. Moreover, the Jacobian matrix is symmetric and has orthogonal eigenvectors. Lensed images are stretched along the two eigendirections and are thus distorted both in shape and in size: the first kind of distortion is caused by the shear (the tidal gravitational field), whereas the latter is caused by both isotropic focusing (by the convergence) and anisotropic focusing (by the shear). Finally, since the Jacobian matrix is symmetric, notice that there is no rotation of the images. As shown in Fig. 1.3, the distortion of images, and in particular their lengthening, could be extremely strong, becoming particularly evident in galaxy cluster cores, where spectacular GL events (giant arcs) are observed.

A relevant consequence of the distortion is the image magnification (or demagnification). Let us consider the specific intensity of radiation, I_ν, at a frequency ν, flowing in a given direction, as measured in a specified local Lorentz frame. It can be demonstrated from the kinetic theory in curved space-time (sect. 22.6 in Misner, Thorne and Wheeler, 1973) that, in the vacuum, the quantity I_ν/ν^3 is conserved. Since in a gravitational lensing event the photon frequency does not change, this implies the conservation of the surface brightness, B, which is simply the specific intensity integrated over frequency. Since B is conserved, but the solid angle under which the source is seen ($d\Omega$) varies, the (monochromatic) flux of the image ($F = Bd\Omega$) has to differ from that of the source ($F_S = Bd\Omega_S$). The change in flux is measured through the *magnification* and can be related to the determinant of the Jacobian matrix:

$$\begin{aligned} |\mu(\boldsymbol{\theta})| &\equiv \frac{F}{F_S} = \frac{d\Omega}{d\Omega_S} \ , \\ \mu(\boldsymbol{\theta}) &= \frac{1}{\det A(\boldsymbol{\theta})} = \frac{1}{(1-\kappa(\boldsymbol{\theta}))^2 - \gamma(\boldsymbol{\theta})^2} \ . \end{aligned} \tag{1.18}$$

The sign of $\mu(\boldsymbol{\theta})$ is called *image parity* and a negative parity differs from a positive one, because the corresponding distorted image is specularly transformed in the former case. In general, the intrinsic luminosity of sources is unknown, hence the magnification is not an observable and cannot be used to constrain the lens model. Instead, the magnification ratio can be exploited, since it corresponds to flux ratios of different images of the same source.

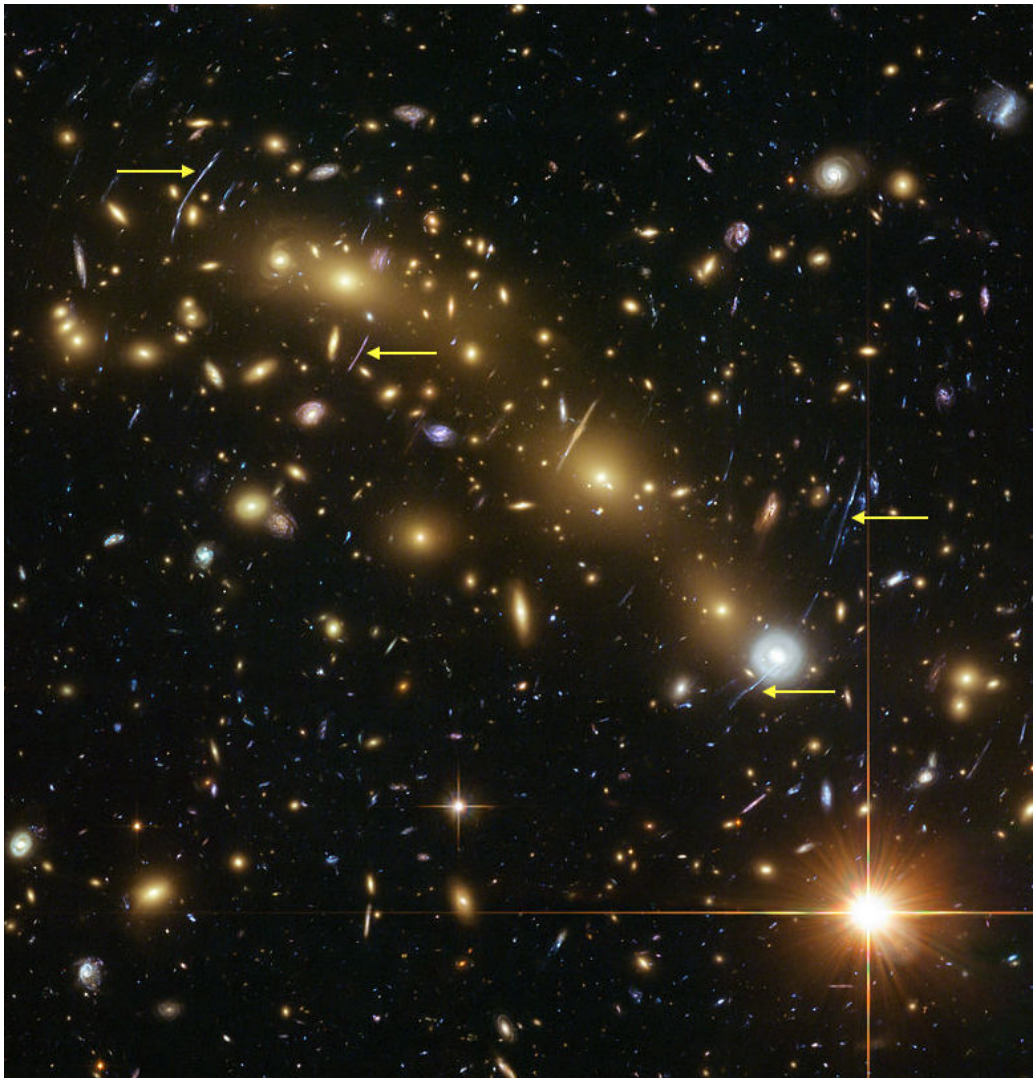


Figure 1.3: Color-composite image of the core of the galaxy cluster MACS J0416.12403 from *HST* data; the yellow arrows indicate the position of some giant arcs. In contrast to the arcs in Fig. 1.2, they are extremely elongated across very large scales, reflecting the substantial mass difference between a cluster and a galaxy and the consequent grater *lensing power* of former compared to the latter. The characteristics of galaxy clusters producing giant arcs are described in Narayan & Bartelmann (1995).

Due to the last relation in Eq. (1.18), a matrix A^{-1}, called *magnification matrix*, is considered, which corresponds to the inverse of the Jacobian one; in this way, the eigenvalues of the former are the inverse of those of the latter and it is then possible to interpret $\mu(\boldsymbol{\theta})$ as the result of two contributions, a radial and a tangential one:

$$\begin{aligned} \mu(\boldsymbol{\theta}) &= \mu_r(\boldsymbol{\theta})\mu_t(\boldsymbol{\theta}) \ , \\ \mu_r(\boldsymbol{\theta}) &= \frac{1}{\lambda_r(\boldsymbol{\theta})} = \frac{1}{1-\kappa(\boldsymbol{\theta})+\gamma(\boldsymbol{\theta})} \ , \\ \mu_t(\boldsymbol{\theta}) &= \frac{1}{\lambda_t(\boldsymbol{\theta})} = \frac{1}{1-\kappa(\boldsymbol{\theta})-\gamma(\boldsymbol{\theta})} \ . \end{aligned} \tag{1.19}$$

Thus, the eigenvalues $\mu_r(\boldsymbol{\theta})$ and $\mu_t(\boldsymbol{\theta})$ measure, respectively, the amplification in the tangential and in the radial direction (with respect to the reference frame with origin in the lens center and axes aligned with these two directions) and, theoretically, they can even diverge when, respectively, $\lambda_r(\boldsymbol{\theta}) = 0$ and $\lambda_t(\boldsymbol{\theta}) = 0$. The closed and smooth curves on the lens plane which satisfy these conditions are known as *radial critical lines* (rcl) and *tangential critical lines* (tcl), respectively, and their projection via the ray-tracing equation (1.7) onto the source plane are called *caustics* (in general, they are cuspy curves). The critical line names refer to the fact that every vector tangential to the tcl and every vector perpendicular to the rcl is an eigenvector of the Jacobian matrix, with a null eigenvalue. So, an image which is very close to the rcl will be extremely distorted perpendicularly to the rcl itself, while those images lying nearly along the tcl will be very stretched tangentially to this critical line. This holds even in complex systems, such as merging galaxy clusters, where the critical lines can be very curvy, and represents the origin of the giant arcs shown in Fig. 1.3.

1.2 Lens models of an extended mass distribution

1.2.1 Axisymmetric lenses

In the thin lens approximation, a three-dimensional mass distribution is completely described by its surface mass density $\Sigma(\boldsymbol{\xi}) = \int_{\mathbb{R}} \rho(\boldsymbol{\xi}, z)dz$, that is the density $\rho(\boldsymbol{\xi}, z)$ projected onto the lens plane. The deflection angle can be written in terms of the surface density as

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4G}{c^2} \int_{\mathbb{R}^2} \Sigma(\boldsymbol{\xi'}) \frac{\boldsymbol{\xi} - \boldsymbol{\xi'}}{|\ \boldsymbol{\xi} - \boldsymbol{\xi'}\ |^2} d^2\boldsymbol{\xi'} \ . \tag{1.20}$$

When the mass distribution is axisymmetric, the surface mass density depends only on the modulus of the impact parameter, $\xi := |\boldsymbol{\xi}|$. In this case, the deflection angle, which points along the radial direction, has the same absolute value for all the points at a distance ξ from the lens center. This vector becomes one-dimensional, because one of its two components can be considered null from symmetry. The lens equation itself is one-dimensional since all light rays of a (point) background object lie on the plane passing through the source, the lens center and the observer. Using Eq. (1.20), the modulus of

the deflection angle is

$$\begin{aligned} \hat{\alpha}(\xi) &= \frac{4GM(\xi)}{c^2\xi} \ , \\ M(\xi) &= 2\pi \int_0^{\xi} \xi' \Sigma(\xi') d\xi' \ , \end{aligned} \tag{1.21}$$

thus the deflection angle only depends on $M(\xi)$, the mass enclosed inside a circle of radius ξ, but not on its radial profile, i.e., it is independent of the angle which defines the impact parameter position on the lens plane. Notice that when the lens is strong and within a circle where the average value of the surface mass density is equal to Σ_{cr}, from Eqs. (1.21), it follows that

$$\begin{aligned} M(\xi \leq R_E) &= \pi R_E^2 \Sigma_{cr} \ , \\ \alpha(\theta_E) &= \theta_E \end{aligned} \tag{1.22}$$

and the lens equation simply implies that $\beta = 0$ for $\theta = \theta_E$. The *Einstein angle* θ_E (or, equivalently, the *Einstein radius* $R_E = D_L\theta_E$) is thus a typical length scale of strong lensing: it corresponds to the radius of a circular image on π_L which encircles the lens and is ray-traced into the origin of the source plane. In other words, when the observer, the axisymmetric lens and the source are aligned, the latter is seen as a circular image, known as *Einstein ring*.

1.2.2 A benchmark example: the isothermal sphere

We use the framework developed in the previous sections to illustrate two lens models, the singular isothermal sphere (SIS) and the cored (or softened) isothermal sphere (CIS).

The SIS model is derived assuming that the lens matter content behaves like an ideal gas, immersed in its own gravitational potential with spherical symmetry. The self-gravitating gas is in both thermal and hydrostatic equilibrium and is characterized by a constant, one-dimensional velocity dispersion, σ_v. Indicating with r the radius of the sphere, the three-dimensional mass density and the corresponding surface mass density (derived by the former through a projection along the *los*) are

$$\begin{aligned} \rho(r) &= \frac{\sigma_v^2}{2\pi G r^2} \ , \\ \Sigma(\xi) &= \frac{\sigma_v^2}{2G\xi} \ . \end{aligned} \tag{1.23}$$

This density profile diverges for $\xi = 0$ (hence the adjective *singular* in the SIS name), nevertheless it is useful, at first approximation, to describe different astrophysical objects (e.g., this profile can be used to reproduce the flat rotation curves of spiral galaxies).

From the expression of the surface mass density (1.23), we can rewrite Eqs. (1.21) in the SIS case

$$\begin{aligned} M(\xi) &= \frac{\pi\sigma_v^2}{G}\xi \ , \\ \hat{\alpha} &= \frac{4\pi\sigma_v^2}{c^2} \ , \end{aligned} \tag{1.24}$$

where the latter is independent of θ. Then, the lens equation and the Einstein angle can

be written as

$$\begin{aligned}\beta(\theta) &= \theta\left(1-\frac{\theta_E}{|\theta|}\right) \ , \\ \theta_E &= \frac{D_{ls}}{D_s}\frac{4\pi\sigma_v^2}{c^2} \ ,\end{aligned} \tag{1.25}$$

where the Einstein angle is calculated substituting the deflection angle expression, Eq. (1.24), in the lens equation and imposing $\beta = 0$; notice that it equals the scaled deflection angle. In terms of the Einstein angle, one obtains

$$\begin{aligned}\alpha(\theta) &= \theta_E\frac{\theta}{|\theta|} \ , \\ \kappa(\theta) = \gamma(\theta) &= \frac{\theta_E}{2|\theta|} \ .\end{aligned} \tag{1.26}$$

Due to the spherical symmetry, the SIS lens has a point-like caustic at $\beta = 0$ and a critical line given by the Einstein ring. By means of the above relations, we can thus conclude that the SIS model produces two images for a source with an angular position $0 < \beta < \theta_E$: they are located on opposite sides of the lens center, at $\theta_+ = \beta + \theta_E$ and $\theta_- = \beta - \theta_E$, and are always separated by the constant value $|\theta_+ - \theta_-| = 2\theta_E$; their magnifications are, respectively, $\mu_+ = 1 + \frac{\theta_E}{\beta}$ (positive) and $\mu_- = 1 - \frac{\theta_E}{\beta}$ (negative). When the source is at $\beta > \theta_E$, then only the image at θ_+ and with magnification μ_+ will be produced. The Einstein radius defines the characteristic scale of strong lensing, which discriminates between the presence or absence of multiple images.

One drawback of the SIS, however, is that it has an infinite density at its center. This singularity can be removed by introducing a small core, resulting in the finite central mass density of the CIS. For this model, the counterpart of SIS Eqs. (1.23) is given by

$$\begin{aligned}\rho(r) &= \frac{\sigma_v^2}{2\pi G(r^2+r_c^2)} \ , \\ \Sigma(\theta) &= \frac{\sigma_v^2}{2GD_l}\frac{1}{\sqrt{\theta^2+\theta_c^2}} \ ,\end{aligned} \tag{1.27}$$

where we have introduced the radius core, $r_c = D_L\theta_c$. The CIS Einstein radius, corresponding to one critical line, is

$$\theta_{E,CIS} = \theta_{E,SIS}\sqrt{1-\frac{2\theta_c}{\theta_{E,SIS}}} \ . \tag{1.28}$$

Another critical line forms at an angular radius smaller than the Einstein radius. In the CIS model, the presence of a core radius leads to the formation of three images, instead of two, when $0 < \beta < \theta_{E,CIS}$.

1.2.3 Non-circularly symmetric lenses: the elliptical isothermal model

To make lens models more realistic and well representative of extended astrophysical systems, such as galaxies and clusters of galaxies, a term of ellipticity in the mass density of circularly symmetric lenses can be considered. Another possibility is to introduce

the ellipticity directly in the projected gravitational potential, which guarantees simpler expressions for the deflection and magnification. However, already from small values of the ellipticity (~ 0.2), the resulting mass distributions have unphysical shapes (peanut shaped isodensity contours; Blandford & Kochanek 1987; Kassiola & Kovner 1993).

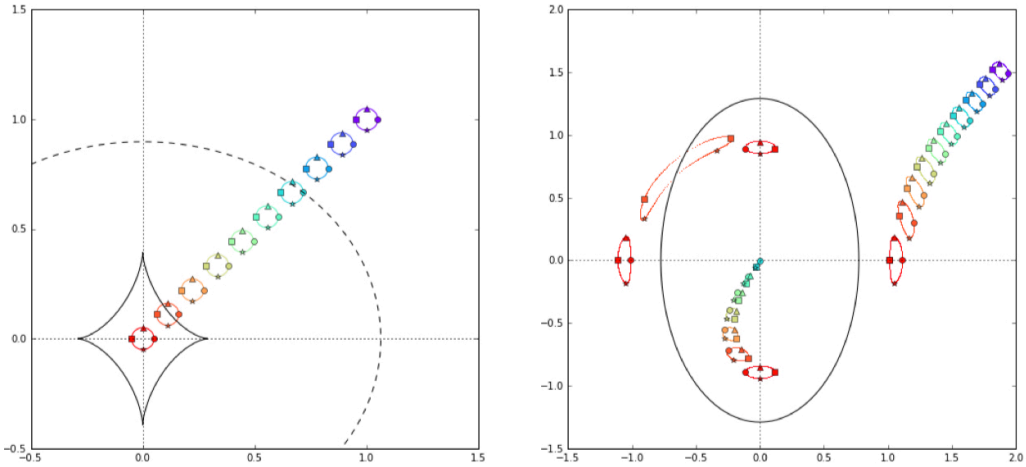


Figure 1.4: Caustics, critical lines and the formation of multiple images for an elliptical lens. The right and left panels show, respectively, the lens and the source planes; the little symbols on the circle of each source allow one to understand the parity of the corresponding multiple images.

Elliptical lenses also generate caustics and critical lines more complex than the circularly symmetric ones (Kovner 1987). As shown in Fig. 1.4, for a singular isothermal elliptical profile, the caustic is an astroid, a diamond shaped figure with four cusps, instead of being a central point-like caustic, as for systems with circular symmetry; besides, the circular radial caustic is now deformed. The presence of a further caustic line modifies the image multiplicity: since at each caustic line passage the number of images increases/decreases by two, elliptical lenses can form one and three images, as for the circular case, but also five images; in the latter case, the source is enclosed within the two caustics and a faint image appears close to lens center. In the same figure, we also illustrate the configuration and the shape of the multiple images produced for different source positions. If, for example, the source is close to the inner caustic, an arc tangentially oriented with respect to the center of the lens is formed. Finally, notice from the geometrical details on the source circumference that the images inside and outside the critical line are inverted.

A more complicated (although more general) lens mass profile is often adopted in strong gravitational lensing modeling: it is the dual pseudo-isothermal elliptical (dPIE; Elíasdóttir et al. 2007) profile, which is that mostly used later in this thesis[3]. This lens

[3]Sometimes (e.g., Limousin et al. 2005), it can be found in literature under the acronym of PIEMD, pseudo isothermal elliptical mass distribution.

model is described by the following mass densities

$$\begin{aligned} \rho(r) &= \frac{\rho_0}{(1+r^2/r_c^2)(1+r^2/r_t^2)} \ , \\ \Sigma(R) &= \frac{\sigma_0^2}{2G}\frac{r_t}{r_t-r_c}\left(\frac{1}{\sqrt{r_c^2+R^2}}-\frac{1}{\sqrt{r_t^2+R^2}}\right) \ . \end{aligned} \tag{1.29}$$

Here, R is the projected radial distance from the mass distribution center, r_t is the truncation radius and σ_0 is the central velocity dispersion for a circular potential, related to ρ_0 by the expression (Limousin et al. 2005)

$$\rho_0 = \frac{\sigma_0^2}{2\pi G}\left(\frac{r_t+r_c}{r_t r_c^2}\right) \ . \tag{1.30}$$

In the center, the profile has a core with central density ρ_0, i.e.,

$$\rho \simeq \frac{\rho_0}{1+r^2/r_c^2} \ . \tag{1.31}$$

In the transition region ($r_c < r < r_t$) the profile is approximately isothermal, $\rho \simeq r^{-2}$, and, in the outer parts, the density falls off as $\rho \simeq r^{-4}$. The enclosed two-dimensional mass interior to radius R and the total mass are given, respectively, by

$$\begin{aligned} M(<R) &= \frac{\pi r_t \sigma_0^2}{G}\left(1-\frac{\sqrt{r_t^2+R^2}-\sqrt{r_c^2+R^2}}{r_t-r_c}\right) \ , \\ M_{tot} &= \frac{\pi\sigma_0^2}{G}\frac{r_t^2}{r_t+r_c} \ . \end{aligned} \tag{1.32}$$

Including a possible elongation term on the plane of the sky, R can be redefined in terms of the ellipticity, ε, and indicated with R_ε

$$\begin{aligned} R_\varepsilon^2 &\equiv \frac{x^2}{(1+\varepsilon)^2}+\frac{y^2}{(1-\varepsilon)^2} \ , \\ \varepsilon &\equiv \frac{1-q}{1+q} \ , \end{aligned} \tag{1.33}$$

where $q \equiv b/a$ is the minor to major axis ratio. Finally, the advantage of adopting this profile is that, by introducing a truncation radius, the total mass becomes finite. In fact, the surface mass density in Eq. 1.29 is given by the difference between two cored isothermal profiles and, thus, the deflection can be easily calculated by the superposition principle.

1.3 Lensing by galaxy clusters: a cosmological perspective

GL is a promising and exciting research line. With its wide phenomenology, this discipline is a unique multi-purpose tool, serving not only to quantify DM and to provide hints on its cross-section properties, neither only as a natural telescope to study magnified background sources, but it can accommodate a diagnosis of cosmological models and of galaxy formation and evolution theories.

Some scientific researches are best suited to one of the three regimes, namely micro, strong, and weak lensing: extrasolar planet detection (microlensing), DM distribution in galaxies and in galaxy cluster cores (strong lensing), cosmological parameter estimate (weak lensing). Moreover, the combination of GL studies in different regimes results in constraining better the astrophysical system under analysis. In this respect, deriving with weak plus strong lensing precise maps of the DM distribution in complex systems like clusters of galaxies offers one relevant research opportunity. In this section, we present two important examples of galaxy cluster studies which take advantage of weak and strong GL for cosmological purposes. The applications of the microlensing regime are beyond the scope of this thesis, which deals with extended mass distributions[4].

1.3.1 Cosmological parameter estimate using weak lensing

Understanding how the entire cluster population evolves with time, namely how its properties change with redshift, provides some of the most powerful constraints on cosmological models. An important instrument in this respect is the cluster mass function, giving the number density of clusters in a comoving volume element as a function of the mass and the redshift. In particular, to know the dependence on redshift of the observables which trace the cluster mass function can considerably constrain the cosmological parameters. In fact, the evolution of this function is controlled entirely by Ω_M, Ω_Λ and ω, respectively, the matter and dark energy densities and the parameter of the cosmological equation of state. Indicating with X a certain observable, the quantity to measure is the number of clusters within a given solid angle and redshift interval that falls into the range $[X, X + dX]$. The distribution of clusters of a certain mass with redshift is then derived directly from the observations if the relation between the mass and the generic observable X is known. Thus, constraining the cosmological parameters through the redshift distribution of clusters depends on the ability to understand the evolution of the mass-observable relations.

The X-ray temperature and the X-ray luminosity are among the common observables which trace the cluster mass and are related to the latter through simple parametric relations. In cosmology, calibrating the normalization and the scatter in these relations is thus a major motivation for measuring the mass of clusters using gravitational lensing. The weak GL technique used to measure the mass of a cluster is based on the fact that background galaxies are distorted due to the lensing shear, that is the distortion effect on the shape of these far away galaxies in some particular directions. The way in which the shear varies on the sky depends on the cluster mass distribution, so that understanding how the distortion of background galaxies changes allows us to measure the mass of the cluster. Although the intrinsic shapes of these sources are not known, assuming that their ellipticities are randomly distributed implies that the mean ellipticity is zero. Considering a large number of background galaxies, it is thus possible to deduce in a statistical way the effect of the distortion due solely to the weak lensing. It is a non-easy task, because the galaxy shape measurements are affected by different factors, such as the fact that astronomical images are inevitably pixellated and with a certain degree of noise. Moreover, the instrument resolution is a crucial factor for an accurate shape measurement, so that this technique needs large area surveys with a well controlled correc-

[4] We refer the reader to Tsapras (2018), who describes the fundamental theoretical concepts in microlensing, reviews the method for exoplanet searches, addresses how observations are performed in practice and discusses the future prospects of microlensing.

tion of the point spread function. On the other hand, weak lensing mass measurements are largely independent of the X-ray temperature and luminosity observables and do not depend on baryonic physics. The main systematic problem is the excess of mass due to the sources along the *los* that are outside the cluster virial radius.

1.3.2 Dark matter profiles of clusters using strong lensing

Galaxy cluster observations have long indicated that member galaxies have a relatively constant velocity dispersion, implying a mass density profile $\rho(r) \propto r^{-2}$. The singular isothermal sphere, introduced in Sect. 1.2.2, represents the simplest analytical model which is consistent with this profile. Although approximated analytical estimates of cluster properties can be inferred using the SIS model, more sophisticated forms for the density profile are needed to reproduce the observations. DM halos in numerical simulations show density profiles which deviates from the isothermal one, in particular at small radii. They can be represented by the following generic form

$$\rho(r) \propto r^{-p}(r + r_s)^{p-q} \ , \tag{1.34}$$

where p and q describe the inner and outer slopes and r_s is the radius where the profile steepens. The case in which $p = 1$ and $q = 3$ corresponds to the most widely used fitting formula to represent cosmological simulation results: the Navarro, Frenk and White's profile. Its description is detailed in Sect. 2.2.2, where we also discuss the issue due to the discrepancy between simulations and observations. Here, we want to highlight the contribution of strong lensing studies in constraining the asymptotic inner slope p, which, as explained, is of great observational interest to test the CDM paradigm.

With its 102 spectroscopically-confirmed multiple images, the *Hubble Frontier Fields* galaxy cluster MACS J0416.1-2403 has the best dataset currently available for strong GL analyses. Annunziatella et al. (2017) present a decomposition at high resolution of its projected total mass distribution, using *HST* imaging, VLT/MUSE spectroscopy, and *Chandra* data to separate the stellar, hot gas, and DM mass components in the inner 300 kpc of the cluster. Their strong lensing analysis, including about 200 cluster members (75% confirmed spectroscopically), shows a baryon fraction of ∼10% in the outermost region and that, from 10 to 300 kpc from the cluster center, the stellar mass fraction decreases from 12% to 1% and the gas mass fraction increases from 3% to 9%. The total mass in the cluster substructures due to the stellar component represents ∼30%, near the cluster center. They divide the DM component in two halo families, diffuse halos and sub-halos, and consider three kinds of surface mass density profiles: the one including all the mass components (the "total" mass profile), that of all the DM components (the "global" DM mass profile) and that due only to the mass of the cluster-scale DM halo. Thus, in the latter and in the global profile, the mass contribution of the baryonic component is not included. As shown in Fig. 4 of that paper, within few times the BCG effective radius from the cluster center (coincident with the BCG position), the surface mass density profile of the total mass is steeper than a cored or a NFW profile, as well as that of the global DM, while the profile of the cluster-scale DM halo is flatter than a NFW profile. The difference among these three profiles is thought to be due to the inclusion of the DM halo of the BCG.

Strong lensing accurate studies like that presented above will contribute in increasing the sample of galaxy clusters with robust reconstructions of the inner DM distribution. Currently, this sample is not yet large enough to discriminate between different structure formation scenarios. Specifically, one of the difficult aspects in determining precise mass profiles is to separate the dark and baryonic components. As explained, strong lensing can be used to overcome this problem: to disentangle the DM only component is possible by subtracting the baryonic mass contribution from the cluster total mass profile measured via strong lensing. Nevertheless, we also remark that GL measures the cluster projected total mass.

Finally, galaxy cluster mass diagnostics by strong GL can complement those by weak lensing (e.g., Bradač et al. 2005a,b) since the use of these two tools is valid over a specific radial interval: the latter can (safely) constrain the outer density profile until about $\sim$200 kpc, while the former is adopted for scales below $\sim$150 kpc, where the lens reaches the critical density. The density profile of the most internal cluster region ($<$30 kpc) is constrained by stellar kinematics studies of the BCG. An additional advantage in using GL for cluster mass diagnostics is the exquisite mass measurement, deriving from the independence of assumptions on the cluster dynamical state. Moreover, strong GL does not require the condition of the hydrostatic equilibrium, which is necessary for X-ray mass estimates and, thus, it is especially useful in the case of clusters with non-negligible merging processes.

CHAPTER 2

Galaxy clusters

This chapter addresses the definition, the classification(s) and the general properties of galaxy clusters (Sect. 2.1), together with a comprehensive description of the primary and minor components of these systems, respectively, the dark matter, the intracluster medium, the galaxies (Sects. 2.2, 2.3, and 2.4) and, from the other hand, the intracluster light and the dust (Sect. 2.5). During the analysis, we also discuss the reliability of the most common assumptions at the base of cluster mass diagnostics.

2.1 Generality

2.1.1 Galaxy cluster formation

In the current, standard paradigm of cosmic structure formation, galaxy clusters are the final result of hierarchical merging processes (see e.g., Bond et al. 1991; Lacey & Cole 1993; Somerville & Kolatt 1999). They primarily involve small structures of CDM, which represents more than 80% of the matter content of the present universe (see Sect. 2.2.1). The perturbations of the initial, smooth universe thicken due to the gravitational attraction and decouple from the universe's expansion as a consequence of gravitational instabilities. CDM dominates the gravitational field and drives a bottom-up assembly process: sub-halos, that are smaller structures, form sooner (with the first-formed galaxies inhabiting their cores), then merge to form bigger ones on larger scales; this hierarchical process continues until the high-mass tail of cosmic structures decouples from the *Hubble flow*[1].

Galaxy clusters are thus relatively young systems at the interface of cosmological large-scale structures, dominated by DM and dark energy, and individual small-scale objects, strongly affected by baryonic matter physics. The only larger systems are the super-clusters (aggregations of clusters of galaxies), that form at the cross-roads of the filamentary structure of the *cosmic web*. For example, the Milky Way neighborhood contains a small number of galaxies (about 35), called the *Local Group*, and, relatively close to it, there is the *Virgo cluster*, containing up to 2000 galaxies: both are part of the *Virgo super-cluster*.

In theoretical models, the general bulk properties of galaxy clusters are determined solely by three ingredients: the cosmological initial conditions, gravity and dissipationless DM halos. Models of this kind, such as the self-similar model of clusters (Kaiser

[1] This term defines the galaxy motion due solely to the expansion of the universe, which is currently dominated by dark energy.

1986), despite their simplicity, can predict the abundance and the clustering of clusters; the functions defining the correlations between cluster properties can also be derived and they, in turn, allow to describe how the scaling relations of clusters evolve. In general, although the issues presented in the Introduction, the results from cosmological N-body simulations are consistent with several observed quantities of real clusters, but there are still some discrepancies, mostly related to the physics of baryons[2], e.g., the *cool-core problem* (Peterson & Fabian 2006) and the *over-cooling problem* (Balogh et al. 2001). Regarding the formation history of galaxy clusters, several lines of evidence suggest that observed clusters of galaxies in the Universe might have assembled their mass earlier than in simulations. Two examples are the detection of very massive galaxy clusters at $z > 1$ (in Jee et al. 2011) and the measurements in clusters at intermediate redshift ($z \sim 0.3$) of values of mass concentration well in excess of the cosmological predictions (Oguri et al. 2009). In conclusion, cosmological N-body simulations predict well the overall phenomena governing the formation of galaxy clusters, but not in every phase. In particular, the non–linear phase in the collapse and the baryonic dissipative processes have to be considered and modeled in detail, in order to reach a complete description of the cluster formation.

2.1.2 Definitions

It is not trivial to establish distinctly what a galaxy cluster might be. The most common definitions refer to a cluster as a region within a certain well-defined radius or containing a certain number of galaxies. We briefly analyze these possibilities in the present section, but preliminarily and as a general rule we can simply consider them as gravitationally bound systems consisting of mostly DM ($80\% - 85\%$), a hot and diffuse intracluster medium (ICM) ($10\% - 15\%$) and bright galaxies[3] ($\sim 5\%$), whose number is greater than ~ 50. Hence, as the mass balance in clusters shows, they are DM-dominated structures.

It is particularly difficult to define uniquely when a set of galaxies forms a clustered structure, since different criteria for setting its radius exist. The picture has gotten progressively more complicated due to observations in additional wavelength bands (other than the optical one). They show new cluster mass components with different distributions, for instance, the gas observed in the X-ray images: it is a diffuse and continuous component in the cluster core, compared to galaxies, which instead populate also the outer regions in a discrete way. The structure formation model, within the framework of ΛCDM cosmology, has an undeniable predictive power and also proposes one of the possible definitions of the cluster radius. In general, it is the radius within which the average density is greater, by a specified factor, than the universe critical density:

$$\rho_{\rm crit} = \frac{3H^2}{8\pi G} \ , \tag{2.1}$$

where H is the Hubble parameter, varying, at redshift z and for a matter density parameter $\Omega_{m,0}$, as

$$H(z) = H_0[\Omega_{\rm m,0}(1+z)^3 + \Omega_{\Lambda,0}]^{\frac{1}{2}} \ . \tag{2.2}$$

[2]Within the *Standard Model of Particle Physics*, the term baryons denotes half-integer-spin hadrons consisting of three quarks (like the proton and the neutron, the most widely known); on the other hand, in astrophysics, this word is a blanket term that stands for (practically) all the particles described by this standard model.

[3]Note that with the term "galaxy" we refer only to a baryonic structure, without any surrounding DM halo.

Under the spherical approximation and choosing 200 as the reference numerical factor[4], the cluster radius, r_{200}, and the corresponding mass, $M_{200} := M(r < r_{200})$, are defined through the equation

$$M_{200} = \frac{4}{3}\pi r_{200}^3 \cdot 200\rho_{\text{crit}} \ . \tag{2.3}$$

A common alternative to these definitions is to fix the relation between the radius and the mass of a cluster as a result of the scalar condition of the virial equilibrium:

$$2T - |U| = 0 \ , \tag{2.4}$$

where T is the kinetic energy and U the gravitational potential energy. The *virial radius*, r_{vir}, i.e., the radius within which the mass distribution is in virial equilibrium, and the *virial mass*, $M_{\text{vir}} := M(r < r_{\text{vir}})$, are defined rewriting more explicitly Eq. (2.4) as

$$\langle v^2 \rangle - \frac{GM_{\text{vir}}}{r_{\text{vir}}} = 0 \ , \tag{2.5}$$

where we have used $\langle T \rangle = \frac{1}{2}m\langle v^2 \rangle$ and $U = -\frac{GmM_{\text{vir}}}{r_{\text{vir}}}$, with m and v, respectively the mass and the velocity of galaxies. Eq. (2.5), for spherically symmetric systems with an isotropic velocity distribution, translates into

$$M_{\text{vir}} \sim \frac{3\sigma_{\text{los}}^2 r_{\text{vir}}}{G} \ , \tag{2.6}$$

where σ_{los} is the one-dimensional velocity dispersion of the cluster members, which can be measured from the galaxy spectroscopic redshifts.

As anticipated, the definition of galaxy clusters can also be based on the galaxy count, rather than on fixing a well-defined radius. In particular, the classical description of clusters as significant concentrations of galaxies (in the optical band) represents the original approach to define these systems (e.g., Abell 1958). We describe in subsection 2.1.4 how such a definition was exploited to detect and classify clusters, e.g., leading to the famous Abell catalog and its southern extension, the *ACO* catalog (Abell et al. 1989), still heavily used. The condition to be satisfied by a collection of N galaxies to be considered a galaxy cluster (Bahcall 1977a) makes use of the mean ratio between the surface number density, n, and that of a uniform background, n_{bg}:

$$\left\langle \frac{n}{n_{bg}} \right\rangle \geq N \ . \tag{2.7}$$

The value of N is usually set to ~ 50: too low or too high values would result, respectively, in satisfying that condition more or less easily and would imply an overabundance of galaxy clusters or the opposite situation. Small galaxy over-densities ($N < 50$) are called *galaxy groups* and they differ from galaxy clusters for important characteristics (such as the kind of galaxies they contain), making the reference value of N even more meaningful. Nevertheless, there is not an unambiguous distinction between groups and clusters.

[4] This value is roughly the typical over-density in a *spherical top-hat collapse*, one of the well-known models proposed to explain the formation of cosmic large-scale structures (Gunn & Gott 1972).

2.1.3 Properties

Here, we quote some typical values of the properties of clusters, derived by an order-of-magnitude analysis; we also circumscribe the validity of the main assumptions of the formulae used for such calculations.

Scaling conveniently Eq. (2.6) and transforming the three-dimensional radius to the projected one, R, the total mass of a cluster is directly related to the observables:

$$\frac{M_{\rm vir}}{10^{14}\,{\rm M}_\odot} \sim 7 \left(\frac{\sigma_{\rm los}}{10^3\,{\rm km/s}} \right)^2 \frac{R}{\rm Mpc} \quad . \tag{2.8}$$

The virial equilibrium assumptions usually hold for single galaxies and for the core of galaxy clusters (if no merging processes are present), but they result less representative for peripheral cluster regions, which are often yet in the gravitational collapse phase. However, it would be questionable, in general, to consider cluster phenomena as independent of the cluster environment, so that virial mass estimates should be applied only in specific cases (for *regular* clusters, see subsection 2.1.4). The first dynamical analysis of clusters using the virial theorem (Zwicky 1933) showed that the existence of an excess of gravitational material compared to that indicated by the stellar content of galaxies. Pointing out the first evidence for the existence of DM, this study inaugurated a series of other works, which compared the dynamical (i.e., total) and luminous mass of clusters and quantified the amount of this *missing mass* through their total *mass-to-light ratio*. For galaxy clusters, typical values of this quantity are greatly higher than in individual galaxies and amount to $M/L \sim 200 - 500\,{\rm M}_\odot/{\rm L}_\odot$, where L and ${\rm L}_\odot$ indicate the optical luminosity of a cluster and that of the Sun, respectively.

Another relevant characteristic of a galaxy cluster is the temperature of the ICM, $T_{\rm ICM}$. If we approximate the gas as a system of n protons of mass $m_{\rm p}$, which move similarly to the galaxy component[5], we can roughly estimate its temperature by equating the thermal and kinetic energy of the gas:

$$\frac{3}{2} n k T_{\rm ICM} = \frac{1}{2} n m_{\rm p} \langle v^2 \rangle = \frac{3}{2} n m_{\rm p} \sigma_{\rm los}^2 \quad , \tag{2.9}$$

where k is the Boltzmann constant. After rescaling this expression, two practical quantities used to measure the ICM temperature can be derived:

$$\begin{aligned} \frac{k T_{\rm ICM}}{\rm keV} &= 10.44 \left(\frac{\sigma_{\rm los}}{1000\,{\rm km/s}} \right)^2 \quad , \\ \frac{T_{\rm ICM}}{10^8\,{\rm K}} &= 1.21 \left(\frac{\sigma_{\rm los}}{1000\,{\rm km/s}} \right)^2 \quad . \end{aligned} \tag{2.10}$$

Observations show a small deviation from the law $T_{\rm ICM} \propto \sigma_{\rm los}^2$ (Xue & Wu 2000), so the idealized picture for which both the gas and the galaxies move in the same common gravitational potential is a reasonable representation.

For a typical virial radius and common values of the velocity dispersion of cluster galaxies, a (order-of-magnitude) summary of characteristic values which apply to clus-

[5] This means to assume an equilibrium, for which the galaxy velocities equal the gas thermal velocities.

ters can be easily derived from Eqs. (2.8) and (2.10):

$$\begin{aligned} r_{\rm vir} &\sim 1 - 4\,{\rm Mpc}\ , \\ \sigma_{\rm los} &\sim 5 \times 10^2 - 10^3\,{\rm km/s}\ , \\ M_{\rm vir} &\sim 10^{14} - 10^{15}\,{\rm M}_\odot\ , \\ \frac{M}{L} &\sim 200 - 500\,\frac{{\rm M}_\odot}{{\rm L}_\odot}\ , \\ {\rm k}T_{\rm ICM} &\sim 3 - 10\,{\rm keV}\ \ (T_{\rm ICM} \sim 10^7 - 10^8\,{\rm K})\ . \end{aligned} \tag{2.11}$$

2.1.4 Cluster classification(s)

The first (optical) classification of galaxy clusters was performed by Abell (Abell et al. 1989). He identified them by visual inspection of photographic plates, using data from the *Palomar Observatory Sky Survey* (POSS), a photographic atlas of the Northern sky. Abell excluded the region of the Milky Way disk because of the dust extinction and the high stellar density, which make observations considerably more complicated. His catalog of 2712 clusters has been extended thanks to observations from the Southern sky, leading to the ACO catalog, which contains a total of 4073 objects (including the members of the original catalog).

Abell's criteria for the cluster detection can be summarized as follows:

- the cluster must lie north of declination $-27°$;
- the number of members has to be ≥ 50;
- the apparent magnitude of the latter has to be in the range $[m_3, m_3 + 2]$, where m_3 refers to the third brightest galaxy of the cluster[6];
- galaxies must lie within an angular radius $\theta_{\rm A} = 1'.7/z$ (*Abell radius*) from the cluster center, where z is the redshift, estimated assuming the same luminosity of the tenth brightest galaxy for each cluster[7]. The physical Abell radius usually corresponds to ~ 2 Mpc;
- the cluster redshift must be within $0.02 < z < 0.2$.

The *cluster richness*, a parameter introduced by Abell, is commonly used as a qualitative measure of the mean number density of galaxies, since it is the number of members within the Abell radius. It is thus possible to group clustered galaxies in six richness classes, from 0 to 5, all corresponding to galaxy clusters, except class 0, on the basis of the Abell's definition and the values reported in Table 2.1.

Other systematic classifications of clusters take into account the morphology of the galaxy distribution. They could provide a general picture of the dynamical state of the cluster (dynamical equilibrium state, interactions due to merger processes, etc.). Specifically, the following two categories of clusters can be considered (with continuous and intermediate types), respectively, in a relaxed state and still in the process of evolution:

[6] This choice overcomes two problems: the variation of the luminosity of the brightest galaxy among clusters and the finite probability for it to be bright because it is actually a foreground interloper.

[7] It has been found that these redshifts have an error of about 30%.

- **regular or early-type (more dynamically evolved)**. Compact clusters, with circular symmetry and a large central galaxy density; they are completely dominated by early-type galaxies, especially in the core, where often a cD galaxy (see below) resides. They rarely have any sub-cluster, but a high richness, with at least 10^3 galaxies in the brightest six magnitude range.

- **Irregular or late-type (less dynamically evolved)**. Non-symmetric clusters, not particularly dense in the center, but often with strong sub-structures; they have fewer cluster members than regular clusters ($< 10^3$), with a population of spirals and irregulars nearly as large as in the field galaxy distribution.

Examples of a regular and of an irregular cluster are given, respectively, in Figs. 2.1 and 2.2.

Class	galaxy number
0	$30 - 49$
1	$50 - 79$
2	$80 - 129$
3	$130 - 199$
4	$200 - 299$
5	≥ 300

Table 2.1: Abell's richness classes.

The first, extensive morphological catalog was the Zwicky's "Catalog of Galaxies and Clusters of Galaxies" (Zwicky et al. 1961-1968). It had no limit on the cluster redshift and contained 9700 clusters, with 30000 galaxies (brighter than magnitude 15.7) identified on the *Palomar Sky Survey* plates. Clusters classified by Zwicky as *compact* and *open* correspond, respectively, to the regular and irregular types above introduced. His selection criteria, listed below, were less strict than Abell's ones:

- the clusters must lie north of declination $-3°$;

- the number of members has to be ≥ 50;

- the apparent magnitude range of the latter must be in the range $[m_1, m_1+3]$, where m_1 refers to the brightest galaxy;

- these galaxies must lie within a region inside which the value of the projected galaxy density is twice that of the neighboring field.

Due mainly to the different and less restrictive selection criteria, the Zwicky's catalog, compared to the Abell's one, has more clusters and they are mostly larger, with a lower density and with more multiple dense clumps. Finally, Zwicky's populations of clusters depend systematically on their redshifts (unlike the Abell's classification).

The morphological criterion was also followed by Bautz and Morgan (Bautz & Morgan 1970) in a classification based on the type of luminous galaxies which dominate the cluster, as well as in the successive Rood and Sastry's (RS) catalog. In particular, RS developed a classification method, which considers the projected distribution of just the brightest ten members (Rood & Sastry 1971). Clusters in their classification appear



Figure 2.1: The Coma cluster (at a distance of $\sim$99 Mpc) is a classical example of a regular cluster, with practically no spiral galaxies and a core dominated by a pair of super-giant elliptical galaxies: NGC 4874 and NGC 4889 (picture taken at the *Mount Lemmon SkyCenter*, using the 0.8 m Schulman Telescope).

to have affinities with the regular-irregular scheme already presented, when they are arranged in the "tuning fork" diagram, shown in Fig. 2.3. Here the concentration increases leftward, thus the more irregular (regular) clusters are on the right (left). More in detail, the RS morphological classification distinguishes among the following cluster kinds:

- **cDs (super-giant)** are dominated by the brightest central galaxy, labeled with *cD*. It differs from normal elliptical galaxies in that it has a much more extended brightness profile; a cD galaxy is usually three times larger than any other cluster member;
- **Bs (binary)** have a dual system in the core, with two giant and bright galaxies;
- **Ls (line)** have more than three galaxies (among the ten most luminous) nearly arranged in a line;
- **Cs (core-halo)**: have at least four galaxies, dominant in luminosity, located in the cluster inner region and forming a single core;
- **Fs (flat)** have most of the brightest galaxies distributed in a flattened configuration;
- **Is (irregular)** do not have a well-defined center and show an irregular galaxy distribution.



Figure 2.2: The Virgo cluster (at a distance of ∼16.5 Mpc), with its irregular galaxy distribution, is the largest cluster and the nearest to our Local Group; its core hosts at least three sub-clusters, centered around the galaxies M87, M86, and M49 (picture by Dr. F. Neyer, Antares Observatory).

RS also find that irregular clusters are the most common and that the B and L types, occurring with the same frequency, are the most unusual (Bahcall 1977a); the remaining types are observed in this order of frequency: cD, F and C.

Clearly, each of the classifications here introduced has its own properties (with limits and advantages), but some similarities exist, i.e., to stress the presence/absence of a dominant central galaxy. Moreover, different kinds of classification systems correlate with respect to many of the properties of galaxy clusters, as well as the morphological ones are related to the evolutionary stage of the cluster. Different classes of the presented classifications are reported in Table 2.2 of Sect. 2.4.2, together with other cluster properties we will introduce in the following sections.

2.2 Dark matter

2.2.1 Characterization and distribution

Although the idea of DM had been suggested in the early 1930s (Zwicky 1933), it was not until the 1970s that it was taken seriously by the broader astronomical community. Gradually, particle physicists realized the potential implications for their discipline as well, and a particular set of DM features was developing. Firstly, it must be really dark, namely with no interactions with photons, and cannot dissipate energy through electromagnetic interactions. Moreover, self-interaction of DM scattering between two DM par-

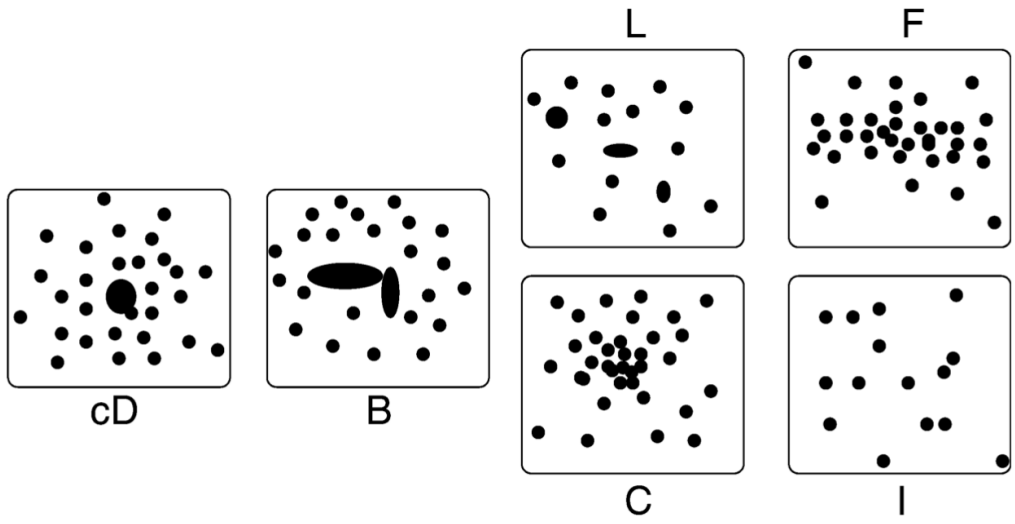


Figure 2.3: Tuning fork classification of galaxy clusters by Rood & Sastry (1971). Image taken from Schneider (2015).

ticles must be small: this is deducible from cosmological simulations which, otherwise, would predict structures with properties different than what we observe. In addition, the so called Bullet Cluster, a system of two colliding clusters of galaxies, is considered the best evidence that DM-DM cross section is small or negligible (Markevitch et al. 2004). Hence, DM differs from a perfect gas, being DM composed of non-collisional particles: the only interaction is gravitational and the only resistance to this force is due to the motion of DM particles, whose energy can be converted only from potential to kinetic. Finally, DM cannot be made by the constituents of the *Standard Model of Particle Physics*, because most leptons (electron, muon, tau) and baryons are charged, while neutral particles are not all viable candidates, since they are not long-lived. The only potential suitable candidate was thought to be the neutrino, neutral and stable, but, by the mid-1980s, this attractive possibility had been excluded: beyond the reasons based on general arguments dealing with theoretical mass constraints and observed properties of dwarf galaxies, Standard Model neutrinos would form structures in a top-down fashion, contradicting the most accredited scenario of cosmic structure formation. As a consequence, the DM existence implies the extension of the *Standard Model of Particle Physics*, unless DM has not a particle nature[8].

Although the baryonic matter in the form of galaxies and hot gas is the most easily observable and distinguishable evidence of a galaxy cluster, as anticipated, it makes up only a fraction of the total cluster mass. On smaller scales, galaxies are hosted by DM halos, which extend to many times the galactic radius and contain many times the stellar mass of the visible galactic region. Even considering the sum of the masses of these DM halos (plus their minor baryonic content), the overall matter budget of clusters is not accounted for. For instance, galaxies and their DM halos contribute only about 10% of

[8] In this scenario, the phenomenology attributed to DM could be explained in the context of other models involving, for example, the modification of gravity theory or the existence of extra dimensions.

the total mass of a cluster core. Results like this came from the early detailed analyses of galaxy clusters, which were possible, as shown first by Kneib et al. (1996), thanks to the refurbishment of the *HST*, which initially experienced an aberrated state (Sect. 4.2.1). In the light of the above, to recover the observed amount of matter in galaxy clusters, these systems must have not only DM halos on small scales (so called *sub-halos*), but they must be also characterized by a diffuse DM component, which accounts for the majority of their mass content.

2.2.2 Comparing simulations and observations

Density profile. In 1995, Navarro, Frenk and White (NFW) performed dissipationless simulations of cluster formation, showing that DM halos have a roughly universal 3D density profile (Navarro et al. 1996)

$$\rho_{\rm NFW}(r) = \frac{4\rho_{\rm s}}{x(1+x)^2} \ , \ x \equiv \frac{r}{r_{\rm s}} \ , \tag{2.12}$$

where $\rho_{\rm s}$ is the characteristic density at $r = r_{\rm s}$ and $r_{\rm s}$ is the scale radius at which the logarithmic slope of the profile $\alpha \equiv {\rm d}\ln\rho/{\rm d}\ln r = -[1+2x/(1+x)]$ is -2. The slope value varies with radius from $\alpha = -1$ at $x \ll 1$ (asymptotic slope) to $\alpha = -3$ at $x \gg 1$. In later simulations (e.g., Graham et al. 2006), the Einasto profile (Einasto 1965) and other similar models seemed to provide a more accurate description of the DM halo density profiles. The Einasto profile is characterized by a power-law variation of the logarithmic slope with radius:

$$\rho_{\rm E}(r) = \rho_{\rm s} \exp\left[\frac{2}{\beta}(1-x^{\beta})\right] \ , \ x \equiv \frac{r}{r_s} \ , \tag{2.13}$$

where $\rho_s \equiv \rho_{\rm E}(r = r_{\rm s})$ and again $r_{\rm s}$ is such that $\alpha = -2$ at $x = 1$; β is an additional parameter describing the power-law dependence of the logarithmic slope on radius: $\alpha = -2x^{\beta}$. There is no asymptotic slope at small radii (unlike the NFW profile) and α is increasingly shallower at $r \to 0$ at the rate controlled by β. Moreover, the total mass corresponding to this profile is finite (contrarily to the NFW profile). Progressively, observational constraints on density profiles of clusters and their concentrations were derived in a number of studies (e.g., Vikhlinin et al. 2006; Sereno & Zitrin 2012). In particular, lensing analyses (e.g., Bartelmann & Meneghetti 2004; Newman et al. 2011) indicated a specific issue between model predictions and observations at the scale of cluster cores (~ 350 kpc): the slope of the inner density profile of some galaxy clusters was found shallower than predicted. This discrepancy is even more intriguing since different values of the slope parameter are predicted depending on the specific DM particle candidate considered. Currently, its precise value is yet intensely debated. In some cases, the analysis of the same cluster by different groups is controversial. For example, Umetsu et al. (2016) find mass density profiles of 20 clusters consistent with the NFW profile, while cluster profiles in Newman et al. (2013a,b) are significantly flatter than NFW ones, once the stellar component mass contribution is subtracted. Possible explanations of this issue have been related to baryonic effects flattening the cluster inner profiles, or to the presence of unaccounted systematics in the observational analyses (e.g., Dalal & Keeton 2003). Solutions to discrepancies such as those described above have been proposed (e.g., Fedeli et al. 2012; Rudd et al. 2008), mostly preventing drastic changes to the ΛCDM scenario of structure formation. However, it is not clear, currently, which solution could be, at the same time, general to preserve the advantages of the standard ΛCDM paradigm and sufficiently specific to solve the majority of its discrepancies with the observations.

Halo shapes. Simulations suggest that the halo shape is tightly correlated with its merger history: low-mass halos are more spherical and, on the contrary, the high-mass ones seem to deviate more from this trend. The reason derives from the later assembly of massive halos, which form after experiencing interactions with the less massive ones, flattening their shapes in the direction along which major mergers occurred (e.g., Vitvitska et al. 2002; Hetznecker & Burkert 2006). However, also the first and less massive halos are expected to deviate from sphericity. This symmetry can be an assumption of spherical collapse models, but, in general, halos are represented by the peaks of a random density field and have a different curvature along different directions. Several authors have fitted simulated DM halos with ellipsoids, characterized by the following dimensionless shape parameters

$$s = \frac{c}{a} \quad , \quad q = \frac{b}{a} \quad , \quad p = \frac{c}{b} \ , \tag{2.14}$$

where $a \geq b \geq c$ are the halo axes. The triaxiality parameter, T, is also used to define an oblate ($T = 0$) or a prolate ($T = 1$) halo

$$T = \frac{a^2 - b^2}{a^2 - c^2} = \frac{1 - q^2}{1 - s^2} \ . \tag{2.15}$$

In simulations, CDM halos typically have different axis values, with $0.5 < T < 0.85$, hence they are triaxial ellipsoids more prolate than oblate, resembling a cigar, with the intermediate axis length usually closer to that of the minor axis. Axis ratios greater than 2 are also common, with major axis aligned along the *los* (Oguri & Blandford 2008; Meneghetti et al. 2010). Towards the center, the CDM halo elongation is even more pronounced, i.e., the axis ratios become smaller while the orientations remain fairly well aligned with the *los*. Finally, Allgood et al. (2006) found that halo shapes have a dependence on the halo mass and redshift, well characterized by the relation

$$s(M, z) = (0.54 \pm 0.03) \left[\frac{M}{M^*(z)} \right]^{-0.050 \pm 0.003} \ , \tag{2.16}$$

where $M^*(z)$ is the characteristic halo mass at redshift z and s is measured at $0.3 r_{\rm vir}$.

Halo angular momentum and spin. The angular momentum growth of proto-halos was described by White (White 1984), in the *linear tidal torque theory*, where DM halos acquire angular momentum (J) in the linear regime due to tidal torques from neighboring overdensities; then, the angular momentum acquisition stops once a proto-halo starts to collapse. The overall behavior of J in this model is consistent with simulations, but its numerical value is overpredicted by a factor of ~ 3. This can be explained by considering that the halo angular momentum continues to evolve due to the accretion of other halos and the merging with them (Maller et al. 2002; Vitvitska et al. 2002). Traditionally, the angular momentum of a DM halo is defined through the dimensionless spin parameter:

$$\lambda = \frac{J |E|^{1/2}}{G M^{5/2}} \ , \tag{2.17}$$

where E and M are the energy and the mass of the halo, respectively. Simulations show a median spin parameter of $\lambda \simeq 0.03$, which is about an order of magnitude

smaller than the value for a typical spiral galaxy. This indicates that DM halos are not rotation-supported structures, so their flattening is not due to rotation, but to the velocity anisotropy. Recent major mergers can increase the spin parameter value, but only for a *short* time before λ stabilizes again to the average, small value (e.g., D'Onghia & Navarro 2007).

Sub-structures. Both observations and numerical simulations of galaxy clusters agree on the presence of hundreds or even thousands of sub-halos with masses that are orders of magnitude lower than that of the smooth DM distribution. These sub-structures form earlier in the cosmic evolution, as a direct consequence of the bottom-up hierarchical structure formation. They also appear qualitatively similar over a wide range in mass (or scale), as a consequence of the power spectrum of density fluctuations (Schneider 2015), thus DM halo properties are scale-invariant. The sub-halo spatial distribution is not uniform since they are less concentrated towards the halo center. This can be explained by the strong tidal forces in the cluster's innermost regions, where sub-halos can be disrupted in the course of evolution. When baryons and gas physics are included in simulations this disruption is less pronounced and more sub-halos survive for a longer time. However, the number of observed halos associated with cluster galaxies is greater than that in simulations and this mismatch is more severe in the cluster core (< 200 kpc). More specifically, simulated clusters have a statistically significant deficiency in more massive sub-structures (Grillo et al. 2015; Munari et al. 2016).

2.3 Intracluster medium

2.3.1 Discovery and applications

The discovery of the diffuse hot ICM was possible after the launch, in 1970, of the *UHURU* X-ray satellite, which detected its high-energy X–ray radiation in massive galaxy clusters, such as the Coma cluster (e.g., Cavaliere et al. 1971). The same kind of emission was then found in low mass clusters and galaxy groups, thanks to observations made with the *Einstein* and *ROSAT* X-ray satellites. The latter allowed the identification of several galaxy clusters up to $z < 1$ (e.g., Burenin et al. 2007). Since the total X-ray luminosity of the extended ICM can reach up to $10^{43} - 10^{45}$ erg s^{-1}, it is not surprising that the first observations in this band were also devoted to detect clusters. The most recent *Chandra* and *XMM-Newton* X-ray observatories have widely confirmed ICM observations, have made X-ray measurements routine and have also allowed observations of galaxy clusters at higher ($z > 1$) redshifts (e.g., Rosati et al. 2002). The search for galaxy clusters continues with these new generation instruments, whose angular resolution is comparable to that of ground-based optical images, e.g., the *Chandra Observatory* has roughly one arcsecond resolution. A relevant project aimed to collect a statistically complete sample of X-ray luminous galaxy clusters is the MAssive Cluster Survey (MACS). It selected distant candidates from the *ROSAT Bright Source Catalogue*, spectroscopically confirming a sample of 124 clusters at $0.3 < z < 0.7$ (more than 80 were new discoveries).

X-ray analyses of galaxy clusters are revealing themselves crucial to shed light on complex aspects of different astrophysical and cosmological issues. The observed thermodynamic properties of the ICM can be confronted with the predictions of baseline models, like the popular self-similar model of clusters, developed by Kaiser (Kaiser 1986); by this comparison, if the chosen model is simple, without cooling and heating, it is possible to isolate the impact of these physical processes in the formation history

of a cluster. Information on the present dynamical status of clusters (phases of pre- or post-merging) can also be derived from X-ray data, measuring the position of the ICM X-ray peak (Ogrean et al. 2015).

From a cosmology perspective, ICM studies can indicate the effective contribution of non-gravitational processes, constraining the theories of large-scale structure formation. The ICM is also a record of the nucleosynthesis in the universe since measurements of both the elemental abundances and their evolution provide fundamental data to trace the origin of the elements. The ICM has even played a role in better defining the "missing mass" problem in galaxy clusters, as theoretically conjectured by Limber (1959) and van Albada (1960). In fact, currently we know that the total mass of the ICM is too low to account for the "missing mass", although this gas component accounts for the most significant portion of the baryonic material, surpassing that of cluster stars. Furthermore, the high ICM temperature provided an independent way to confirm that the cluster gravitational potential requires a predominant dark component.

2.3.2 ICM features

In this section, we briefly describe how the ICM is created, how it is distributed across galaxy clusters and its most relevant properties derived (mainly) by X-ray observations. We also present some of the several correlations between optical and X-ray quantities, found in galaxy clusters.

ICM origin and distribution. The collapse of the large-scale density peaks during the initial phase of cluster formation affects not only the characteristics of the DM halo of a cluster, but also the thermodynamic properties of its ICM. When the dynamically dominant DM captures baryonic matter, the latter enters the DM halo and during its collapse (and merging) acquires a supersonic motion which converts the gas kinetic energy into thermal energy via adiabatic compressions and shocks. This is only the primordial component of the ICM, which is never captured by cluster galaxies and have a cosmological chemical abundance. The ICM mixture also includes a metal-enriched gas, ejected later from galaxies by stellar processes (supernova explosions and stellar winds) and by violent phenomena, such as galaxy stripping (see Figs. 2.4, 2.5 and their captions). The final result of these processes is a diffuse, hot gas distribution, which extends between 200 and 3000 kpc and is not locked in individual galaxies, as shown from X-ray observations (Fig. 2.4). For regular clusters, the morphology of the X-ray emission has a central maximum (usually coincident with the BCG) and a flat inner distribution, followed by a decrease in the regions out of this core. Irregular clusters may show several X-ray peaks, typically centered on galaxy sub-structures.

ICM composition and X-ray emission. Due to the deep potential well of a cluster, the shock-heated ICM reaches very high temperatures of $T_{ICM} = 10^7 - 10^8$ K (or energies of $\sim 1 - 10$ keV) as it enters the DM halo. Thus, it becomes fully ionized and emits copious X-ray radiation via thermal bremsstrahlung. More precisely, the emission is non-variable and comes from a rarefied gas of free electrons: they emits because of their acceleration via Coulomb interactions with a mixture of protons (released by the predominant hydrogen) and ions (the nuclei of helium and heavier elements). Depending on the precise temperature of the ICM, the heaviest atoms can be only partially ionized, hence they produce emission lines superimposed on the continuum bremsstrahlung spectrum. Clearly, this is valid especially for less massive clusters, whose ICM energies are below ~ 2 keV.

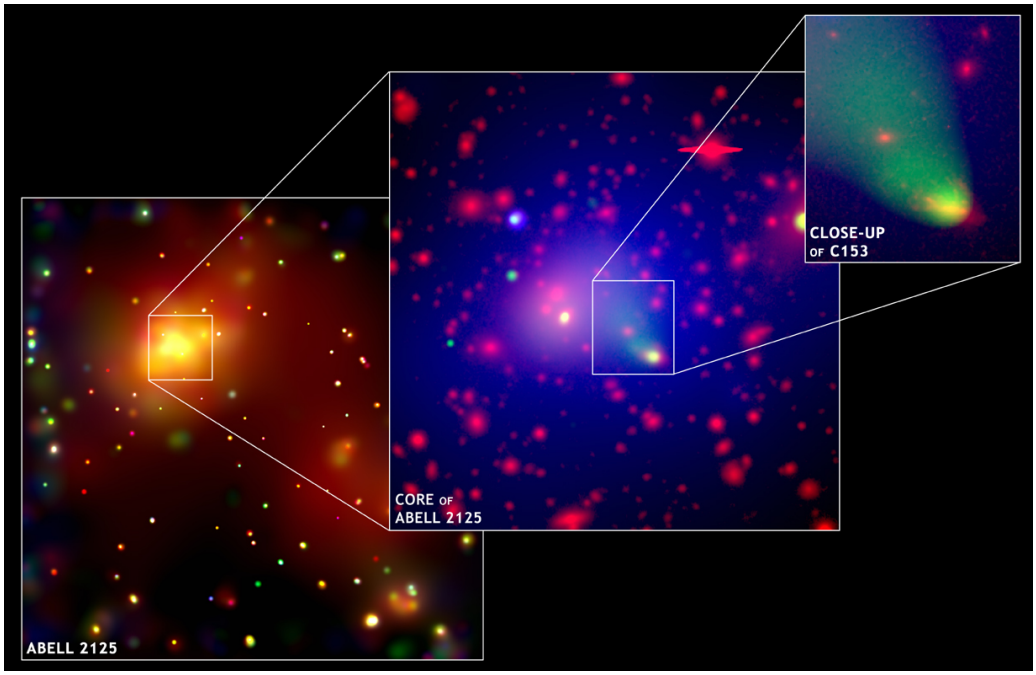


Figure 2.4: Three-panel image of Abell 2125. Left: a *Chandra*'s X-ray view of the entire galaxy cluster, where red, green, and blue represent, respectively, low, medium, and high-energy X-ray emissions. Middle: Abell 2125 core, where the X-ray emitting gas distribution is in green and blue, while the red color refers to the optical emission. Right: Galaxy C153 in visible light (red) from *HST* and X-ray emission (green and blue) observed by *Chandra* (Credit: X-ray: NASA/CXC/UMass; Optical: NASA/STScI and NOAO/Kitt Peak).

However, heavier elements represent, on average, a fraction of one per cent; they have typical metallicities which amount to one-third of the solar abundance and are concentrated, in an inhomogeneous way, in the inner cluster regions (especially in merging systems, Werner et al. 2008).

Correlations. Although the diffuse plasma of the ICM is not directly associated with cluster galaxies, their properties are correlated (e.g., Lin et al. 2012). In general, different scaling relations with a small scatter exist, both between the ICM observable integrated quantities and between observable quantities and the total mass. These correlations are a key ingredient especially in cosmological applications of galaxy clusters. In particular, one of them involves the X-band luminosity, L_X, in the energy range $[1-10]$ keV, which is related strongly with $\sigma_{\rm los}$, the *los* velocity dispersion of cluster galaxies (Quintana & Melnick 1982):

$$L_{\rm X[1-10]keV} = 4.2 \times 10^{44}\ {\rm erg\ s^{-1}} \left(\frac{\sigma_{\rm los}}{1000\ {\rm km\ s^{-1}}} \right)^4 . \tag{2.18}$$

L_X is also correlated to the kind of galactic content. The most X-ray luminous clusters are characterized by a small fraction of spiral galaxies ($f_{\rm sp}$), whose concentration and number, respectively, increases in the outer regions and decreases with the velocity dis-

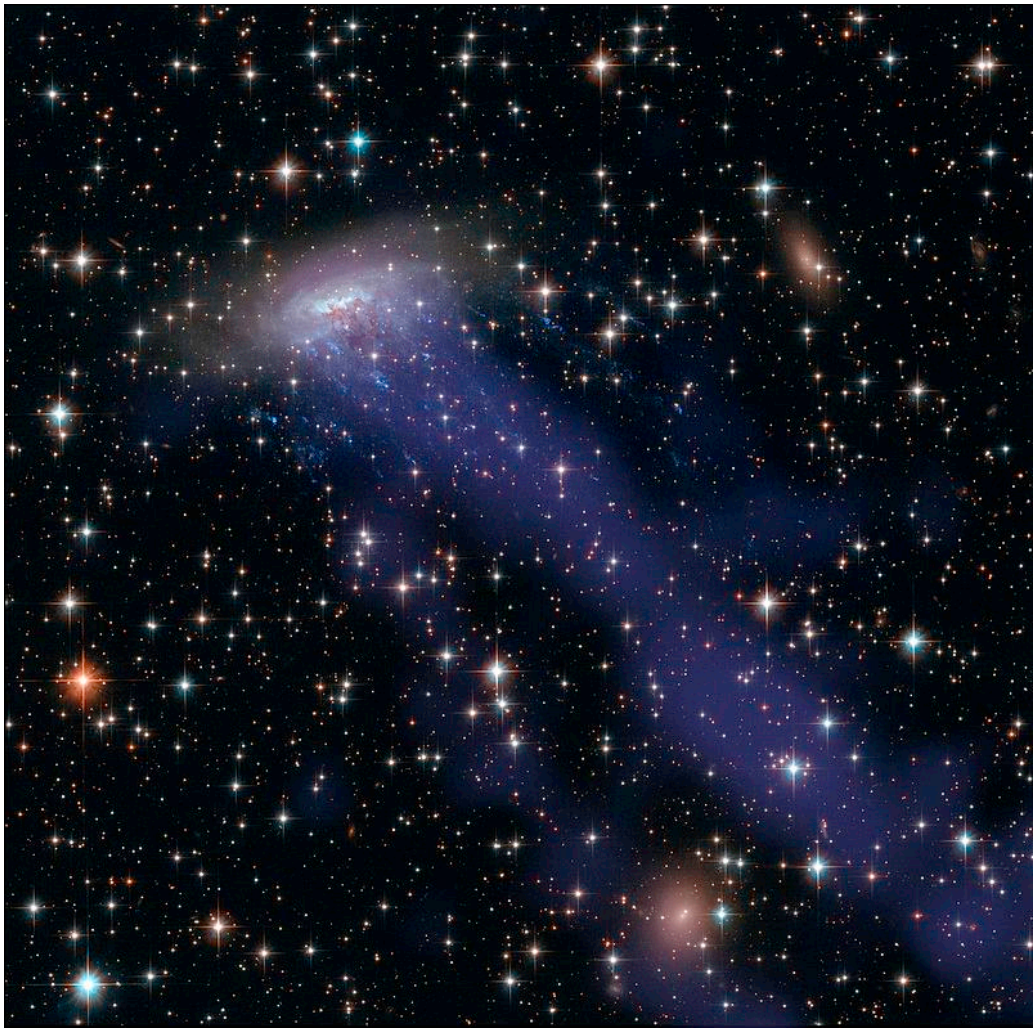


Figure 2.5: The spiral galaxy ESO 137-001, in the Norma cluster located at a distance of about 68 Mpc. This image combines NASA/ESA *HST* observations with data from the *Chandra* X-ray Observatory (in blue). The gas trail of this spiral (also known as the *jellyfish galaxy*) is a remarkable example of a process known as *ram pressure stripping*: high-velocity galaxies traveling through the hot ICM lose their gas, preserving their stellar content, which instead remains intact due to the binding force of their higher gravity.

persion. The spiral fraction and the X-ray luminosity are related (Bahcall 1977c) by

$$f_{\rm sp} \simeq 0.37 - 0.26 \, \log \left[\frac{L_X}{10^{44} \text{ erg s}^{-1}} \right] \quad . \tag{2.19}$$

This relation is consistent with the interpretation of cluster S0 galaxies being a by-product of the gas stripping of spiral galaxies due to ICM interactions (however, this theory has not been proved). Other correlations between the optical and X-ray properties of clusters can be found in Bahcall (1977a, b).

2.3.3 Cluster mass estimate from X-ray data

One of the applications of ICM studies (of our interest) consists in estimating the total mass of a cluster from X-ray data. This method relies on the assumptions of hydrostatic equilibrium of the emitting gas and the spherical symmetry of its spatial distribution. The following equation describes how the force of the gas pressure (P) is balanced by the gravitational force (corresponding to the potential Φ):

$$\frac{dP}{dr} = -\frac{d\Phi}{dr}\rho_{\rm g}(r) = -\frac{GM(<r)}{r^2}\rho_{\rm g}(r) \ , \tag{2.20}$$

where $M(<r)$ is the total enclosed mass (i.e., not just the gas mass) and $\rho_{\rm g}(r)$ describes the variation of the gas density with the radius. We also assume an ideal cluster gas and use the equation of state

$$P = n{\rm k}T = \frac{\rho_{\rm g}{\rm k}T}{\mu m_{\rm p}} \ , \tag{2.21}$$

where n is the number density and μ is the mean molecular weight, defined as the average mass of a gas particle in units of the proton mass, $m_{\rm p}$. By inserting Eq. (2.21) in (2.20), we obtain the X-ray mass estimate:

$$M(<r) = -\frac{{\rm k}Tr}{G\mu m_{\rm p}}\left(\frac{d\log\rho_{\rm g}}{d\log r} + \frac{d\log T}{d\log r}\right) \ . \tag{2.22}$$

Thus, the total mass profile can be derived from the radial profiles of both the density and the temperature, which are provided by the observations of the X-ray luminosity and the *spectral temperature* (i.e., temperature through spectral fitting). In order to obtain the mass from these observables, two techniques are usually applied: the *forward* and the *backward* method (a comparison between them can be found in Meneghetti et al. 2010). In the former (e.g., Vikhlinin et al. 2006, Pratt & Arnaud 2002) a parametric form of the gas-density profile is considered; it is usually the β-model (Cavaliere & Fusco-Femiano 1978), in which the density profile of the total matter (dark and luminous) is described by an isothermal distribution (or its generalizations, e.g., Pratt & Arnaud 2002, and references therein):

$$\rho_{\rm g}(r) = \rho_{\rm g,0}\left[1 + \left(\frac{r}{r_{\rm c}}\right)^2\right]^{-3\beta/2} \ . \tag{2.23}$$

Here, $\rho_{\rm g,0}$ is the central gas density and $r_{\rm c}$ is the core radius; the index β depends on the ratio of the dynamical temperature, measured by the velocity dispersion $\sigma_{\rm los}$, and the gas temperature

$$\beta \equiv \frac{\mu m_{\rm p}\sigma_{\rm los}^2}{{\rm k}T} \ , \tag{2.24}$$

where usually $\mu = 0.63$, assuming solar abundances. The temperature profile has also a parametric form (Meneghetti et al. 2009):

$$T(r) = T_0\frac{(r/r_{\rm t})^{-a}}{[1 + (r/r_{\rm t})^b]^{c/b}} \ , \tag{2.25}$$

where, T_0, a, b, and c correspond to fit parameters[9],[10].

[9] Note that a, b and c, despite the notation, are not the halo shape axes introduced in Eq. (2.14).

[10] In Allen et al. (2001) and Vikhlinin et al. (2006), a more general expression is considered: the right-hand side of Eq. (2.25) is multiplied by a further term, $T_{\rm cool}$, which is dimensionless. It takes into account the possible presence in galaxy clusters of a "cool core", namely a temperature profile falling towards the center.

The density and the temperature profiles are then projected along the *los* (Mazzotta et al. 2004) and, finally, fitted to the observations. Instead, in the backward approach (Ettori et al. 2002; Morandi et al. 2007) the observed quantities are deprojected and compared to the predictions of an analytical three-dimensional mass profile of the cluster. However, numerical simulations suggest that both the hydrostatic equilibrium and the spherical symmetry are too strong assumptions, which do not describe well the shape and the dynamical state of galaxy clusters; for example, the ICM can be in hydrostatic equilibrium only provided that a sufficient time has passed from the last major merging event or no relevant and continuous heat source exists in the cluster core, like an active galactic nucleus (AGN) hosted by the BCG. This could explain why the mass estimates obtained with these methods are sometimes different by a factor of two compared to those from other techniques, like gravitational lensing (Morandi et al. 2010, and references therein).

2.4 Galaxies

2.4.1 Galactic content in cluster cores

The morphological classification of clusters presented in Sect. 2.1.4 describes regular, compact clusters as structures dominated in their core by a giant cD galaxy or a pair of central luminous galaxies. About 20% of all rich clusters contain these kinds of ellipticals[11], which, together with the surrounding bright galaxies in the cluster core, make up a significant fraction of the entire stellar mass of the cluster and of its visible light. More precisely, excluding the presence of peculiar sources in the galactic nucleus like AGNs, cD galaxies are the most luminous galaxies known, about 60% brighter than common giant ellipticals (Malumuth & Kirshner 1981). cD galaxies are also more extended than the typical giant ellipticals and have a core region with effective radii roughly a factor of two or three larger than those of other elliptical galaxies. The dynamical mass measurement of cD galaxies are typically $\sim 10^{13}\,\mathrm{M}_{\odot}$ and their velocity equals the mean velocity of cluster galaxies (Quintana & Lawrie 1982). The kinematic properties of these peculiar galaxies suggest that they have been either formed or modified by cluster dynamical processes. Gallagher & Ostriker (1972) have proposed that cDs represent the debris from galaxy collisions. On the other hand, their extended nucleus can be explained on the basis of their formation history. Ostriker & Tremaine (1975) were the first ones to suggest that different massive galaxies are guided by dynamical friction towards the cluster core, where they merge and form a single supergiant galaxy. The latter can subsequently become more extended and luminous, swallowing other galaxies which pass through the cluster center, in a process known as *galactic cannibalism* (Ostriker & Hausman 1977). This also leaves the resulting cD galaxy close to the center of the cluster and nearly at rest relative to the other cluster galaxies. Finally, a further confirmation of this phenomenon comes from the presence in many cD galaxies of multiple cores, which are rarely observed in other cluster members.

2.4.2 Overall galactic component

As shown in Table 2.2, adapted from Abell (1975) and Bahcall (1977a), galaxy clusters show a systematic variation in their galactic content: in regular compact clusters, elliptical and S0 galaxies predominate over spirals, while the opposite is true both in irregular clusters and in the field (e.g., Oemler 1974; Capak et al. 2007). More precisely, $\sim 60\%$ of

[11] Nevertheless, some of them have been also found in poor clusters and groups (Morgan & White 1975).

Table 2.2: Comparison between clusters in different classification schemes.

Property/Class	Regular (early)	Intermediate	Irregular (late)
Zwicky type	Compact	Medium-Compact	Open
Rood-Sastry type	cD, B, L, C	L, C, F	F, I
Richness	Rich	Rich-Moderate	Rich-Poor
Symmetry	Spherical	Intermediate	Irregular shape
Content	Elliptical-rich	Spiral-poor	Spiral-rich
E:S0:S ratio	3:4:2	1:4:2	1:2:3
Central concentration	High	Moderate	Low
Sub-clustering	Absent	Moderate	Significant
X-ray luminosity	High	Intermediate	Low
Examples	A2199, Coma	A194, A539	Virgo, A1228

the bright galaxies within the virial radius are bulge-dominated compared to $\sim 30\%$ of galaxies with similar luminosity present in very low-density environments (Whitmore & Gilmore 1991; Postman et al. 2005). The question naturally arises as to whether red, early-type galaxies represent a *special* component of clusters.

Many explanations of the over-abundance of early-type galaxies in clusters have been proposed and, in general, they fall into two broad classes. In the former, *pre-processing* phenomena are responsible for the formation of these galaxies in the environment prior to cluster formation or, anyway, these processes are important in determining the early-type fraction in clusters (e.g., Zabludoff & Mulchaey 1998). In the latter, galaxy morphology evolves in response to the cluster environment after its formation and, thus, this change is related to the overall picture of galaxy evolution within clusters, detailed, for their inner regions, in section 2.4.1. In particular, cluster galaxies experience more mergers than field galaxies and this explains the differences between these two populations (Toomre & Toomre 1972). The fact that the environment density plays a major role is also suggested by two observational findings. The first is that the overall mix of galaxies in clusters evolves with redshift (e.g., Capak et al. 2007) and this means, for example, that the late-type galaxy fraction is larger for higher redshifts. The second is that, at all redshifts, early-type galaxies in clusters are typically older than field galaxies by $0.5-1.5$ Gyr (Clemens et al. 2006), suggesting that galaxy evolution was somehow boosted by the cluster environment. The excess of blue color galaxies in nearby (low redshift) clusters compared to similar systems at high redshift was first recognized in 1978 (Butcher & Oemler 1978) and is now known as the *Butcher-Oemler effect*. Their initial measurements were then refined, leading to the conclusion that the blue galaxy fraction in rich clusters rises from $\sim 3\%$ for nearby ($z < 0.1$) clusters, to 25% at $z \sim 0.5$ and reaching 70% by $z \sim 1$. This fraction also depends on the number of galaxies that the cluster contains, the galaxy magnitude limit adopted and the distance from the cluster center.

Regarding the stronger lack of blue galaxies in a cluster core, several phenomena have been considered to explain the possible transformation of star-forming spirals into red, dead ellipticals. However, it is still unclear how these processes interplay, when they act and which one is dominant. Beyond ram-pressure stripping (Gunn & Gott 1972) and galaxy mergers (Murante et al. 2007), presented in Sect. 2.4.1, other mechanisms are harrassment (Farouki & Shapiro 1981), i.e., rapid tidal encounters, strangulation (McCarthy et al. 2008), i.e., loss of the hot halo, and AGN feedback (Nesvadba et al. 2006).

2.5 Minor components

The whole baryonic content of galaxy clusters, primarily in the form of ICM and stars, is then completely accounted for considering two minor mass components: the intracluster light (ICL), an inter-galactic diffuse (optical) radiation due to the presence of stars in between the cluster members, and the dust. Here we briefly describe the characteristics of these two further mass components and why they were both unexpected, at first.

ICL. Outside the cluster center, the detection of the ICL is extremely difficult, due to its low surface brightness of ~ 30 mag arcsec^{-2} at a distance of several hundred kpc from the core center. For ground-based observations, corrections are required to eliminate the contribution by the night sky brightness (~ 21 mag arcsec^{-2}, in the V-band). In particular, for the ICL to become visible the correction must be better than 0.1% (Schneider 2015). Furthermore, the galaxies of the cluster and possible interlopers must be both masked out in the images. To appreciate distinctively this diffuse light component in galaxy clusters, we refer to Fig. 2.6, taken from an animation on the *HST* website[12], in which the ICL of the cluster MACS J0416.1–2403 is highlighted compared to the original *HST* image. Although observed in clusters up to intermediate redshifts, the ICL is best studied in nearby groups and clusters, for which individual objects can be detected. It was found that the ICL accounts for about 10% of the total optical cluster light, but for few dense clusters it can be even three time higher. This mass component is thus not negligible and this is unexpected, as well as its own existence: in the star formation theory, stars only originate in the inner, dense core of molecular clouds, hence they cannot be formed in the intergalactic medium, where, instead, the stars of the ICL are observed. The way in which the ICL stars become an intergalactic population is closely related to processes responsible for galaxy morphological transformation, tidal stripping, and galaxy disruption. During gravitational interactions between galaxies, stars can be stripped and tidal interactions can act likewise on the outermost galaxy stars; clearly, the latter cannot be subject to ram pressure stripping, in contrast to the galaxy gas. A diffuse intergalactic light is also observed in (compact) groups where the strength of tidal interactions is stronger than in clusters (Schneider 2015). What we can learn from the ICL about cluster formation and evolution will probably come from cosmological N-body and hydrodynamical simulations; they are now considering the kinematics and the origin of the ICL and are beginning to make predictions.

Cluster dust. More unexpected than the ICL due to the high temperatures in clusters, the dust in these systems not only exists, but it has been recently considered in galaxy cluster analyses. Dust grains absorb and redden radiation from background sources, hence an accurate knowledge of the dust large-scale distribution is relevant. However, beyond the pure correction of astrophysical catalogs for the effects of extinction, phenomena related to dust have found several other applications, although observations show that this component is negligible in terms of mass. Dust is expected to be stripped from the galaxy interstellar matter and then totally or partially destroyed by the hostile environment in which the hot ICM is distributed: dust grains *sputter* and eject their atoms due to the collisions with sufficiently energetic gas particles. Dust is not only present in the cluster periphery, but it can be observed also in inner regions, where, depending on the density and temperature of the environment, it sputters on time-scales of $10^7 - 10^9$ yr (Draine & Salpeter 1979). This is a short time implying that dust must have

[12] https://www.spacetelescope.org/videos/heic1820a/



Figure 2.6: Image of the galaxy cluster MACS J0416.1–2403 (observed by the HFF team), in which the ICL is highlighted in blue. Credit: ESA/Hubble, NASA, *HST* Frontier Fields team (STScI), and M. Montes and I. Trujillo.

been accreted recently by the cluster, but it is also a long time for dust to be observed. Clearly, larger dust grains are favored since they survive longer. This peculiar composition of galaxy cluster dust could imply a different reddening law compared to that commonly found for the galactic interstellar material (McGee & Balogh 2010). In fact, the extinction coefficient depends on the properties of the dust, such as the chemical composition and, indeed, the size distribution of the dust grains (Schneider 2015). The predicted reddening law can be tested in gravitational lensing systems, since different colors of the multiple images of the same background source indicate dust absorption. Another interesting application is to use dust analysis and gravitational lenses to study the interstellar medium of background (high-redshift) galaxies (Malhotra et al. 1997). In addition, it is interesting to assess the impact of cluster dust emission on the results of studies in which this contribution was not considered. Using observations at millimeter wavelengths has allowed Melin et al. (2018) to constrain the spatial profile of the dust emission. Comparing gas plus dust and gas-only simulations, they have revisited *Planck* cluster cosmology results. They have shown that dust emission has a negligible effect on the recovery of individual cluster parameters for the *Planck* mission, for instance, it biases *Planck* cluster flux measurements at only the $1-2\%$ level. Nevertheless, dust emission affects the completeness of the cluster catalog: it causes a reduction in the number of cluster detections, for redshifts in the range $[0.3-0.8]$, with a maximum loss of $\sim 9\%$ of clusters between $z=0.5$ and $z=0.8$. However, correcting for this incompleteness has a negligible effect on *Planck* cosmological parameter measurements.

CHAPTER 3

Our strong gravitational lens: Abell 2163

We focus on the specific galaxy cluster which is the main subject of this thesis, A2163. This massive system is characterized by several mass components, which can be identified using the multi-wavelength observations described in Sect. 3.1. The characteristics of these matter sub-clumps are then considered in Sect. 3.2 to discuss a possible merging scenario for this cluster, on both small and large scales.

3.1 Multi-band cluster overview

Located at $z \simeq 0.201$, A2163 is the most massive galaxy cluster of the RELICS Survey (see Sect. 4.1.3 and Fig. 4.2), with a mass $M_{500} \simeq 1.6 \times 10^{15}\,\mathrm{M}_{\odot}$, as estimated from the *Planck* collaboration (Ade et al. 2016b) and reported in Cerny et al. (2018), hereafter C18. We show in Fig. 3.1 a schematic representation of the A2163 sub-structures, observed in different bands and discussed in the paragraphs below.

Maurogordato et al. (2008) conducted an in-depth optical analysis of A2163, whose central region is shown in the *HST* image of Fig. 3.2. In this thesis, we will use the same names for its sub-components. In the main central cluster, A2163-A, galaxies are distributed in two main clumps: A2163-A1 (R.A. 16:15:47.8, Decl. $-$06:08:11) in the north-east and A2163-A2 (R.A. 16:15:42.5, Decl. $-$06:09:53) in the south-west. Also weak lensing analyses (Okabe et al. 2011; Soucail 2012) found evidence of a bimodal total mass distribution in the cluster core. At larger distances from the main mass clumps, there are more sub-structures (see Fig. 3.1): beyond the north-east radio relic (discussed below), the most significant one, A2163-B (R.A. 16:15:48.8, Decl. $-$06:02:21) is in the north, nearly coincident with a secondary X-ray peak.

A2163 is among the most luminous clusters in X-rays ($L_{X\,[2-10\,\mathrm{keV}]} = 6.0 \times 10^{45}$ erg s^{-1}, Elbaz et al. 1995), with exceptionally high gas temperatures, varying between 11.5 and 14.6 keV (Arnaud et al. 1992; Elbaz et al. 1995; Markevitch et al. 1996; Markevitch & Vikhlinin 2001). The gas distribution is non-isothermal, with a high temperature gradient in the center and a strong temperature decline in the outer regions (Markevitch et al. 1994; Govoni et al. 2004; Ota et al. 2014); its generally complex gas distribution shows features similar to those observed in the Bullet cluster (e.g., Soucail 2012), with a central X-ray peak between the two optical sub-clumps of the cluster core. We report in Fig. 3.5 the temperature profiles for A2163 core sub-clusters: the behavior of the two curves is similar within 6′, but clear deviations are found farther out. In particular, the SW cluster shows lower temperatures, in general, and the presence of more discontinuities (see Sect. 3.2 for more details and the interpretation of these temperature profile features).

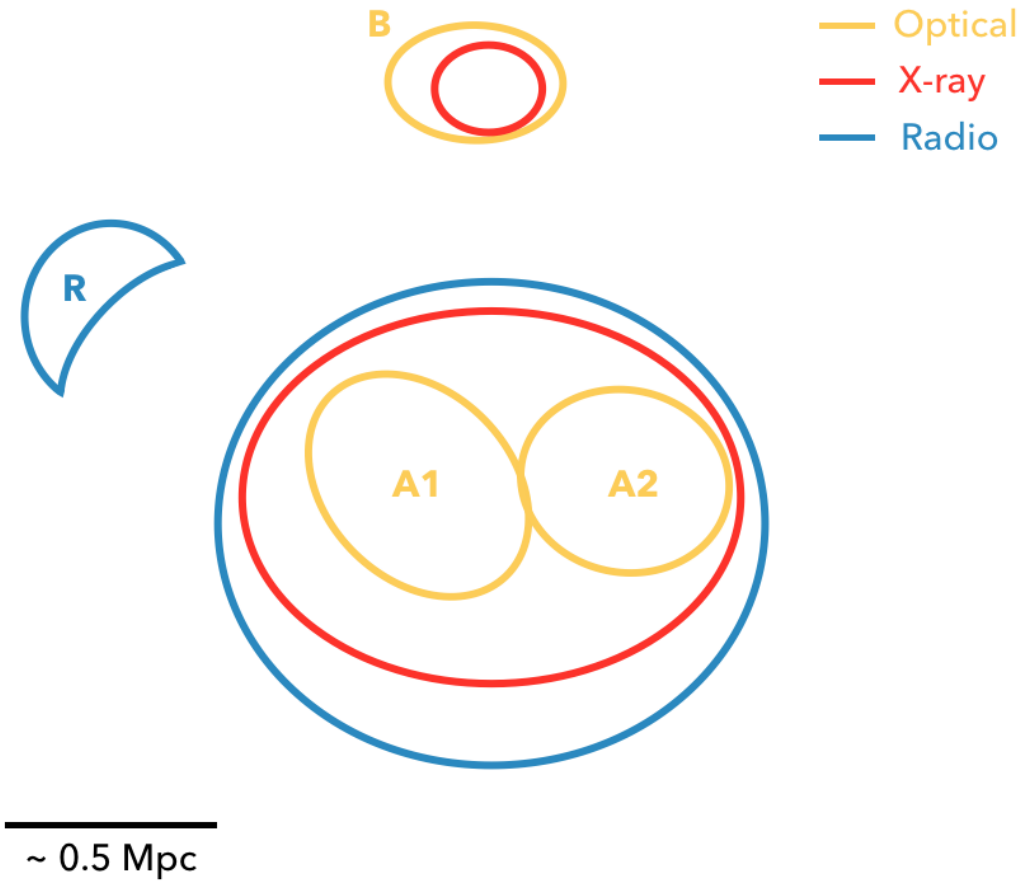


Figure 3.1: Spatial distribution of the A2163 sub-structures, as observed in the three bands reported in the upper-right legend. The optical mass clumps (A1, A2 and B) and the radio relic (R) are described in the text.

X-ray surface brightness is difficult to observe in the outskirts of clusters. However, observations in the microwave can be nicely complementary to X-ray data (and are often combined with them), since the signal of the former remains substantial also in the cluster peripheral area. In detail, the external gas in clusters can be observed through its effects on the cosmic microwave background (CMB). The very young Universe consisted of a compact hot plasma, namely a mix of particles and photons, which were constantly scattered by electrons and protons in the plasma. The Universe expansion decoupled photons by the matter and the former freely traveled through space, becoming less energetic (due to the expansion of space) and generating an isotropic radiation: it is known as CMB, because of its current temperature of about 2.75 K. The CMB photons passing through a cluster interact with the high energy electrons of the ICM via Compton scattering. As a result, there is a measurable difference between these boosted photons and the original ones. The CMB has a nearly perfect thermal blackbody spectrum, whereas scattered photons show spectrum deviations. This process generating CMB map distortions is called the Sunyaev-Zel'dovich (SZ) effect (Sunyaev & Zel'dovich 1970).



Figure 3.2: Optical image of A2163, taken by *HST* ACS and WFC3 cameras, partially for the RELICS observing program. The image FoV is 6.26×3.10 arcminutes and the North is $56^{\circ}.1$ left of horizontal. The list of the 7 filters used and further information can be found on the *HST* website, at the link `https://www.spacetelescope.org/images/potw1750a/`.

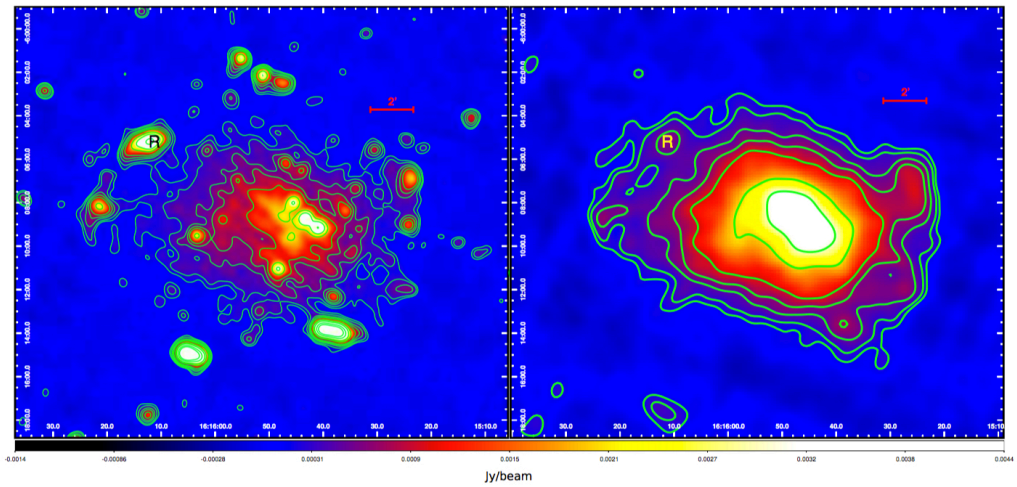


Figure 3.3: Radio maps of A2163 at 20 cm from Thölken et al. (2018), where the letter R marks the position of the radio relic, as identified by Feretti et al. (2001), and the contour levels are 0.2, 0.3, 0.5, 0.7, 1, 2, 3, 5, 7 mJy/beam. The right part of the figure shows the extended emission after the removal of compact sources which instead are present in the image on the left.

Nord et al. (2009) presented A2163 maps based on the SZ effect at two frequencies, 150 and 345 GHz, from observations with the APEX-SZ bolometer and the LABOCA bolometer camera, respectively. In combination with archival *XMM-Newton* X-ray data, SZ measurements were used both to model the radial density and temperature distributions of the intracluster medium and to obtain the total mass profile and the gas-to-total mass fraction. Under the assumption of hydrostatic equilibrium and spherical symmetry, the calculated total mass was $M(r < 100 \text{ kpc}) = 4 - 6 \times 10^{13} \text{ M}_{\odot}$ and the gas-to-total mass ratio (enclosed within 300 kpc) was $\sim 0.10 - 0.15$.

Finally, as first reported in Herbig & Birkinshaw (1994), A2163 radio emission reveals one of the most powerful and extended halos ever detected (see Fig. 3.3). This structure, centered in the cluster core, has a quite regular shape and an elongation in the E-W direction, similar to the X-ray emission (see Fig. 3.1). An elongated, diffuse source, interpreted as a potential radio relic (Feretti et al. 2001), is also present in the N-E region (see Fig. 3.3).

3.2 Merging history

The data from several observational windows presented in Sect. 3.1 suggest a complex mass structure for A2163, with interacting sub-systems, which make the reconstruction of a realistic merging scenario very difficult. Studies from spectroscopic and imaging data (Maurogordato et al. 2008) confirm that A2163-B is part of the cluster and that it is dynamically separated from A2163-A. Nevertheless, between A2163-A and A2163-B, a *bridge* of faint galaxies (along the N-S axis) suggests that the latter is probably infalling into the cluster core (Maurogordato et al. 2008). As an alternative, it can have a tidal origin: A2163-B has already undergone a collision with A2163-A, but at a peripheral region of the latter, so that the gas core of A2163-B would have been preserved, but the tidal in-

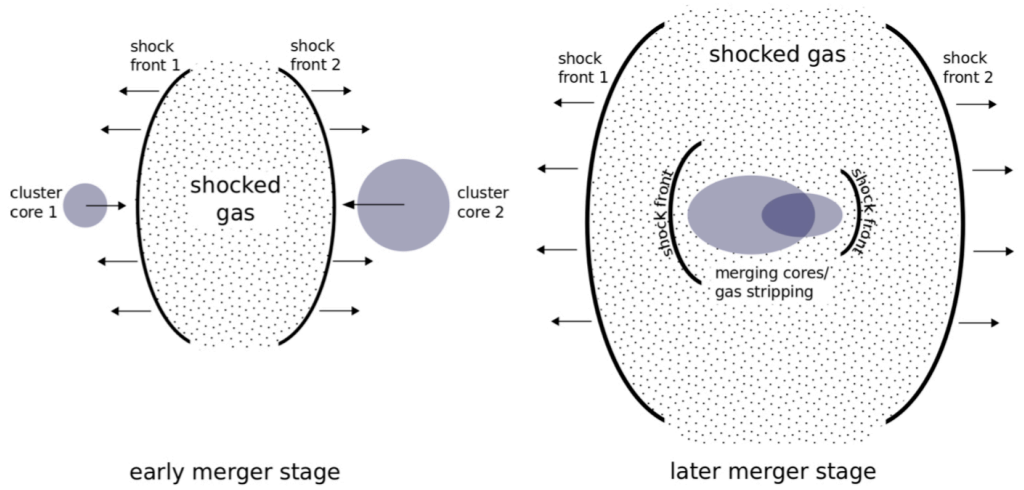


Figure 3.4: Formation of shock fronts in a merger event from Thölken et al. (2018). In an early stage of the merger (left), shock fronts develop after the shock heating of the gas from the two clusters; successively, the cluster core merger can create further shock fronts, while the initial shock waves travel outwards.

teraction would have left the N–S galaxy bridge. On the other hand, in the main cluster, A2163-A, the two clumps A2163-A1 and A2163-A2 can be regarded as a pair of colliding structures that will eventually merge into a bigger one. The best evidence for that is the elongation of the central radio halo, the presence of a main X-ray peak between the two clumps and the detection of shock fronts around the latter (Thölken et al. 2018).

As described in Sect. 2.1.1, the filamentary large scale structure of the universe is characterized by matter knots hosting galaxy super-clusters, whose matter accretion is the result of energetic merging processes of galaxy clusters (and groups). These powerful events dissipate a significant fraction of their energy into shock heating of the ICM (see Sect. 2.3.2 and, e.g., Ha et al. 2018), with shock fronts visible in X-ray observations (Thölken et al. 2018); a schematic representation of shock formation in merging events is given in Fig. 3.4. Shocks can be traced by the presence of discontinuities in the temperature, density and surface brightness profiles. For both sub-clusters in A2163 core, such discontinuities are observed in the temperature profile: specifically, A2163-A1 and A2163-A2 exhibit one discontinuity in common (at about $6'.5$), but the latter shows a second one in the inner region around $3'.3$; this is shown in the temperature profiles of Fig. 3.5, where the shocks are labeled (in the upper panel) with the colored letter "S", in blue for A2163-A2 and in red for A2163-A1; the position of shock fronts is given in Fig. 3.6.

The direction of the merger in A2163-A is confirmed by radio observations and the velocity field. Green contour levels in Fig. 3.6 well match the positions of the X-ray detected shocks and delineate the elongation in E-W direction of the radio emission, which supports the interpretation by Bourdin et al. (2011) and Maurogordato et al. (2008) that the sub-cluster interaction is proceeding along this direction. Moreover, the coexistence in the same cluster of a central radio halo and a peripheral relic, as already observed by

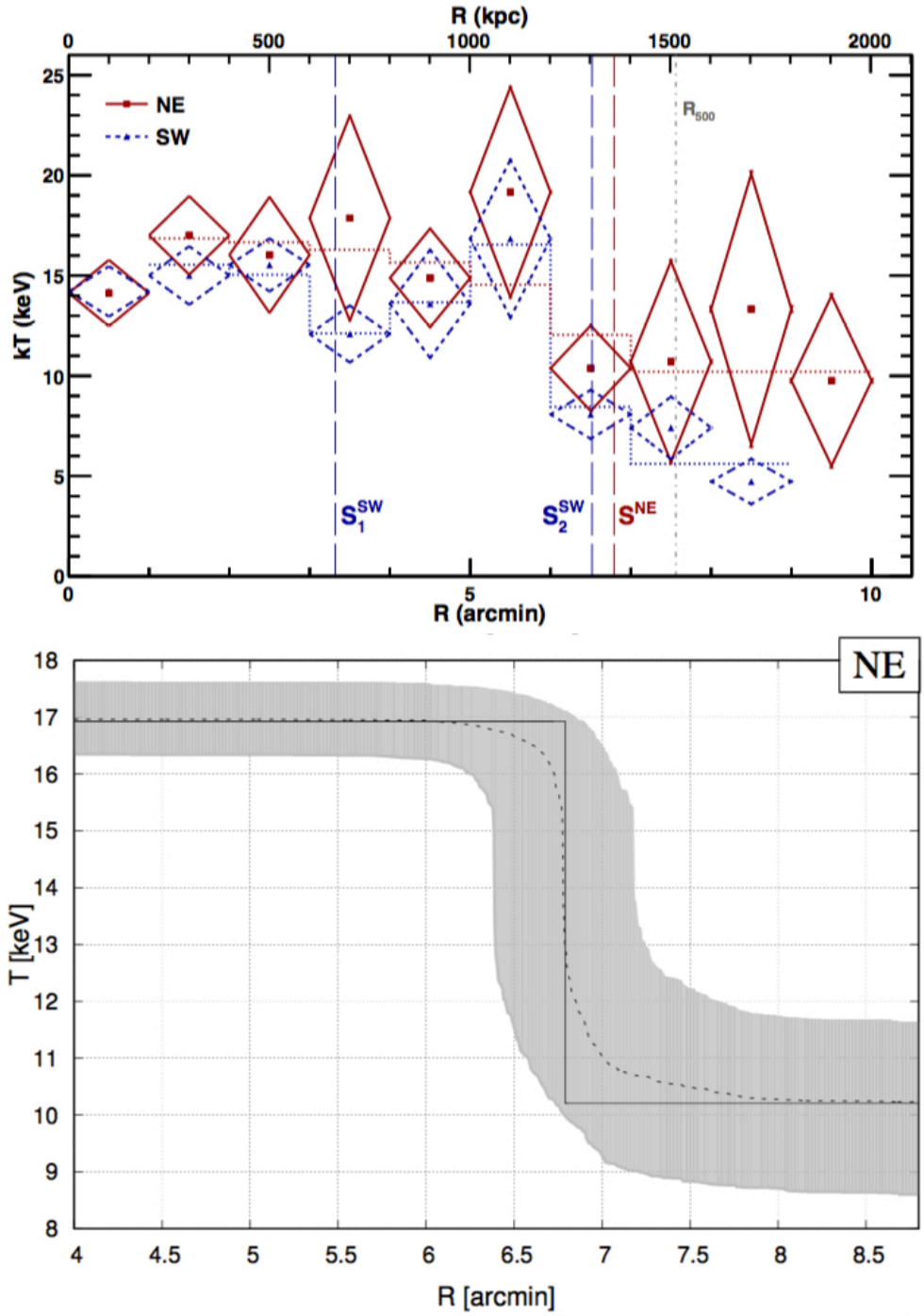


Figure 3.5: Temperature profiles of A2163 core clusters from Thölken et al. (2018). Top panel: projected profiles of the sub-clusters at NE (red diamonds) and SW (blue diamonds); the horizontal red and blue dotted lines are from the best-fit models, while vertical dashed lines indicate the discontinuity positions (S stands for "Schock"). Bottom panel: deprojected temperature profile of A2163-A (solid line) around the shock position; the gray shaded area is the 68% confidence interval from 1000 Markov Chain realizations and the dashed line shows the median of all of them.

Feretti (1999), would suggest a common origin for both structures, to be better confirmed by detailed investigations about the cluster merger state (Feretti et al. 2001). Nevertheless, A2163-A shows a velocity field with a strong gradient which follows the galaxy distribution and is also elongated in the NE/SW direction (Soucail 2012).

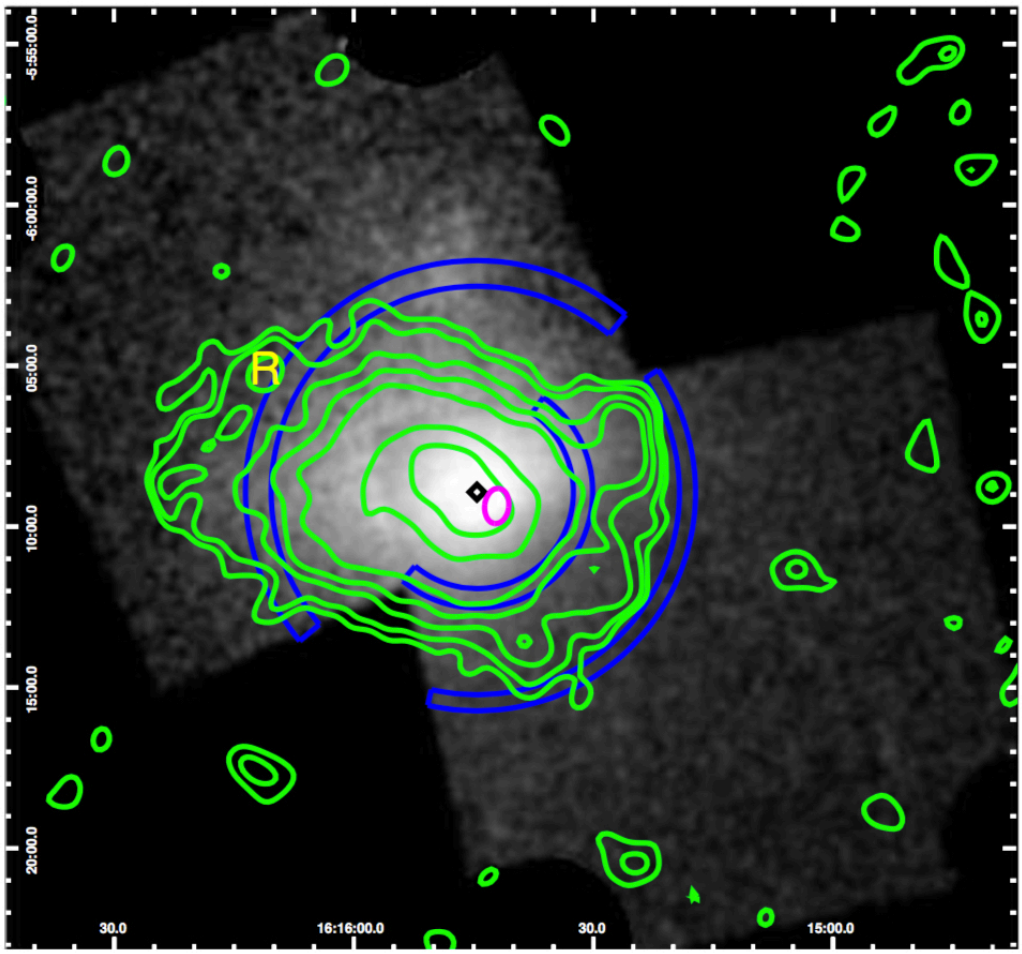


Figure 3.6: Regions of the shock fronts in A2163 (blue), whose widths correspond to the position uncertainties. The black diamond and the magenta ellipse are, respectively, the X-ray peak and the approximate position of the cool core (remnants of the merging constituents). Green contour levels (the same as in the right panel of Fig. 3.3) show the extended radio emission.

The merging processes in A2163 core are also very energetic. The energy activity of the ICM during cluster mergers is quantified by the *Mach number*, M, defined as the ratio between the fluid speed, q, and the sound wave velocity which propagates in the medium, c_s: in explicit terms, $M := q/\sqrt{\gamma kT}$, where γ and k are, respectively, the adiabatic coefficient and a constant which refers to the specific mixture of the gas. Typical values for galaxy cluster ICM are $M \leq 3$ (Markevitch & Vikhlinin 2007). For example, Hydra A has $M \sim 1.3$ (Simionescu et al. 2009) and the shocks in all the 15 clusters recently checked by Botteon et al. (2018) have $M < 2$. Some remarkable exceptions are

Abell 2744 ($M \sim 3.7$, Hattori et al. 2017), Abell 3376 ($M \sim 3$, Akamatsu et al. 2018) and the Bullet cluster ($M \sim 3$, Springel & Farrar 2007). In A2163, the shocks in the two core clumps have Mach numbers estimated from the temperature profile jumps and their values are 1.6 (inner SW shock), 3.2 (outer SW shock), and 1.7 (NE shock); the two lowest values correspond to typical shock velocities $> 2000\,\mathrm{km\ s^{-1}}$, while the remaining shock is among the strongest ones and comparable to those in the Bullet cluster. However, the uncertainties on the values of the Mach numbers are relatively large (up to 0.7), so that the strong shock velocity has values within a large range of $3200 - 4800\,\mathrm{km\ s^{-1}}$. Mach numbers are also important indicators of the deviation of the merging processes from an ideal situation: their different values, as well as those of the shock velocities, indicate a collision which is likely not head-on and/or with merging constituents having unequal masses. This, together with projection effects, could play a role in explaining why the two outer shocks are at roughly equal distances from the X-ray peak even if they have different velocities. Additionally, another factor is that, during the merger of the two dense cores, the X-ray peak position can change, so that it does not necessarily coincide with the *epicenter* of the initial shock waves.

In summary, all the discussed multi-wavelength observation properties of A2163 would suggest that its core is in a non relaxed state and is likely undergoing a post-merging process, which adds complexity to the dynamical properties of this system. In particular, the cluster core structures, A2163-A1 and A2163-A2, have undergone a recent merger along the NE-SW direction, while A2163-B is infalling into A2163-A or already has collided with the peripheral region of A2163-A. The different Mach numbers and shock velocities suggest that the real merger is an energetic, non-ideal collision: it likely occurred off-axis and between merging clumps with unequal masses. A further factor contributing to possible deviations from the ideal case (shown in Fig. 3.4) is the presence of turbulent gas motions. Finally, the situation is even more complex since multiple minor mergers may occur, because of the detection of several sub-structures identified in radio and optical observations. For these reasons, although the post-merging scenario presented above is a reasonable option, the merging processes in A2163 are not conclusively defined and need a better-constrained characterization: the present thesis adds an important contribution in this sense, giving the most accurate strong lensing mass measurements of A2163-A1, the first based on spectroscopically confirmed multiple images.

CHAPTER 4

Observations and data

We describe the *HST* galaxy cluster surveys in Sect. 4.1, with emphasis on RELICS, from which part of our data are taken. In Sects. 4.2 and 4.3, we give an overview of the instruments supplying our imaging and spectroscopic data, respectively. A2163 observations and the process of data collection and reduction are finally illustrated in Sects. 4.2.2 (for imaging) and 4.3.2 (for spectroscopy).

4.1 Large galaxy cluster surveys with the *HST*

4.1.1 CLASH and CLASH-VLT

The CLASH survey, together with the successive CLASH-VLT project, gave important contributions to astrophysics and cosmology, expanding common concepts about structure formation, galaxy evolution and DM distribution[1]. In general, the precision of CLASH measurements permitted to test the predictions of the ΛCDM model, via a systematic comparison with cosmological simulations. This program intended to accomplish four main goals, which we detail in the following paragraphs.

Precise and high-resolution measures of cluster (primarily dark) matter profiles and sub-structures. Each CLASH cluster mass profile was constrained tightly by complementing the highest resolution strong GL maps (of the inner cluster region) with those from weak lensing, mm-wave and X-ray observations (on larger scales; Postman et al. 2012). Their central density concentration was measured robustly, since the CLASH sample was chosen to minimize the bias due to lensing-based selection (which would favor clusters with dense cores; Umetsu et al. 2014). Finally, in this survey, the degree of DM sub-structures was inferred through galaxy-scale lensing analyses.

To detect and to characterize some of the most distant (z>7) sources. The high mass of CLASH clusters makes them excellent gravitational telescopes: their strong magnifying power enables the detection of dozens of galaxies with $m_{\mathrm{F775W}} < 26$ mag, existing when the universe had an age $< 6\%$ of the current one (Bradley & CLASH Team 2014).

To discover in parallel fields new Type Ia supernovae out to z∼2.5. The aim is double: to test the dark energy's repulsive force over time and possible evolutionary effects in supernovae (Graur et al. 2013; Salzano et al. 2013).

[1]The CLASH and the CLASH-VLT program websites can be found, respectively, at the links: http://www.stsci.edu/~postman/CLASH/Home.html and https://sites.google.com/site/vltclashpublic/

To analyze the galaxy structure and evolution. The survey is well suited for studying galaxies in and behind CLASH clusters and the galactic assembly at $z < 3$. The stellar content of the most distant sources can be studied, thanks to their high magnification and spatial stretching. This enables the measurement of the star formation rates and stellar ages in several galaxies, as well as the study of sub-structures (clumps, spiral arms, etc.) and the comparison between inner and outer galaxy regions (Zitrin et al. 2012).

The CLASH survey was carried out under a 524-orbit *HST* Multi-Cycle Treasury program with an observing plan of 2.7 years (over Cycles 18, 19, and 20), from November 2010 to July 2013. Specifically, 474 orbits were for cluster imaging and supernova search and the remaining 50 for supernova follow-up observations. The CLASH clusters were observed with two *HST* cameras, the Wide Field Camera 3 (WFC3; Kimble et al. 2008) and the Advanced Camera for Surveys (ACS; Ford et al. 2003), in 16 broadband filters, covering the near-UV up to the near-IR ($2000 - 17000$ Å). These imaging measurements yielded photometric redshifts with a precision of $0.02(1+z)$ for clusters and background galaxies with $m_{\rm F775W} < 26$ mag.

The survey targeted the cores of 25 galaxy clusters, mostly in common with the MACS (14) and Abell (6) catalogs and the majority imaged to a depth of 20 *HST* orbits. They have a high-mass, in the range $5 \times 10^{14} < M_{\rm vir}/M_{\odot} < 3 \times 10^{15}$, and intermediate redshifts within the interval $[0.19, 0.89]$. More precisely, two sub-samples of 20 and 5 clusters were selected based on their X-ray and strong lensing properties, respectively. The first set was defined using *Chandra* X-Ray Observatory observations and consists of nearly spherical ($\langle\epsilon\rangle < 0.2$), hot ($\langle T\rangle \simeq 7.7$ keV) and mostly dynamical relaxed systems; the remaining clusters have large Einstein's radii ($\theta_E > 35''$, at $z_{\rm sp} = 2$) and were included to efficiently detect highly magnified and very distant ($z > 7$) galaxies. General and *HST* observing plan information of all CLASH clusters are summarized in Table 3 of Postman et al. (2012).

The southern-sky sample of the CLASH survey consists of 14 massive clusters at $z = 0.2 - 0.6$; 13 of them[2] were the focus of an extension of the CLASH project: the large ESO program called CLASH-VLT. It was granted 225 hours of observational time, 200 of multi-object spectroscopy and 25 of pre-imaging, initially distributed over periods P86–P89. Its goals were a natural continuation of those of the CLASH program, but with more ambitious objectives. Several new in-depth analyses on galaxy clusters were possible thanks to the more precise spectroscopic and/or spectro-photometric data. In fact, CLASH-VLT members performed a comprehensive spectroscopic campaign using the VIMOS spectrograph, very well suited for this project, with its FoV of approximately 10 Mpc at the (median) redshift of ~ 0.4. Thanks to this instrument, an area of $15 - 20$ arcmin^2 was embraced, with eight to twelve pointings per cluster, and keeping one quadrant fixed on the cluster core to favor the detection of faint lensed sources (Rosati et al. 2014). In 2014, the program mostly observed all of its targets and ended up confirming spectroscopically a very large sample of sources, observed within twice the cluster virial radius (at least): thousands of cluster members and intervening *los* galaxies, including lensed galaxies (giant arcs and multiple images) out to $z \sim 7$. Combining this unprecedented dataset with additional data from *Chandra*/*XMM* and Subaru/SupCam, the accuracy of previous cluster mass profiles was reinforced and new ones were recovered over a large radial range with strong+weak lensing, dynamics, and X-ray methods.

[2]The remaining cluster, MACSJ0429-02, was included at the beginning, but then excluded by the sample.

4.1.2 HFF

The Frontier Fields combine the power of the *HST* and the *Spitzer Space Telescope* to detect the faintest and most distant galaxies, in order to understand how the Universe has evolved over its first billion years. To this aim, massive high-magnification clusters of galaxies were used as natural gravitational telescopes: a source, in fact, can be observed near the critical curves with magnification factors of 10-100, so that objects too faint to be directly observed by *HST* can be detected.

The HFF sample consists of six clusters (four in common with CLASH), three from MACS and three from the Abell catalogs, with redshifts $0.31 < z < 0.55$. They were selected according to the program aim, namely taking into account their *HST*, *Spitzer* (and the upcoming *JWST*) observability and their lensing strength, as well as parallel field suitability and accessibility to ground-based facilities. Further aspects considered in the selection stage were the Galactic extinction and the possible presence of bright stars responsible for polluting the faint light from distant galaxies. The *HST* ACS/WFC and WFC3/IR cameras targeted these clusters in Cycles 22, 23 and 24, for over 840 orbits (140 per cluster). *Spitzer* dedicated more than 1000 hours to them, obtaining IRAC 3.6 and 4.5 micron imaging to $\sim 26.5 - 26.0$ AB magnitudes. All observations occurred over the course of three years.

With the same instruments, in addition to the six clusters, an equal number of nearly contiguous sky regions (*parallel fields*) were also targeted. There, the galaxy distribution is mostly uniform, with few very massive galaxies. Actually, these observations served as an opportunity to explore the deep sky and were not devoted to GL studies, contrarily to the cluster pointings. Thus, parallel fields allow the study of a deep sky *corridor* hosting a wide assortment of galaxies, varying in age, shape, and stellar populations, while cluster field observations offer the opportunity to improve the knowledge about the universe overdensities and the DM distribution in the halos hosting them. Beyond DM and galaxy cluster studies, they are also helping in understanding the star formation history and the galaxy evolution during the epoch of reionization.

The HFF data are publicly available, as well as the cluster lensing models, and can be found at the Mikulski Archive for Space Telescopes (MAST) and the NASA/IPAC Infrared Science Archive.

4.1.3 RELICS

RELICS is part of a *HST* Treasury program (PI Coe D.) and *Spitzer Space Telescope* GO program (PI Bradač M.). This survey allowed the first high-resolution imaging of 41 massive galaxy clusters by infrared *HST* ($0.4 - 1.7\,\mu$m) and *Spitzer* ($3.0 - 5.0\,\mu$m) observations. It was primarily designed and optimized to discover very specific targets: candidates with very high redshifts ($z > 9$), from the Universe's first billion years, exceptionally magnified by gravitational lensing and, thus, bright enough for detailed follow-up study with current and future observatories (like the NASA *James Webb Space Telescope*, *JWST*). CLASH yielded many high-redshift candidates, but was not optimized to deliver the above-mentioned lensed targets; for instance, the first 18 CLASH clusters provided 262 candidates at $z \sim 6 - 8$ (Bradley & CLASH Team 2014), but only a few at $z \sim 9 - 11$ (Zheng et al. 2012; Coe et al. 2013; Bouwens et al. 2014). Likewise, the very high-redshift candidates of HFF were only a few. The efficient detection and later collection of these galaxies its the main goal of RELICS: to build a statistically significant sample of these

targets ($40 - 200$ at $z \sim 9 - 12$) to better constrain the luminosity function and improve subsequent studies of the epoch of reionization[3] (Robertson et al. 2015; *Planck* Collaboration: Ade et al. 2016a). This could centrally contribute to the study of early galaxy formation and evolution: the galaxy luminosity function, in fact, can be used to infer the distribution and statistical properties of galaxies over cosmic time, including the early stages of their evolution (Livermore et al. 2017). This is challenging for sources at about $z > 9$ (Bouwens et al. 2015), since the number of detections at such high redshifts is small and, as a consequence, the epoch of reionization is not yet well understood. RELICS strategy is to bridge this gap, namely to constrain better the high-redshift galaxy luminosity function, exploiting lensing magnification to recover these distant objects, affected by the cosmological dimming. In Coe et al. (2019), the entire RELICS program and its sample are described in detail as well as its ancillary goals, which we summarize in what follows, together with their motivations and implications.

Strong Lens Modeling. To properly quantify the magnification factor of RELICS lensed galaxies, as robust as possible cluster strong lens modeling and spectroscopic redshift measurements are required. In fact, the use of photometric redshifts for multiple image systems can introduce a systematic error in the magnification by up to 20%, while ≥ 3 spectroscopic redshifts reduce it to $< 5\%$ (see Fig. 4.1); the number of multiply-imaged galaxies and sub-structure as well as interlopers can contribute to the uncertainties. For this reason, RELICS also uses ground-based instruments to measure spectroscopic redshifts of strongly lensed galaxies, in order to model via GL all of its clusters with high precision.

Cluster mass measurements. The scaling relations between different estimates of the galaxy cluster mass (from weak, strong lensing, X-ray and SZ analyses) have current uncertainties which can be decreased by the study of RELICS clusters. A weak+strong lensing analysis of 20 CLASH clusters has recently measured masses with a total fractional uncertainty of 25% per cluster, at $\langle M_{500} \rangle = 10^{15}\,\mathrm{M}_{\odot}$ (Umetsu et al. 2016). Performing a similar study for all 41 RELICS clusters would yield a mass calibration with an accuracy of $\sim 4\%$ out to the virial radii, providing a legacy database to test cosmology and structure formation paradigms (Coe et al. 2019). In particular, better calibration of SZ mass estimates is required, since, currently, they underestimate the most massive clusters (Coe et al. 2019).

DM particle constraints. Such a complete study, when focused on merging clusters, can contribute in constraining self-interacting DM (SIDM, Spergel & Steinhardt 2000). This is a kind of DM with a non-null cross section, whose possible interaction (with itself) can be understood when massive sub-structures in cluster cores merge. More precisely, it is crucial to obtain precise measurements of the positions of different mass components: the center of the cluster DM halo, that of the galaxy distribution and the position of the ICM X-ray peak. This possibility is also discussed in the final section (Sect. 7.3), where we explain a possible way to constrain SIDM within the context of our results. Here, we want to remark on the potential of RELICS in this respect: this survey includes new merging cluster candidates and, among the 17 confirmed ones, some have existing *Chandra* X-ray imaging. Together with the lensing models, these data help to constrain the merging processes and, thus, to tighten limits on the SIDM parameter space.

[3] It is the period ($\sim 1 Gyr$ after the Big Bang or $z > 6.5$) when the first stars and galaxies formed, releasing large amounts of ultraviolet light, which ionized the neutral intergalactic medium.

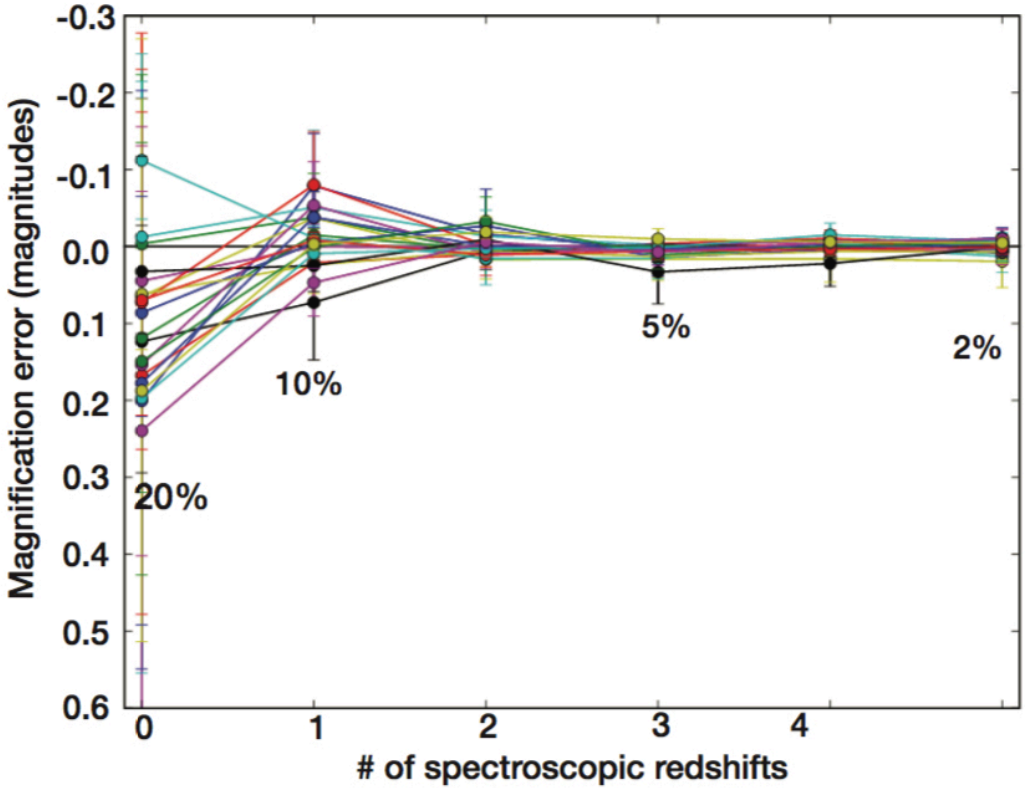


Figure 4.1: Variation of the systematic error of the magnification as a function of the number of spectroscopically confirmed multiple images. The data refer to several gravitational lensing models (shown in different colors) and the statistical errors are not included (Credits: T. L. Johnson, C. Cerny, American Astronomical Society, AAS Meeting 229, 01/2017).

Supernova searches. RELICS is contributing significantly to increase the numbers of known distant and lensed supernovae identified in other surveys. After the first pointing, all RELICS clusters are observed in a second epoch to discover transient phenomena, like supernovae. The cadence (one or two months) has been thought to detect supernovae near the brightness peak and at $z \sim 1-2$. Their light curves can then be traced using the 20 follow-up orbits included in the RELICS proposal and, in particular, the lens model magnification estimates can be tested when Type Ia supernova brightness is known. Moreover, further constraints of progenitor properties can be obtained. Finally, RELICS can provide information about the transient phenomena revealed by previous observations: magnifications of individual stars (including possible Population III stars) and stellar eruptions in galaxies (Kelly et al. 2018; Rodney et al. 2018; Chen et al. 2019; Kaurov et al. 2019).

RELICS sample consists of 41 clusters, which were chosen among the most massive ones of the *Planck* program, through two criteria: the high value of their *Planck* PSZ2 mass estimate, $M_{500} > 4 \times 10^{14}\,\mathrm{M}_{\odot}$, and the presence of exceptional strong-lensing features observed with ACS imaging. In particular, based on these conditions, 21 and 20 clusters were selected, respectively. We report in Fig. 4.2 the complete set of *Planck* SZ catalog clusters, with the RELICS sub-sample marked with red boxes.

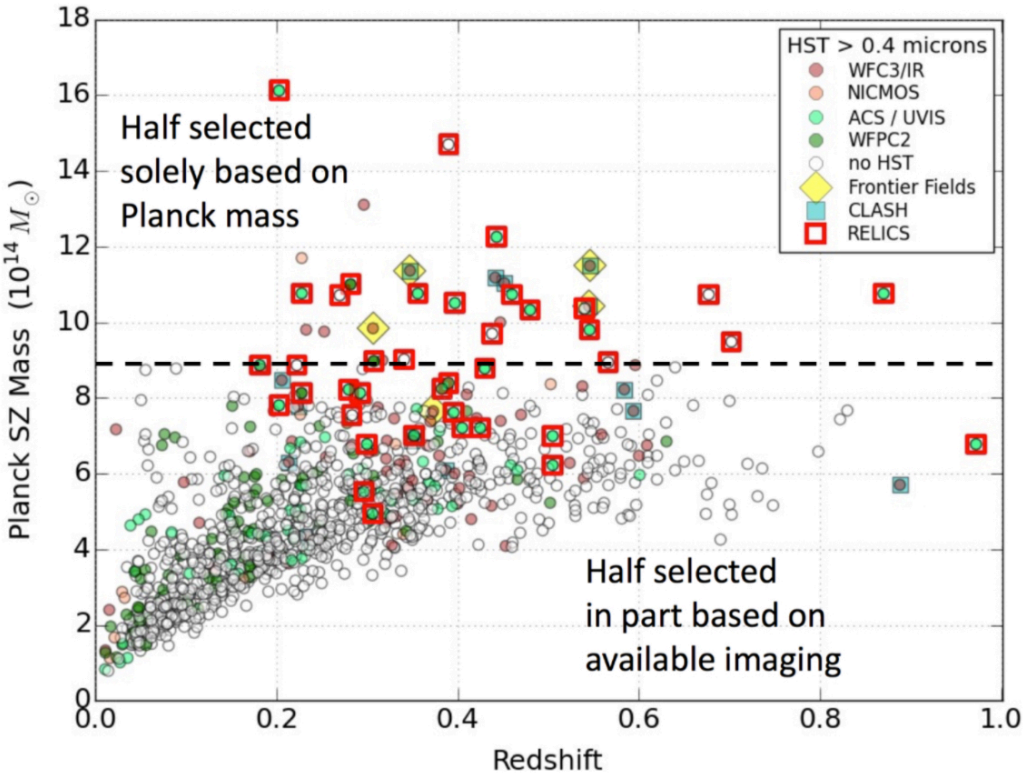


Figure 4.2: The figure shows the cluster mass from the *Planck* SZ catalog (*Planck* Collaboration: Adam et al. 2016) as a function of the redshift, for different observations and astronomical programs (taken from the RELICS project website). The RELICS clusters (squared red boxes) are very massive; most of the boxes are in fact concentrated in the high mass area. In particular, A2163, the most massive and at $z \sim 0.2$, is at the top-left corner.

Building on the successes of CLASH and the HFF, the RELICS program allows all the 34 most massive *Planck* clusters to have both *HST* optical/near-infrared imaging and *Spitzer* infrared imaging. In fact, RELICS provided the first *HST* IR images of these fields. In the 188-orbit *HST* Treasury Program 14096 (PI Coe, STScI), the targets were observed between September 2015 and April 2017, using the cameras ACS (in F435W, F606W, and F814W), with three orbits, and WFC3/IR (in F105W, F125W, F140W, F160W), with two orbits and pointings of $\sim$ 200 arcmin^2. Additionally, cluster *Spitzer* imaging was obtained through the *Spitzer* GO program 12005 (PI Maruša Bradač, UC Davis) and *Spitzer* DDT program 12123 (PI Tom Soifer, Caltech). More details about the data collection can be found in Sect. 4.2.2, where more specific aspects concerning A2163 observations are illustrated. Reduced *HST* and *Spitzer* images and catalogs of all RELICS clusters are available via MAST and via *IRSA* websites, respectively.

The lens model of each cluster is also publicly released, progressively, once it is completed. Specifically, at the time of writing (June 2019), the strong lens models of 28 clusters are available on the RELICS website[4] and some of them have published analyses (e.g., C18). The RELICS team plans to supply all of them before the *JWST* GO Cycle 1 call for proposals. In addition, a further release is scheduled for the future: each lens will be remodel on the basis of new spectroscopic redshift measurements, since refined models are crucial to generate reliable magnification maps to interpret the measurements of background, high-redshift galaxies.

Another present-day result from RELICS is the discovery of > 300 $z \sim 6 - 10$ candidates, which are currently under follow-up observations with facilities such as Keck, VLT, Subaru, Magellan, MMT, GMRT, and ALMA, besides being prime targets for the forthcoming *JWST*. Among these sources, some remarkable examples include the brightest galaxy known at $z \sim 6$ (Salmon et al. 2017) and the most distant galaxy lensed to form a spatially-resolved arc, SPT0615-JD, at $z \sim 10$, namely about 500 million years after the Big Bang (Salmon et al. 2018). At lower redshifts, RELICS is studying in greater detail the current analogs to high-redshift galaxies: compact, low metallicity dwarfs. Finally, as anticipated, for each cluster, observations were split into two epochs, separated by 25-60 days, to enable variability searches, in addition to cluster and high-redshift science. Currently, > 10 supernovae have been discovered, including 3 lensed supernovae, followed up with 20 orbits.

4.2 Imaging

4.2.1 *HST*

The idea of telescopes operating in outer space was conceived in 1923, by the German scientist Hermann Oberth, one of the founders of rocketry, who proposed to launch them into space aboard a rocket. The project to make a space telescope began more than a decade before NASA's foundation, in 1946, when the American astrophysicist, Lyman Spitzer, Jr. (1914-1997), wrote a paper proposing the development of a space-based observatory (Spitzer 1946). He discussed the benefits of such an extraterrestrial-based instrument, like the possibility to observe, through a broad range of wavelengths, astronomical objects which would not be affected by the atmospheric distortion. However, after this idea was conceived, it took about five decades to make the space telescope a reality, a period in which Spitzer's contribution in the design and development of the *HST* was instrumental. His enthusiastic support for the telescope development plan influenced both the Congress and the scientific community and his deep participation in the *HST* program continued even after *Hubble*'s launch, in 1990. *HST* was a multi-billion dollar NASA-ESA project[5], in which the telescope launch was scheduled for October 1986. However, it was delayed of four years, due to a tragic event: the Challenger Shuttle exploded killing seven NASA astronauts during takeoff. After that, Shuttle flights ceased for two years. The Space Shuttle Discovery was then used to put the 11-ton bus-sized telescope in orbit, with five instruments: The Wide Field/Planetary Camera, the Goddard High Resolution Spectrograph, the Faint Object Camera, the Faint Object Spectrograph and the High Speed Photometer.

[4] https://archive.stsci.edu/prepds/relics/

[5] The ESA participated before the *HST* launch under a memorandum of agreement, guaranteeing to ESA 15% of the telescope observation time (*Status of the Hubble Space Telescope Program*, Space Science, May 1988).

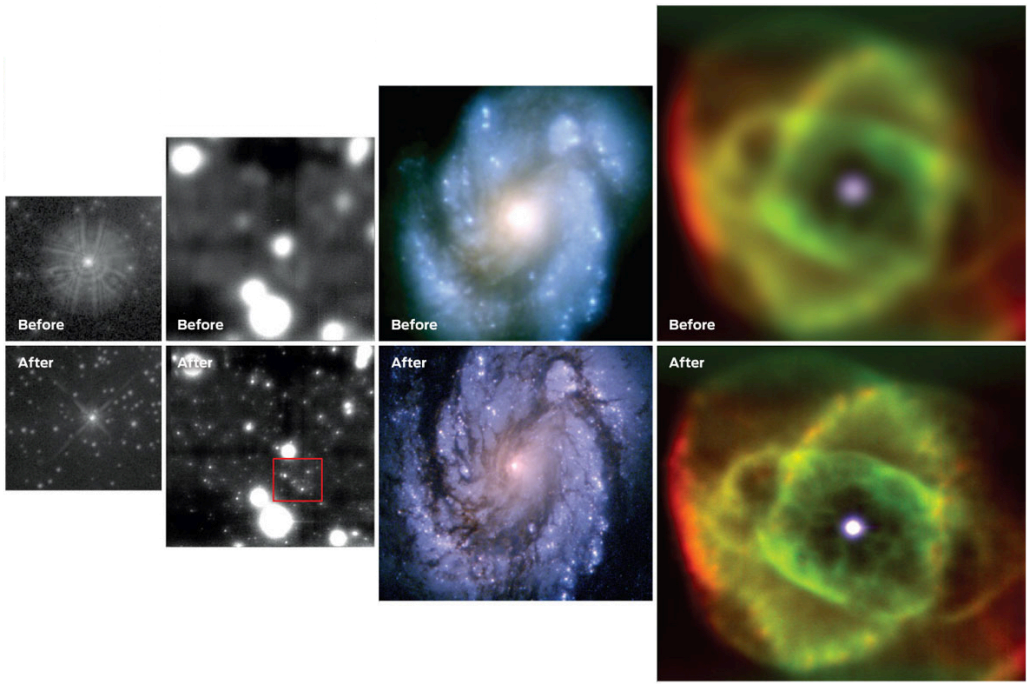


Figure 4.3: Sequence of images taken with the *HST* (columns 1 and 3) and the Palomar observatory telescope (columns 2 and 4). The first and third couple illustrate the spherical aberration effects on the same *HST* pointing, before (upper panels) and after (lower panels) the telescope mirror fixing (Credits: New Eyes and Messier 100, ESA *HST*/NASA). The second and fourth couple of images are described in the text and show, respectively, the globular cluster M13 and the Cat's-Eye Nebula (Credits: Caltech/JPL/Cambridge/Palomar Observatory).

After the very first *HST* observations, NASA scientists discovered a defect in the main focus lens, polished in the wrong way. It was only in December 1993 that this negligence (which costed 1.5 billion dollars to NASA) was compensated for by a repair mission using the shuttle Endeavor. To have an idea of the higher quality of the new observations, we compare them to those before the telescope fixing in the first and third couple of images of Fig. 4.3. Nevertheless, it must be noted that, a decade ago, researchers from the Palomar Observatory achieved an even better resolution than that of the *HST* using a ground-based telescope (a remarkable result considering the significant distortions by Earth's atmosphere). In Fig. 4.3, the couple of images in columns 2 and 4 compare observations before and after the use of their new technique, with which twice the resolution of the *Hubble* was reached[6]. On the bottom panel of column two, we also highlight a specific region (red box) to better appreciate the difference in terms of resolution compared to the same region in the upper panel image.

Nowadays, the *HST* is more technologically advanced than when it was launched, thanks to five NASA space shuttle servicing missions, which handled the maintenance and the upgrades. *Hubble*'s current scientific instruments are shown in Fig. 4.4, together

[6]The combination of these technologies, involving a Shack-Hartmann wavefront sensor, a deformable mirror, and the Lucky Camera (built by the Institute of Astronomy in Cambridge, England) was named one of *Time* magazine's Best Inventions of 2007.

with the information on the wavelength range they cover. All of them are powered by sunlight, directly converted into electricity, some of which is stored in batteries to keep the telescope running when Earth shadow blocks the Sun rays. We report in Table 4.1 a list of general features of the telescope (a Cassegrain reflector) and briefly describe, in what follows, its present-day equipment, consisting of three types of instruments: cameras, spectrographs and interferometers.

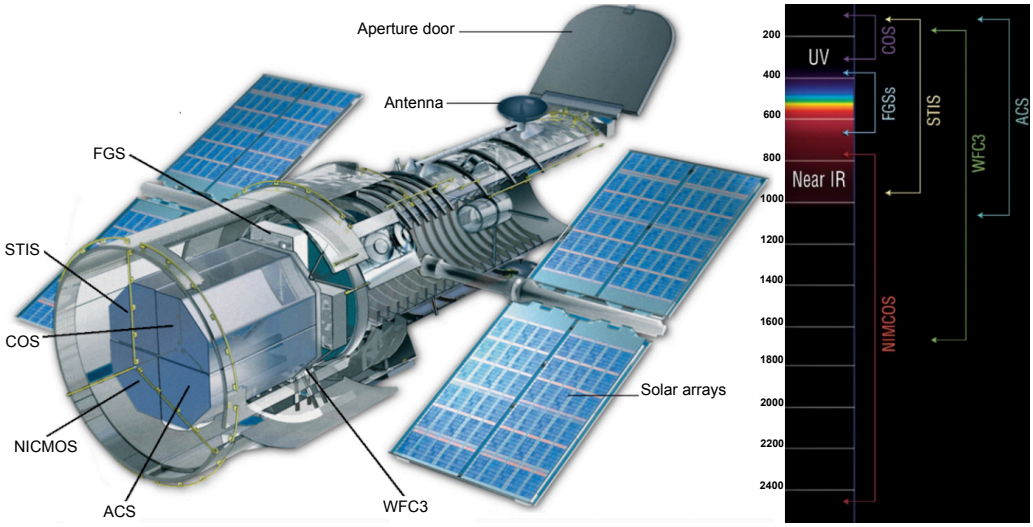


Figure 4.4: Representation of the *HST* and its principal instruments, labeled. On the right part, the vertical numerical scale, in nanometers, indicates which wavelengths each component studies. All these instruments are described in the text, except for the Near Infrared Camera and Multi-Object Spectrometer (NICMOS), since it is currently inoperative: it was only active during about 8 years (1997-2008) for IR imaging and spectroscopy. The figure is a modified and composite version of different NASA images (Credits: NASA's Goddard Space Flight Center).

Cameras. The *HST* primary camera systems are two: the Advanced Camera for Surveys (ACS) and the Wide Field Camera 3 (WFC3). They work together to cover a broad range of wavelengths and to provide superb wide-field imaging. ACS was installed on *Hubble* in 2002 and can detect light from UV to NIR, although it was designed primarily for wide-field imagery in visible wavelengths. It has three cameras (*channels*), capturing different types of images, two of which (the most used) were rendered inoperable due to an electrical short in January 2007. During Servicing Mission 4, in May 2009, one of them was repaired, restoring the capacity of the ACS to perform high-resolution, wide-field observations. In fact, this third-generation instrument is exactly designed and optimized to carry out surveys or broad campaigns of imaging. ACS greatly helps in several astrophysical studies, such as those on the evolution of galaxy clusters and the DM distribution, and detected the most distant objects in the universe, as well as relatively massive planets orbiting close stars. Regarding WFC3, it was installed in 2009 and replaced the Wide Field Planetary Camera 2 (WFPC2). It provides wide-field imagery in UV, visible, and IR light, thus complementing ACS imaging capabilities with deeper IR and UV observations. In addition, WFC3 has a FoV much greater than that of the other *HST* instruments and is also 35 times more sensitive in the UV band than its predecessor. The IR channel increases *Hubble*'s infrared resolution from 6.5×10^4 to 10^6 pixels. WFC3 is used to study dark energy, DM, and star formation. Moreover, it has

been a key instrument for the discovery of extremely remote galaxies beyond *Hubble*'s possibility with its previous IR camera.

Spectrographs. The ability to break light down to its components is crucial to determine, for astrophysical sources, their characteristics, such as temperature, density, chemical composition and motion. The *HST* has currently two spectrographs: the Space Telescope Imaging Spectrograph (STIS) and the Cosmic Origins Spectrograph (COS). STIS and COS provide a full set of spectroscopic tools, especially when they work together, being complementary in the analysis of a variety of astrophysical objects: the former, a versatile, "all purpose" spectrograph, targets bright and large objects, while the latter measures exceedingly faint UV radiation from remote cosmological sources, like quasars, and is best at observing point-like sources. COS focuses exclusively on this band of the electromagnetic sprectrum and is one of the most sensitive ultraviolet spectrographs ever: compared to previous instruments, it increased *HST* UV sensitivity several times, up to 70 times when looking at extremely faint objects. Regarding STIS, it is a second-generation spectrograph/imager (in the UV-NIR range), which was also repaired during Servicing Mission 4, after a technical failure on August 3, 2004. It targets resolved objects and obtains their spectra at a high level of resolution. Thus, STIS can map large objects like galaxies, but it is also known for hunting black holes. This instrument has also the special ability to return spectra from several different points along each target.

Interferometers. The three *Hubble*'s interferometers, called the Fine Guidance Sensors (FGSs), serve a dual purpose: scientific instruments and a means to keep the telescope pointing steadily. Specifically, during a pointing, two of the three FGSs seek out stable point sources of light (known as "guide stars") and then lock the telescope, allowing it to maintain a steady aim. In a preliminary stage, they reject guide star candidates which are demonstrated to be non-point sources (often, faraway galaxy or double-star systems). For certain observations, the third sensor can be used to collect information about a target. So, FGSs can also function as a scientific high-sensitive instrument, precisely measuring astronomical object's angular diameter, relative positions of stars (with ten times more accuracy compared to ground-based telescopes) and their brightnesses. These interferometers are also able to detect possible rapid changes in the brightness of a star and, as anticipated, they can resolve double-star systems that appear point-like even to *Hubble*'s cameras.

4.2.2 Abell 2163 imaging data

A2163, described in detail in Chap. 3, was observed by the RELICS team using only the *HST* camera WFC3/IR, in two epochs separated by 40 days (according to the RELICS follow-up strategy): on March 9th (visits A2 and B2) and April 18th (visits A4 and B4) of 2016. It is not possible to use simultaneously the UVIS and IR channels of WFC3/IR, but one or more *HST* instruments can be used at the same time. Each of them covers different locations in the telescope focal plane, so parallel observations usually sample a sky area several arcminutes away from the WFC3 target. This means that for extended nearby or distant objects, parallel observations may look, respectively, at adjacent regions of the primary target or at mainly random sky areas. In contrast to other RELICS clusters, for A2163 no parallel observations were made. However, we remark that clear and strong justifications have to be provided to be granted observing time in parallel mode, since it is limited during each cycle, essentially to prolong the life of the *HST* transmitters.

Table 4.1: List of some general characteristics of the *HST*. The different kinds of information are organized in groups of rows, separated by an horizontal line; they refer, respectively, to the mission, size, orbit, power needs/storage and mirrors.

Feature	Description
Launch date	April 24, 1990
Space shuttle	Discovery STS-31
First Image date	May 20, 1990
First Image	Star cluster NGC 3532
Data per week	150 gigabits
Length	13.2 m
Maximum Diameter	4.2 m
Weight (at launch)	10,886 kg
Weight (after 2009)	12,247 kg
Orbit altitude	547 km
Orbit inclination	28.5° (equator)
Orbit period	95 min
Speed	27,300 kph
Energy Source	the Sun
Solar panels	2 of 7.6 m
Power Usage (average)	2,100 watts
Batteries	6 NiH
Primary Mirror Diameter	2.4 m
Primary Mirror Weight	828 kg
Secondary Mirror Diameter	0.3 m
Secondary Mirror Weight	12.3 kg

Two different adjacent areas of the cluster were observed, with two orbits each: the north-east region (target 111, R.A. 16:15:48.3265, Decl. −06:07:36.70) and the south-west one (target 112, R.A. 16:15:42.6336, Decl. -06:09:22.12), on visits A2-A4 and B2-B4, respectively. During each orbit, the same four filters were used (twice and with the same exposition time): F140W (178+178 sec), F105W (353+353 sec), F125W (178+178 sec) and F160W (503+503 sec).

Since for A2163 archival data from the ACS were available, the RELICS team did not observed this cluster with that camera. In fact, they considered ACS data collected in the *HST* cycle 18, thanks to the Proposal GO 12253 (PI Douglas Clowe, 2010): "Gravity

in the Crossfire: Revealing the Properties of Dark Matter in Bullet-like Clusters". The goal of this program was to improve substantially the DM cross section calculations, targeting massive Bullet-like clusters in a merging state (three in this case). This aim can be achieved by a gravitational lensing mass reconstruction of the merging clusters and a successive comparison with that from X-ray studies (see Sect. 7.3 for a description about this appealing possibility). A2163 was observed with eight total orbits: in two different pointings (four orbits each), two adjacent areas of the cluster were observed, at south-west (target 7, R.A. 16:15:41.7199, Decl. −06:10:6.94) and at north-east (target 8, R.A. 16:15:50.1331, Decl. −06:07:41.23). In both occasions, the filters F814W, F606W and F435W were used. A summary of the information about A2163 observations and the filters of both the *HST* cameras is reported in Table 4.2, together with exposure times, while in Fig. 4.5 we illustrate the response curves of all the filters used in A2163 observations.

Table 4.2: Information on A2163 observations: *HST* camera used, its filter and "pivot" wavelength (Tokunaga & Vacca 2005), with exposure times related to the two days in which data were collected. The project ID is also listed. ACS camera data were taken from C18.

Camera	Filter	λ [μm]	Exposure time [sec]	Observation date	Project
WFC3/IR	F105W	1.06	706	2016 Mar 09	RELICS
			706	2016 Apr 18	
WFC3/IR	F125W	1.25	356	2016 Mar 09	RELICS
			356	2016 Apr 18	
WFC3/IR	F140W	1.39	356	2016 Mar 09	RELICS
			356	2016 Apr 18	
WFC3/IR	F160W	1.54	1006	2016 Mar 09	RELICS
			1006	2016 Apr 18	
ACS	F435W	0.43	4664	2011 Jul 03	GO 12253
ACS	F606W	0.59	4667	2011 Jul 03	GO 12253
ACS	F814W	0.81	9192	2011 Jul 03	GO 12253

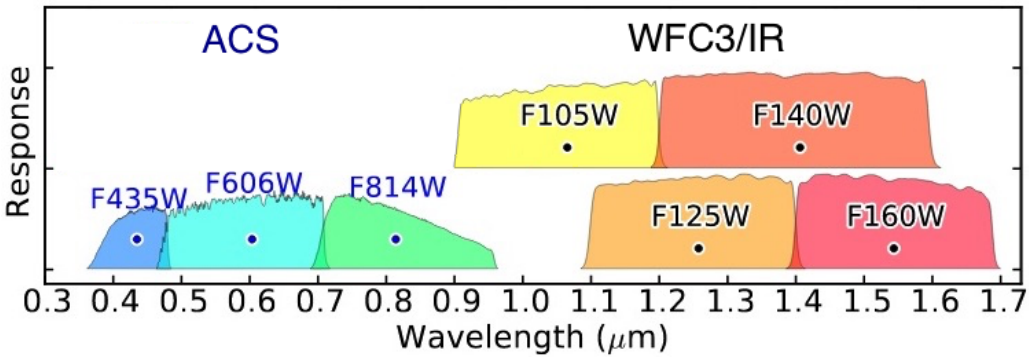


Figure 4.5: Response curves of the ACS and WFC3/IR filters used to observe A2163, as a function of the wavelength. A vertical offset for F105W and F140W was applied for clarity and the black dots indicate the filter effective "pivot" wavelengths (Tokunaga & Vacca 2005), listed in Table 4.2.

The reduction procedure of RELICS *HST* images is similar to that performed on the HFF sample (Lotz et al. 2016). Data were corrected for bias, dark current, flat fields with updated reference files and satellite trails were manually identified and masked, as well as cosmic rays and bad pixels. Moreover, for each RELICS cluster, all images were aligned to two common grids with a resolution of $0''.06$ and $0''.03$. Fig. 4.6 shows a color composite image of the cluster A2163, obtained as the combination of the filters F435W, F606W+F814W and F105W+F125W+F140W+F160W, respectively, for the blue, green and red channels. It was created starting from the $0''.06$ resolution images released by the RELICS team and available via MAST database.

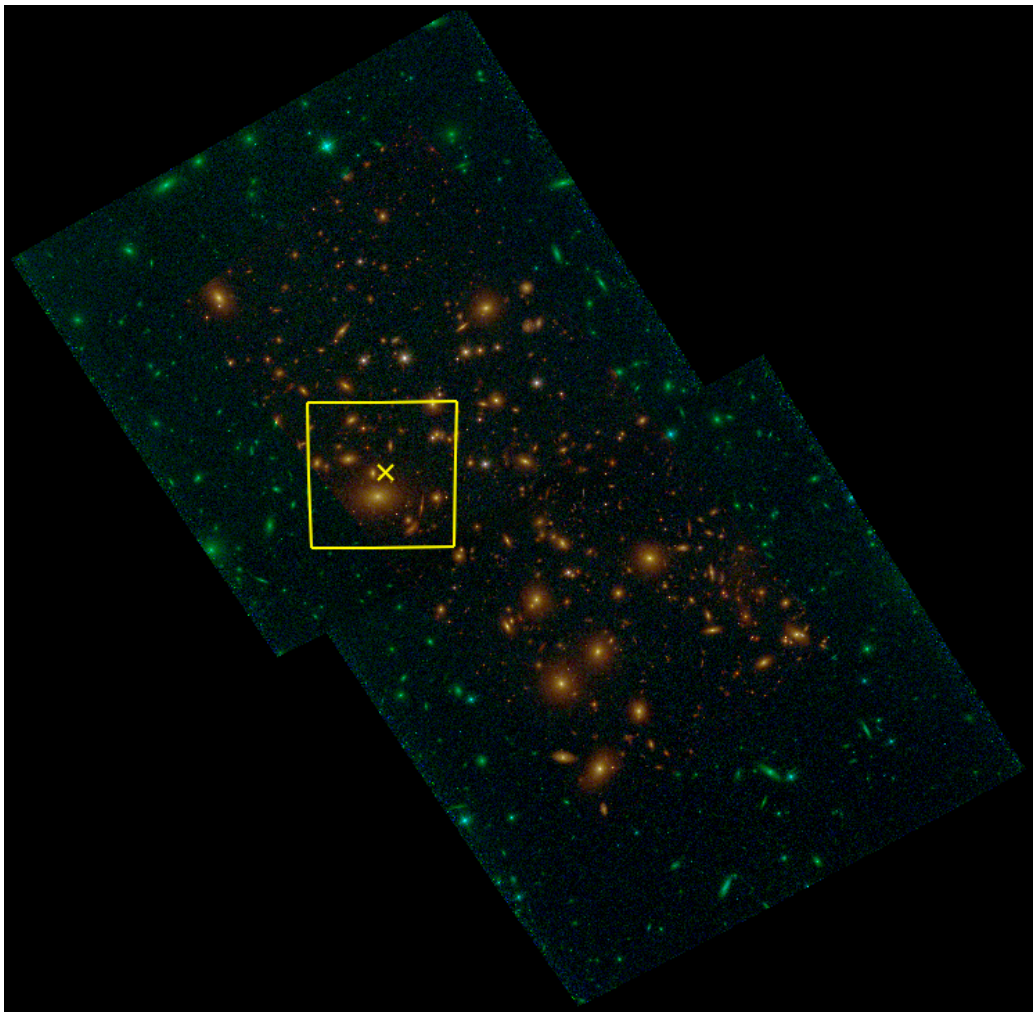


Figure 4.6: Color composite image of A2163. This is a combination of ACS and WFC3 data, which were collected thanks to two different *HST* Proposals, respectively, a 32-orbit Treasury program (ID GO 12253, cycle 18, P.I. D. Clowe) and a 190-orbit program (ID GO 14096, cycle 23, P.I. D. Coe). The yellow region marks the FoV of one of the MUSE pointing, from which our spectroscopic data were taken (see Sect. 4.3.2).

4.3 Spectroscopy

4.3.1 MUSE

In this thesis, we retrieved and used archival data from the Multi Unit Spectroscopic Explorer (MUSE, Bacon et al. 2012), located at the Nasmyth B focus of Yepun, the VLT UT4 telescope, at the cerro Paranal in the Atacama desert, Chile. This instrument implements a relatively recent technology, known as Integral Field Spectroscopy (IFS), which has had a very rapid development in the last few decades. It combines imaging and spectroscopy and allows to simultaneously record many spectra over a 2D FoV. Regardless of the technique used to generate the data, the final product of these spectrographs is (usually) a data-cube, represented in Fig. 4.7.

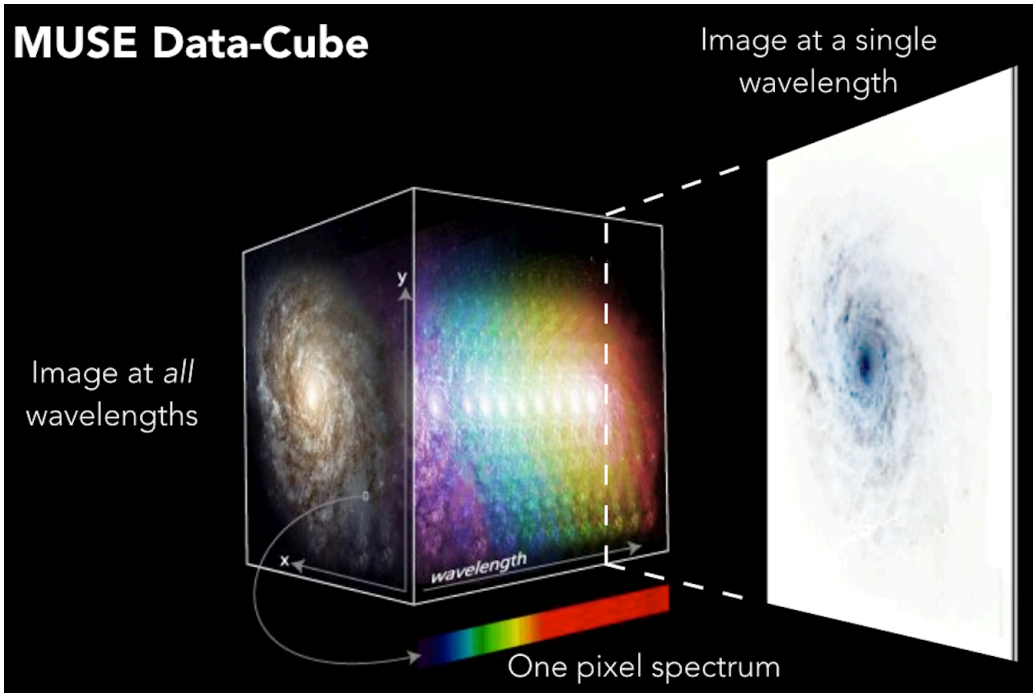


Figure 4.7: Representation of a MUSE data-cube. The figure shows the different cube components and the information that can be extracted from them: the side of the cube at the origin wavelength λ_0 and one cube slice at a certain $\lambda \neq \lambda_0$ give, respectively, the source image at all or at a *single* wavelength. Navigating the data-cube, from a pixel at $\lambda = \lambda_0$ until the end of the wavelength axis, one can also obtain the pixel spectrum.

It has three axes: two correspond to the angular dimensions (R.A., Decl.) and the third one to the wavelength. Thus, each data cube slice perpendicular to the spectral axis is a monochromatic image and each cube column parallel to it represents a spectrum (see Fig. 4.7). In the IFS terminology, some basic concepts to figure out the cube structure are those of pixel, a data-point on a CCD, of *spaxel*, a spectrum in a data-cube, and of *voxel*, a data-point in a data-cube. Hence, many voxels create a spaxel, a shortened version of *spatial pixels*, term used to differentiate between a spatial element on the integral field unit and a pixel on the detector.

IFS was developed to surpass the main disadvantages of long-slit spectroscopy, the traditional one. The latter could record a spectrum from each part of an extended object only in a very inefficient way: varying the slit position across the target, by moving the telescope and recording separate exposures for each position[7]. Conversely, IFS can boast this ability, which implements with no time-consuming phases. Moreover, this technology has the capability to address a number of common issues, such as separating sources which are spatially blended (e.g., a dense star field, foreground emission superimposed on strong background emission). An example of the IFS use to separate blended sources and to detect new ones is shown in Fig. 4.8.

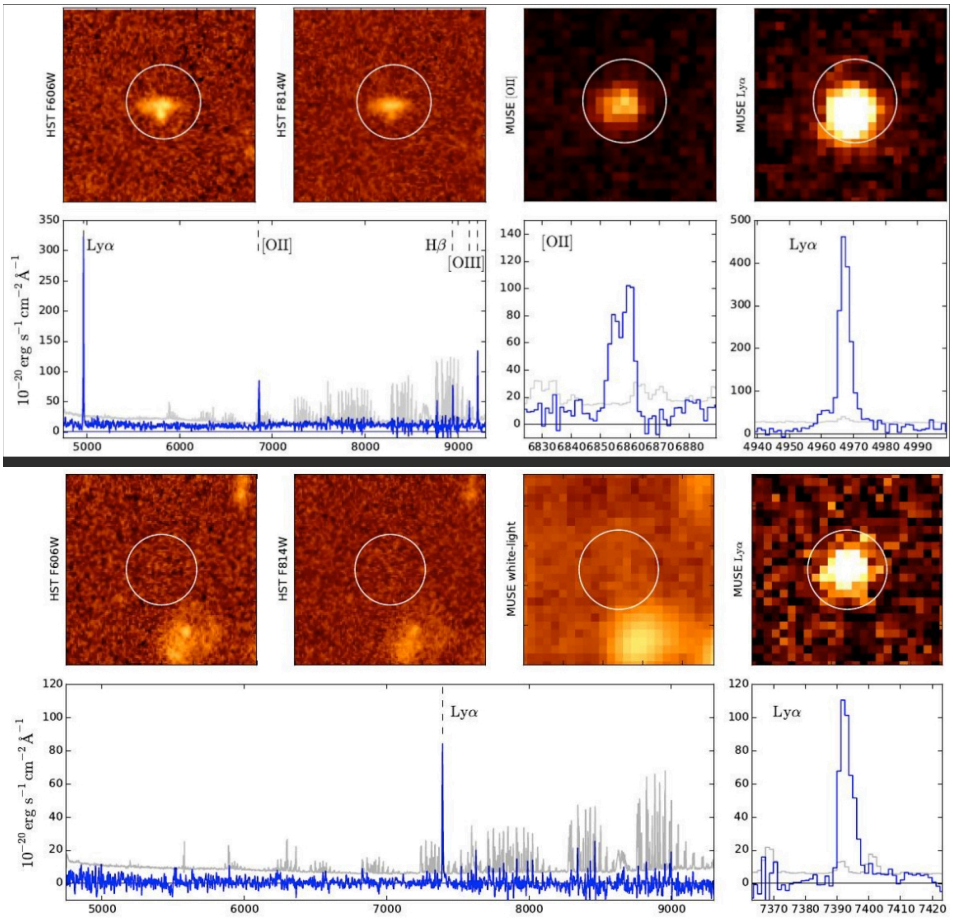
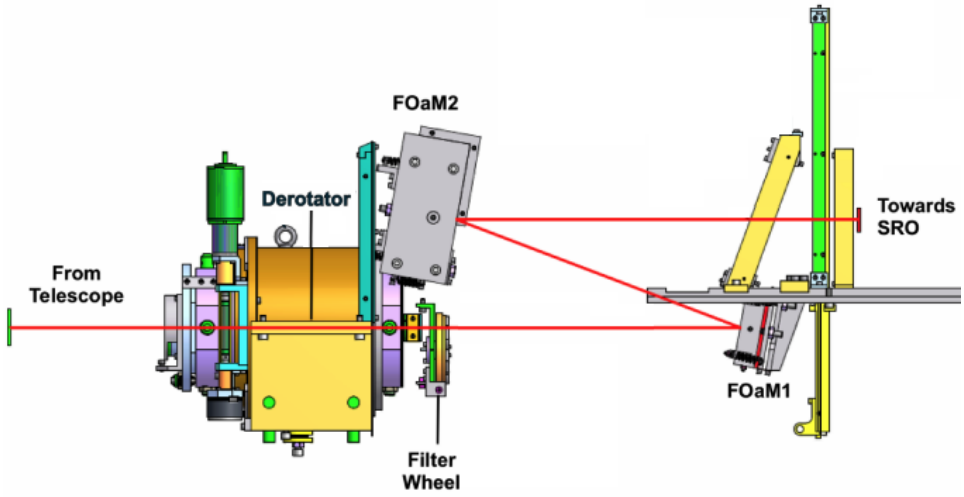


Figure 4.8: Examples of IFS capabilities from Bacon et al. (2015). Upper panel: [O II] and $Ly\alpha$ sources, blended in *HST* images (filters F606W and F814W), were separated looking at two different data-cube slices, corresponding to the wavelength interval of each of those two emission lines. Lower panel: a $(Ly\alpha)$ source discovered with IFS observations, navigating through the data-cube, otherwise invisible in *HST* data or even in the white-light image of the cube.

[7]Successively, more versatile instruments such as FORS1, FORS2, and VIMOS were assembled, although less efficient than IFS.

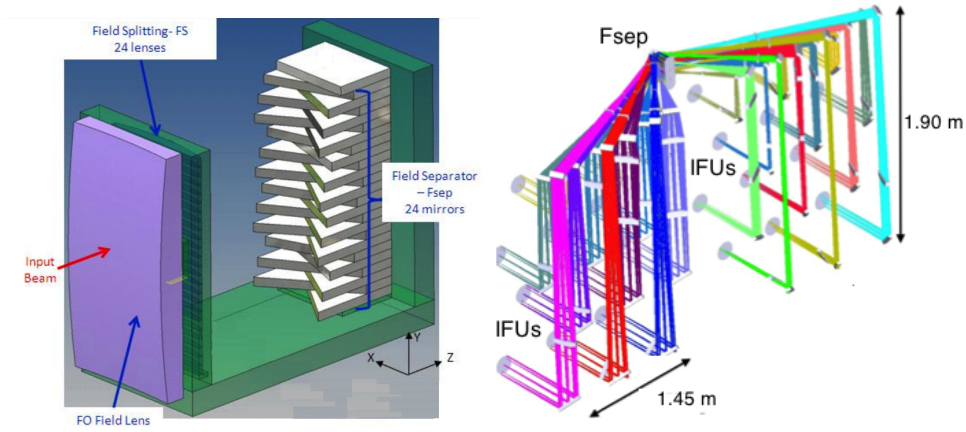
IF units (IFU) can have very different hardware architectures, although preserving the basic components of a spectrograph. We present here a simplified description of the complex structure of MUSE[8]. We summarize the phases from light detection to the outputs characterizing the instrument functioning and describe the light interaction with the main sub-systems of MUSE, following the optical path (from the telescope to the detectors):

- **Fore Optics (tower) (FO).** This component firstly hosts a *derotator*, a two-prism instrument which rotates the image, in order to compensate for the field rotation happening at the VLT Nasmyth focus and due to Earth's rotation. The light beam is then filtered (5 filters are available in the wheel) to select the spectral coverage according to one of the two possible MUSE modes: the Wide Field Mode (WFM), with a resolution (FWHM) of $0''.3 - 0''.4$ and a large FoV of $59''.9 \times 60''.0$, and the Narrow Field Mode (NFM), with an exquisite spatial resolution (FWHM) of $0''.03 - 0''.05$ in a smaller FoV ($7''.42 \times 7''.43$). Finally, light beam is magnified by the *anamorphoser*: it consists of two cylindrical mirrors (FOaM1 and FoaM2) providing the proper magnification ($2\times$) in the vertical direction, thus it reshapes the image into a 2:1 ratio.

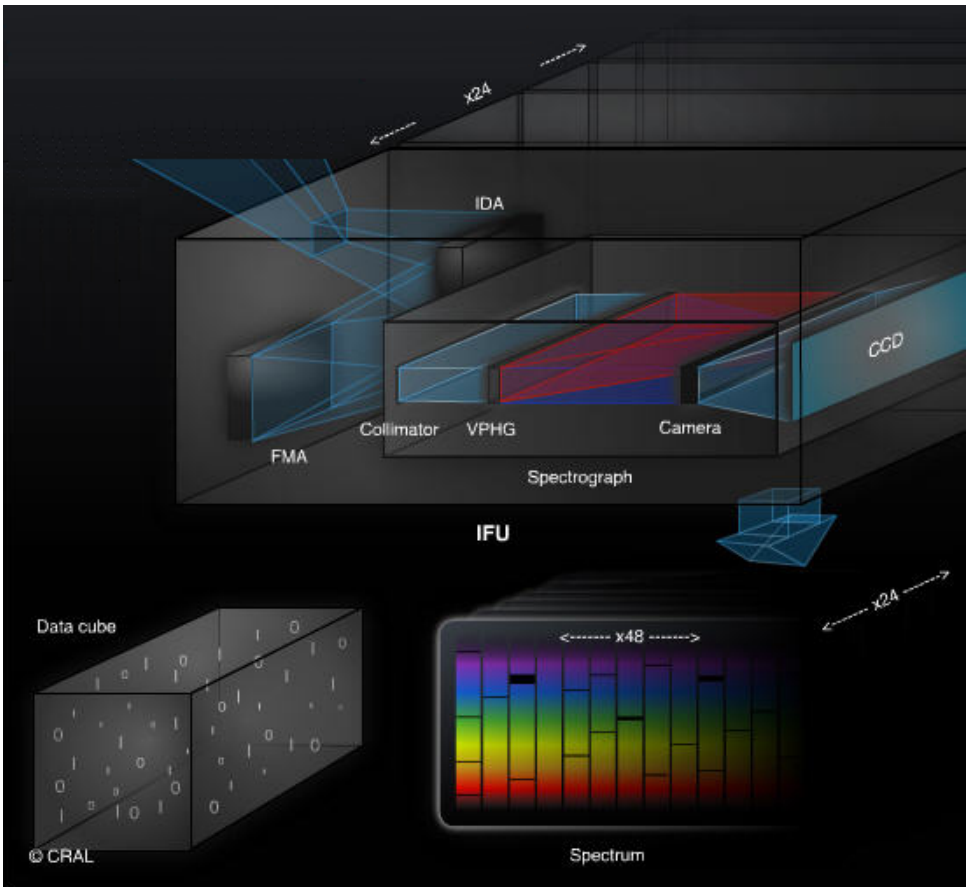


- **Splitting and Relay Optics (SRO).** The Field Splitter Unit (in the left image below) is composed of the *Field Splitting* (FS) and the *Field Separator* (Fsep): the former splits light of the FoV coming from the FO via a 24-lens array into as many rectangular sub-fields (*channels*) and the latter by a 24-mirror array reflects by the appropriate angles these sub-images into 24 different spatial directions. The light from the odd sub-fields is reflected towards $-x$ and that from even sub-fields is reflected towards $+x$. At this point, each of the 24 Relay Optics (right image below) collects one sub-image, corrects for the variations in optical path from one channel to another and redirects the light of the channel towards the entrance of a single IFU.

[8]We also bring the reader back to the following short video (from the University of Lyon), to understand the MUSE processing operations: http://muse.univ-lyon1.fr/IMG/mp4/Decoupeur_Slicer.mp4



- **Integral Field Units (IFUs).** Each of the IFUs observes one of the 24 sub-fields of the FoV and is composed of 3 sub-systems: the Image Slicer, the Spectrograph, and the Detector. The first optical device further subdivides each sub-field into 48 slices (or *mini-slits*), using two sets of off-axis mirrors (the Image Dissector Array, IDA, and the Focusing Mirrors Array, FMA); considering all 24 channels, 1152 mini-slits in total, each of $15''.00 \times 0''.208$, are created.



Each set of 48 mini-slits (namely each channel) is then injected into the same spectrograph, which disperses the light as a function of its wavelength to produce spectra, through a collimator and a Volume Phase Holographic Grating (VPHG). Mini-slit spectra are finally arranged on a single detector, a 4k×4k, 15 μm pixel CCD, cooled with liquid Nitrogen and operating at -130 °C.

- **Data management system and electronics.** A dedicated and powerful data system is responsible to collect and gather data for all the 24 channels. It organizes this amount of data in the specific 3D data format, the MUSE data-cube. Finally, during all the phases described, the detectors and the changes in the instrument configuration are controlled by electronics located in dedicated cabinets, cooled to avoid heat-generated turbulence in the telescope environment.

MUSE is, thus, a second generation optical VLT IF system, which offers a wide simultaneous spectral range ($4650 - 9300$ Å) and a resolving power of $R \sim 1700 - 3500$. The latter varies, depending on which of the two operating modes is considered: 1770 (480 nm) - 3590 (930 nm), in WFM, and 1740 (480 nm) - 3450 (930 nm), in NFM. The instrument was developed by a collaboration among seven international institutes: CRAL and IRAP (France), AIP and AIG (Germany), ETH (Switzerland), NOVA (Netherlands) and ESO. Moreover, MUSE has an Instrument Operation Team made up of two groups, each with very specific roles, at Cerro Paranal (Chile) and in Garching (Germany). Its first light on the VLT was on 31 January 2014 and, in the near future, it will benefit from the improved spatial resolution provided by the Adaptive Optics Facility in development by ESO.

4.3.2 Abell 2163 spectroscopic data

In the commissioning runs by the MUSE team, between February and August 2014, specific targets were selected in order to demonstrate the instrument capabilities, find the most appropriate observation strategies and optimize the data reduction pipeline, particularly in terms of performance. The team collected and then released these datasets, some of which can be used for scientific work. Nevertheless, it must be noticed that, in some observations, exploratory exposure times and no optimal observing strategies were considered.

Specifically, these MUSE data were collected in 2014, in three periods of commissioning: COMM1 (02-22 February), COMM2A (27 April - 7 May) and COMM2B (7 July - 4 August). They are available online, divided by science and calibration files, as well as a set of pipeline reduction tools. We employ MUSE archival data of A2163, collected in July 2014, hence within commissioning period COMM2B (Program ID: 60.A-9100C)[9] and then released in September of the same year. The MUSE survey area includes two pointings, targeting two different regions in the core of this cluster. Observations were in MUSE mode WFM, namely with a FoV of about $1'$x$1'$, and with the characteristics listed in Table 4.3 for both pointings. In this thesis, we model the north-east region only (A2163a, see the pointing footprint in Fig. 4.6), where different lensed galaxy systems are found. Previous spectroscopic observations of A2163 also exist (e.g., see Table 3 in C18), but none have lead to a definitive confirmation of the cluster multiple images.

[9]Data and calibration files for this period can be found, respectively at the link http://archive.eso.org/wdb/wdb/eso/pkg_cont/query?pkg_id=162425&max_rows_returned=1000 and http://archive.eso.org/cms/application_support/calselectorInfo.html.

Table 4.3: VLT Observations of the cluster A2163, using MUSE. Data were obtained in the MUSE commissioning phase, under the program 60.A-9100(C) (PIs: MUSE TEAM, Commissioning).

Pointing ID	R.A.	Decl.	PA	Exposure time	Observations
A2163a	16:15:48.83	−06:08:33.7	0°	16 x 900 s (4.0 h)	16
A2163b	16:15:39.30	−06:09:15.0	30°	6 x 900 s (1.5 h)	6

Our MUSE cube, thus, includes data taken from one single pointing (at $< 1''$ average seeing), with a total exposure of 4 hours. New IFS instruments are extremely complex and their typical data reduction procedures need a significant amount of resources, e.g., reducing a MUSE cube could require 18 GB of RAM. We reduce raw data through the MUSE pipeline (Weilbacher et al. 2014), which provides standard calibrations. Using the same tool, we then combine them into one single final data-cube, taking into account the positional offset of each observation. We apply all the standard calibration procedures provided by the pipeline (e.g., bias and flat field subtractions, wavelength and flux calibration, etc.), following the same procedures described in Grillo et al. (2016) and Caminha et al. (2017a,b). We achieve an overall improvement in the sky subtraction, after experimenting with different configurations of the MUSE pipeline recipes. Sky residual minimization was also the result of applying the Zurich Atmosphere Purge (ZAP, Soto et al. 2016) tool, using a segmentation map of SEx-tractor (Bertin & Arnouts 1996) to define sky regions.

The reduced data-cube has a FoV of ~ 1.13 arcmin^2, a grid of 320×318 pixels, each with a side of $0''.2$. The spectral coverage is $[4750 - 9350]$ Å, with a dispersion of ~ 1.25 Å/pixel, so the cube consists of 3681 slices, along the wavelength axis. From the stars in the white image[10] of the final data-cube, we measure the seeing, finding that it has a fairly constant value of $0''.6$. The quality of the reduced data-cubes allows the spectroscopic identifications of several sources, including both cluster galaxy members and multiple images, beyond foreground and background sources. We illustrate here the overall process of spectra extraction and redshift measurements for all of them, while in Chap. 5 we give more details and concentrate, in particular, on the multiply lensed images, the cluster galaxies and the corresponding catalogs we need for the strong lensing model of A2163 (Chap. 6).

In order to extract spectra and measure redshifts, we consider the RELICS *HST* catalog ACS+WFC3/IR, which is based on the observations and the data described in detail in Sect. 4.2.2. It extends over a total area of ~ 23 arcmin^2 and contains more than 5500 sources. From it, we remove duplications due to segmentation problems and include additional sources, visually identified. We then extract a total of 230 spectra, within circular apertures with radii of $0''.6$, that belong to the *HST* sources inside the MUSE FoV. Finally, we measure the corresponding redshifts through the software Easy-z, *EZ* (Garilli et al. 2010), which finds the best fit value of z comparing a set of (user given) spectral templates with the observed spectrum.

Each measurement is tagged with a Quality Flag (QF), which quantifies its reliability, as in Balestra et al. (2016) and Caminha et al. (2016):

[10] It is a representation of the entire FoV at all wavelengths, created by collapsing the cube along the direction of the third axis.

Very insecure **($QF = 0$).** The source is very faint and noisy, with a flat continuum emission with no features, so that the redshift value is not reliable.

Insecure **($QF = 1$).** The object is very faint, but with some detection on a mostly flat continuum. Although it can probably be matched well by a template, with no clear spectral features the redshift values remains unreliable.

Likely **($QF = 2$).** The redshift is likely to be correct: the spectrum has an intermediate quality, with at least two spectral features which are well identified.

Secure **($QF = 3$).** The spectrum of the source has two or more clear spectral lines (in emission or absorption), that give a non dubious redshift value.

Based on a single emission line **($QF = 9$).** In this case, the redshift measure is also considered reliable, since the MUSE spectral resolution allows to identify the shape of narrow emission lines (e.g., Lyα) and to distinguish fine-structure doublets (e.g., [OII]).

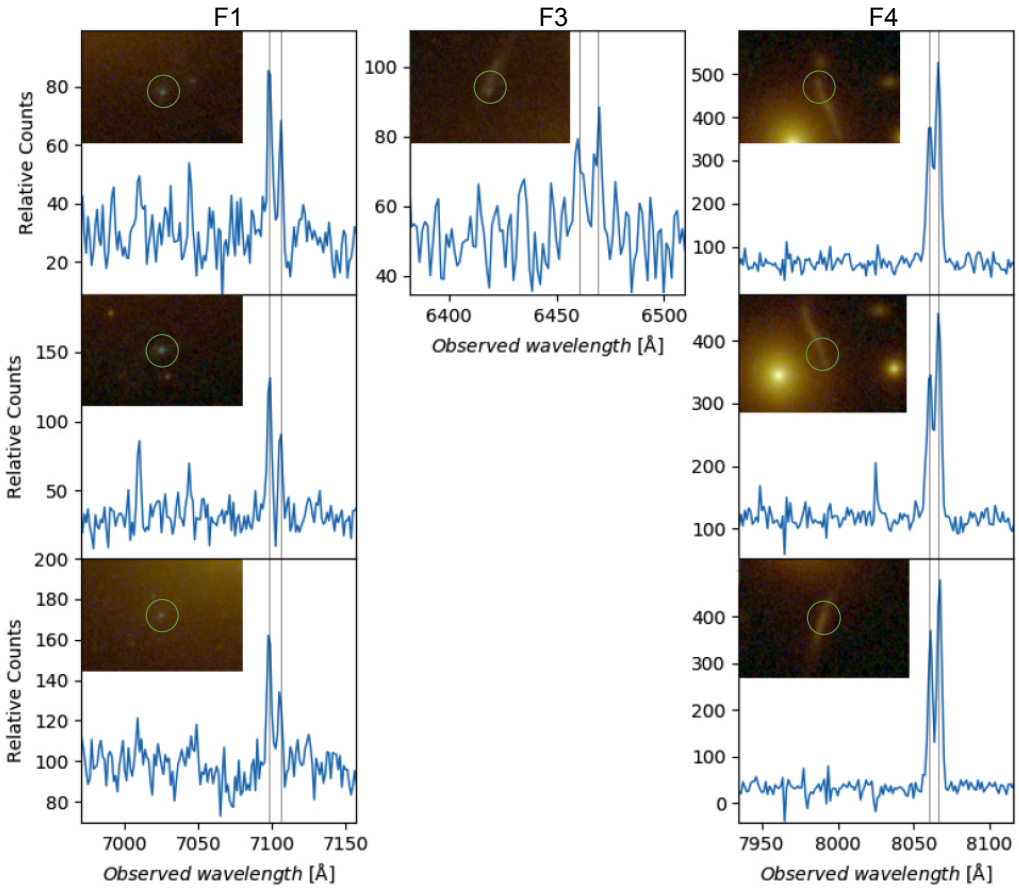


Figure 4.9: *HST* image cutouts, with the circular aperture (green circle) we use to extract the image spectra, and the relevant spectral details for the 7 secure multiple images in A2163 core. The panels of each column refer to the same family and (from the top to the bottom) to the images a, b and c, respectively. QFs for all multiple images are reported in Table 5.2 of Sect. 5.3. The spectral line doublets (in gray) are C III] ($\lambda\lambda$ 1907, 1909 from a semi-forbidden transition) and [O II] ($\lambda\lambda$ 3727, 3729), respectively, for the first two columns and for the last one.

Note that several other factors are taken into account, such as the contamination from nearby galaxies, measurement quality of sources at the edge of the detector or a high level of noise and asymmetry in the shape of Lyα emission lines. With regard to multiple images of A2163, we have measured the spectroscopic redshifts of all its four families (labeled as F1, F2, F3, F4). We can anticipate here that each of the latter has at least one secure measurement ($QF = 3$), except the family F2 (see Sect. 5.3). In Fig. 4.9 we illustrate, in each panel, one of the reference spectral lines used to estimate the image redshift and a *HST* snapshot centered on each image.

CHAPTER 5

The redshift catalogs

In this chapter, we sketch how to take advantage of the *HST* imaging and the MUSE spectroscopic data to build galaxy member and multiple image catalogs. We initially describe the procedure to measure source spectra (Sect. 5.1) and the selection of the cluster members, based on the galaxy spectroscopic and photometric information (Sect. 5.2). We then analyze in detail the families of multiple images, including the new ones we identify (Sect. 5.3). Both the galaxy members and the multiple images are then organized in catalogs, which are used to model A2163 in the method described in Chap. 6. Finally, we also release a further catalog (the table in Appendix A), where we collect the reliable redshifts of all the sources identified as background and foreground objects.

5.1 Extraction of source spectra and redshift measurements

To identify the galaxies belonging to A2163-A1 and the multiple images of the background sources, we need to extract the spectra from all sources in the MUSE FoV and measure their redshifts. In this section, we thus describe in detail these procedures (only introduced in a general way in Sect. 4.3.2), referring to all the following sources within the spectrograph FoV (with few exceptions described from time to time): galaxy members, multiple images, foreground and background objects.

To extract the spectra, we use a simple Python script, whose inputs are a source catalog (a text file) and a MUSE data-cube (a FITS file, already "cleaned" with the reduction procedure described in Sect. 4.3.2). The former is organized in columns reporting a source ID, its coordinates, in degrees, and the radius of the circular aperture from which the spectrum is extracted. For each source, both the center and the radius of the aperture must be set with care. We observe each source in different bands and set the center coordinates to those of the brightest emission peak. The positions we measure are roughly coincident with those in the *HST* photometric catalog, with few exceptions, mostly related to segmentation issues. The latter occur when the same source is indexed more than once, during the automatic procedure used to generate the *HST* catalog; in general, the root of the problem is a relatively extended source size, leading to interpret one single object as a set of contiguous, different sources. We identify all these copies and correct the catalog accordingly (see left panel of Fig. 5.1 for some examples). On the other hand, to set the radius of the aperture, we consider the seeing mean value, since data are taken from different observations. Considering isolated stars in the MUSE FoV, we measure an average value of $0''.6$, which is about the same as the mean seeing calculated from the

ESO portal[1]. Our spectra are thus extracted within 0″.6-radius circular apertures, which we consider a good compromise in the effort to maximize signal-to-noise and minimize source confusion.

With these choices, we extract 230 spectra from the MUSE data-cube. The main script outputs are two files containing, respectively, the source spectra and variance, together with some additional information: the catalog, the data-cube and the source names, the number of extracted pixels, the source positions (in degrees) and the measurement units for the wavelength and the flux values (Å and 10^{-20} erg/s/cm^2/Å, respectively). In addition, a logfile is also generated: it warns about possible discarded sources, for which a spectrum is not extracted. This happens in the case of objects at the edge of the FoV, usually having an insufficient number of readable pixels. For this reason, few sources were discarded, as those in the right panel of Fig. 5.1.

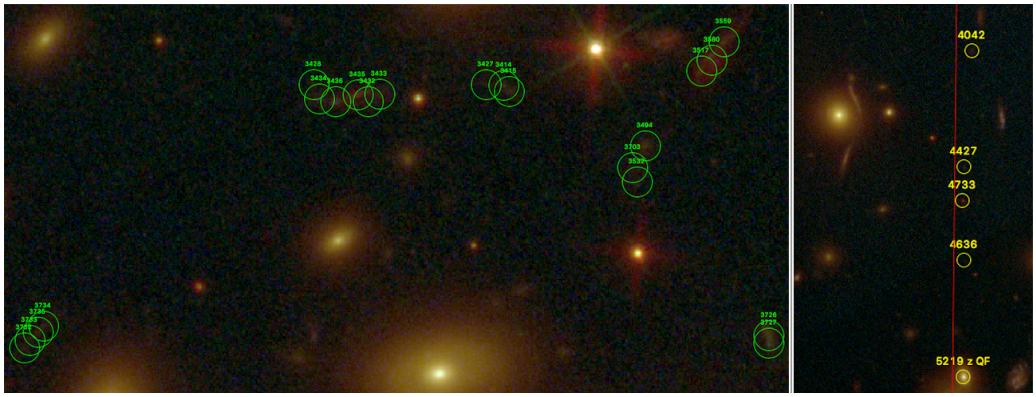


Figure 5.1: Left panel: examples of *HST* catalog sources (in green) which are wrongly segmented. Right panel: some sources (in yellow) discarded because of the small number of readable pixels in the data-cube. The red line is the west edge of the MUSE FoV (its inner region is on the left) in an ideal, conservative representation: the real edge is more jagged, hence including (partially) the pixels within the yellow apertures.

The output files of the code which we use to extract the source spectra are such that they can be directly read by EZ (Garilli et al. 2010), the software we adopt to visualize the spectra and to estimate the source redshifts. As an example, we report in Fig. 5.2 the BCG spectrum of A2163-A1. As expected, it shows the typical spectral features of a bright, elliptical galaxy, with an old, metal-rich stellar population. In particular, at the lowest wavelengths, the HK break, produced by absorption of metals from different elements (such as that of the Ca II K and H lines), clearly indicates the presence of a lot of metals and/or the deficiency of young stars. It is possible to reach this conclusion solely by observing that the break height is very prominent, namely there is no need to measure usual stellar age indicators, such as the ratio between the mean intensity before and after the break[2]. Finally, towards higher wavelengths, several pronounced

[1]The parameter we consider on the ESO website is the "DIMM-Seeing-at-start-time" (the FWHM seeing value measured at the beginning of the observation). There, one can find several information on the tools available to know under which atmospheric conditions the observations are carried out. Other ambient conditions, such as humidity, temperature, pressure, wind speed etc., can be obtained from the ESO Ambient Conditions Database.

[2]In any case, this specific spectrum does not allow for measuring the classical indicator of stellar age,

absorption lines are present, such as Hβ, Mg I, NaD and Hα.

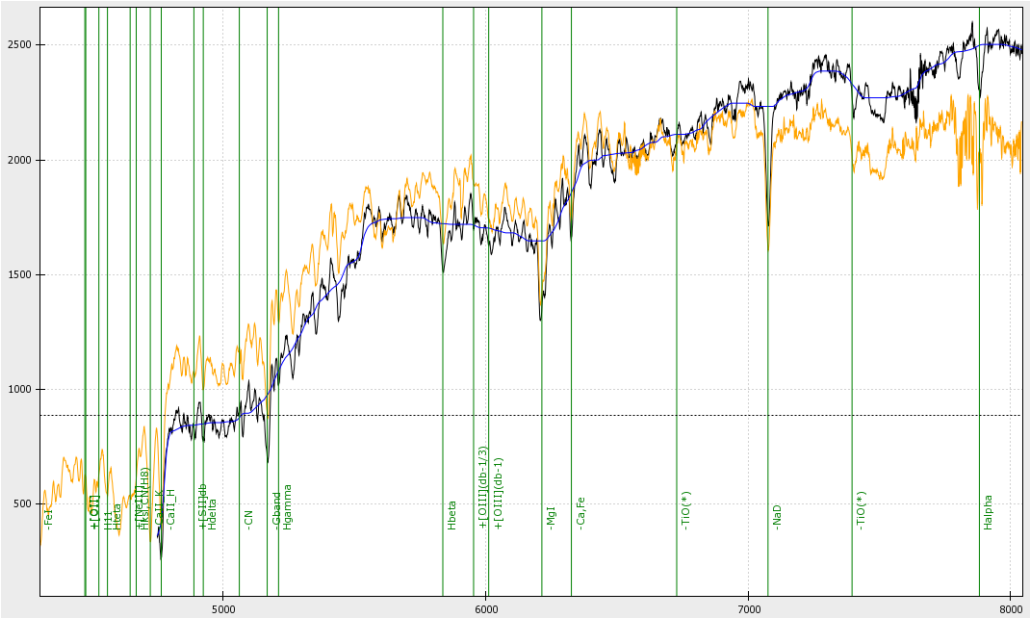


Figure 5.2: Spectrum of A2163-A1 BCG (in black) as visualized in the EZ spectral viewer, after the fit procedure. The vertical axis corresponds to the counts, while the horizontal one to the wavelength, in Å. The orange and the blue curves are, respectively, the galaxy template and the continuum adopted during the fit step; spectral line wavelengths, shifted at the galaxy redshift, are marked in green.

The software EZ (Garilli et al. 2010), or Easy redshift, is an open source program, developed by Marco Fumana and Bianca Garilli. It is devoted to redshift estimation (in an automated way) and belongs to the set of tools created in the PANDORA project by Milan INAF and IASF members[3]. The best redshift value is found by comparing the observed spectrum with a set of spectral templates, in a sequence of operations (the "decision tree"), which mimic the decision path followed by astronomers when assessing a source redshift. Several decision trees coexist in EZ, because these paths usually differ from one dataset to another. One advantage of this tool is the possibility of using different templates provided by a user, so that those more suitable for a particular dataset can be chosen. EZ, thus, provides a high degree of customization, which is important considering the crucial role that spectral templates have during the redshift measurement process. In this thesis, we exploit the templates of the VANDELS survey (McLure et al. 2018), integrated with a further library. The latter, including both low and high-z galaxies, is suited for research projects on galaxy clusters and has been already used in Caminha et al. (2017a) for MACS J0416.1-2403. To accomplish the redshift estimates, in EZ,

namely the ratio of the mean fluxes between two sets of specific absorption lines: those produced by the cyanide molecules in both the red (3750 Å− 3950 Å) and the blue (4100 Å− 4250 Å) regions and highly dependent on the age of the galaxy stars.

[3]More information on PANDORA and on the EZ tool can be found here: `http://cosmos.iasf-milano.inaf.it/pandora`. Unfortunately, due to the lack of manpower and funds to support the PANDORA programs and help-desk, this project has been terminated in March 2011 and the EZ tool is no more freely downloadable from the PANDORA website.

a set of highly configurable and easily extendible algorithms is implemented, through both Python classes and C programs (the latter are also used to increase the computational speed). The software architecture and its algorithms are detailed in Garilli et al. (2010). There, the EZ performances are also evaluated, both on simulated and real data: the redshift measurement success rate is above 95%, for simulated data, and between 70% and 90%, for real data.

EZ has been designed for large spectroscopic surveys and developed particularly to increase the reliability of the results obtained in an automatic way. For example, EZ has been routinely run as part of the reduction process within the VIPERS survey (Garilli et al. 2012). Thus, the results of this tool are thought to be generally used in an unsupervised mode. For this reason, we decide to complement them with visual inspection, in an attempt to possibly improve the redshift measurements: we look for the existence of further emission or absorption lines, unnoticed in the spectral line matching process of EZ, and if there are some, we test whether they match the proposed redshift solution. We also notice sometimes a small shift between the positions of the spectral line peaks and those of the theoretical fitted ones and adjust this by taking as a reference the most secure and strongest spectral line. This translates into a slightly different, but more precise value of the redshift. Moreover, we also make use of the interactive software functions of the EZ graphical user interface; for instance, we use a redshift step of 10^{-3} (or 10^{-4}), set a large z range to be explored and smooth the noisiest spectra (no more than once, to avoid the risk of generating false spectral lines). Our aim, in this phase, is not to perform a complete spectral analysis, but rather to flag each source redshift with a quality flag (see Sect. 4.3.2). To this extent, it is sufficient to identify two or three spectral lines which are secure: we make sure that they are not due to contaminations by galaxies close to the source under inspection, nor to noise effects. The latter, in fact, can be relevant; usefully, in EZ, the variation of each source noise with the wavelength can be visualized together with each spectrum: the alignment between the wavelength axes of both the noise variation and the spectrum makes it easy to recognize and reject the false lines created by the peaks in the noise distribution.

In summary, we have measured the spectra of more than 200 sources in the A2163 core and, at the end of our analysis, we have obtained a sub-sample of 64 spectroscopically confirmed sources, with spectra having $QF \geq 2$: 18 stars, 29 (cluster) galaxies at $z \simeq 0.20$ and 17 (background) galaxies with $z > 0.33$ (with two high-redshift sources at $z \simeq 4.58$ and $z \simeq 4.99$). We report in Appendix A the redshift measurements for all the sources other than cluster galaxy members and multiple images, whose redshift catalogs are instead presented, respectively, in Table 5.1 and 5.2. The latter includes the four images we have discovered (see Sect. 5.3) and consists of 16 multiple images from 4 different sources, with redshifts from 1.16 up to 2.72. This image catalog, together with that of the galaxy members, is used to test five cluster total mass models presented in Sect. 6.3.

5.2 Selection of cluster members

5.2.1 Spectroscopic sample

We initially refer to the automatically-identified sources of the *HST* photometric catalog to have a preliminary list of all the objects within the MUSE FoV. Then, we distinguish between cluster members and non-member galaxies, on the basis of the redshift mea-

surements derived through the procedure of Sect. 5.1. Among the sources with a reliable redshift estimate (i.e., $QF \geq 2$), we identify 29 galaxies as being part of A2163: they are listed in Table 5.1, have a redshift $0.19 < z < 0.22$ and a spatial distribution shown in Fig. 5.3. We also extend the initial list of possible member candidates, including 12 additional sources, which are absent in the *HST* catalog. We identify them through visual inspection of both the *HST* image and the MUSE data-cube and measure their redshifts, looking for further cluster members. Unfortunately, all these sources have very noisy spectra ($QF = 0$), except one, which has a secure ($QF = 3$) redshift measurement of $z \simeq 0.83$. Due to this value, it is classified as a background galaxy (ID 9000 in the table of foreground and background sources of Appendix A).

Table 5.1: Catalog of the galaxy cluster members, with *HST* IDs, celestial coordinates and spectroscopic redshifts with QFs.

ID	R.A. [Deg]	Decl. [Deg]	$z_{\rm sp}$	QF
3355	243.94490	−6.13817	0.1918	2
4431	243.95039	−6.14759	0.1930	3
3206	243.95701	−6.13600	0.1949	3
3623	243.94474	−6.13755	0.1953	3
3461	243.94558	−6.13824	0.1973	3
3518	243.95247	−6.13899	0.1979	3
4659	243.94738	−6.14825	0.1980	2
3820	243.95677	−6.14657	0.1990	3
3760	243.96102	−6.14097	0.1990	3
4825	243.95950	−6.15144	0.1992	2
3686	243.95737	−6.14052	0.1998	3
3397	243.94707	−6.13766	0.1999	3
3824	243.94546	−6.14163	0.2001	3
3829	243.94929	−6.14202	0.2008	3
3822	243.95394	−6.14484	0.2008	3
3480	243.94497	−6.13775	0.2009	3
4051	243.95998	−6.14382	0.2017	3
4101	243.94595	−6.14492	0.2023	3
3570	243.95849	−6.13906	0.2024	3
4544	243.95211	−6.14796	0.2029	2
2878	243.94753	−6.13413	0.2031	3
4104	243.94712	−6.14498	0.2032	3
3821	243.95240	−6.14314	0.2039	3
3358	243.95123	−6.13721	0.2049	3
4504	243.95377	−6.14237	0.2079	3
3933	243.95450	−6.14216	0.2079	3
3543	243.96174	−6.13685	0.2120	3
4461	243.95995	−6.14135	0.2160	3
4890	243.95647	−6.15021	0.2191	3

In detail, these 29 spectroscopic members, which will be included in the mass models of the cluster, are selected through a fit of their redshift values with a Gaussian distribution. The latter has mean and standard deviation values of $\langle z \rangle \sim 0.201$ and $\sigma_z \sim 0.006$, respectively. This corresponds to a cluster velocity dispersion of $\sigma_v \sim 1450$ km s^{-1}.

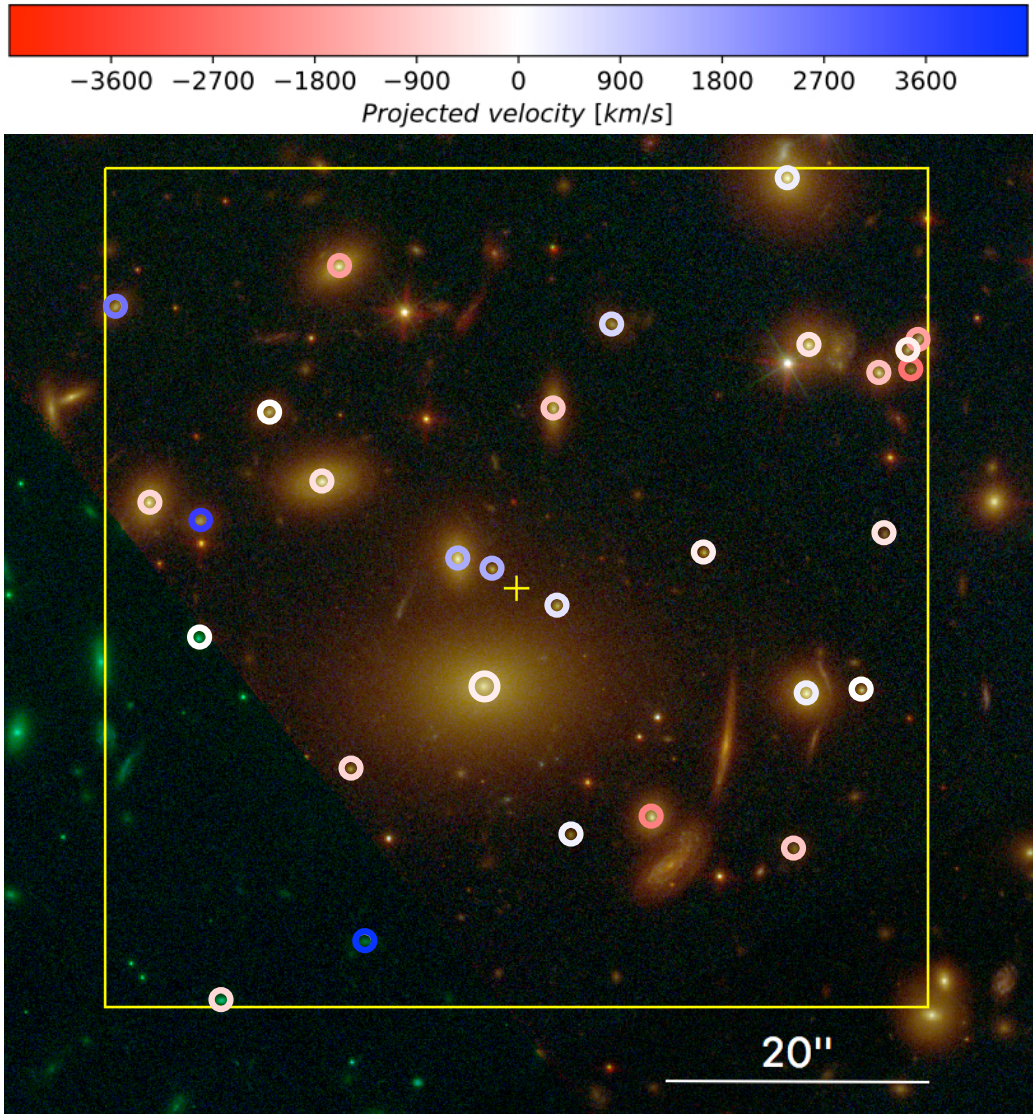


Figure 5.3: Color composite image of Abell 2163 from the *HST* data. The overlaid MUSE pointing (yellow box) is about one arc minute across and is centered on the yellow plus sign position. Circles mark the location of the 29 spectroscopically confirmed galaxy members and they are colored according to the galaxy velocities (relative to the cluster mean redshift). The range of the values is shown on the bar on the top.

From the member redshifts, we calculate the galaxy velocities relative to the cluster mean redshift. They are shown in Fig. 5.3 and are derived as in Harrison (1974), namely taking into account the main factors which determine the observed galaxy redshift (z). Specifically, the latter is determined by the contribution of separated motions (corresponding to different redshifts): that of the galaxy itself ($z_{\rm G}$), the motion of the center-of-mass of the cluster ($z_{\rm CM}$), that of the observer (z_0), and finally that due to the Hubble expansion ($z_{\rm H}$). Harrison (1974) shows that, neglecting the motion of the clus-

ter's center of mass ($z_{\rm CM} \sim 0$), all the remaining redshifts are connected by the following relation:

$$1 + z = (1 + z_{\rm G})(1 + z_0)(1 + z_{\rm H})\,. \tag{5.1}$$

We can then assume an isotropic velocity distribution for the cluster galaxies (so that $\langle z_{\rm G} \rangle = 0$) and that the Hubble expansion is the same for all cluster members (hence $\langle z_{\rm H} \rangle$ has the constant value $z_{\rm H}$). Averaging both sides of Eq. (5.1) and including, there, these conditions leads to

$$1 + z_{\rm H} = \frac{1 + \langle z \rangle}{1 + z_0}\,. \tag{5.2}$$

Inserting Eq. (5.2) in (5.1) and considering the relationship between the galaxy (rest-frame) velocity and its (rest-frame) redshift, $v = cz_{\rm G}$ (in the non-relativistic case), the following expression is found for the i-th galaxy

$$v_i = \left(\frac{1 + z_i}{1 + \langle z \rangle} - 1 \right) c\,. \tag{5.3}$$

Finally, we remark that the spectroscopically confirmed foreground and background perturbers are excluded from the lensing models, due to their minor lensing contribution: in the cluster core, we find background galaxies with $m_{\rm F814W} > 19.5$ mag and do not observe foreground galaxies (see the table of Appendix A). In particular, the background galaxies closest to A2163 ($0.20 < z < 0.41$) are very faint, with $m_{\rm F814W} > 24$ mag. Chirivì et al. (2018) found that this choice does not affect significantly the reconstruction of the total, projected mass profile. Finally, including these *los* structures would require a multi-plane lensing analysis, which is not yet implemented in the software `Lenstool` (Jullo et al. 2007) we use to model A2163-A1.

5.2.2 Photometric sample

As detailed in Sect. 4.2.2, we use archival *HST* photometric data in the optical and near-infrared bands ($0.36 - 1.70$ μm), taken with 7 filters incorporated in the ACS and the WFC3 cameras. They are the same filters adopted in the HFF campaign and we list them in Table 4.2. We have already emphasized in Sect. 4.2.1 the characteristics of these two instruments, from which the imaging data of our work are taken. Here, we describe how we perform a photometric selection of galaxy members, as a complement to the above-mentioned spectroscopic catalog.

We modify the original *HST* ACS/WFC3 catalog to create a new version. There, each set of wrongly segmented sources is aggregated and considered as corresponding to a single one. We also include in the new catalog the (secure) spectroscopic information, namely the available redshift values with $QF > 1$ (also for secure non-member galaxies). We then assign to each source one of three tags: *member*, *unknown*, *non-member*. The first two flags identify the sources which must be included in the selection process, as confirmed spectroscopic members and as possible member candidates, respectively; in detail, the second category consists of sources with no redshift measurements or having a z with $QF < 2$. The *unknown* tag also refers to the secure non-member galaxies. They are then eliminated from the member sample by comparing their redshifts with Δz, the range which defines the cluster membership. Thus, these sources are not directly identified as *non-members*. This choice has a specific function in the code used for the photometric member selection: to allow to easily change the redshift range instead of the input catalog. Finally, the last flag refers to objects which must be excluded, that

is they are not considered as possible member candidates: stars and few objects whose photometry is clearly and strongly contaminated by other bright galaxies are part of this sample. Such a new catalog, containing 5541 objects, is then used as the input for a code, implementing a novel method based on the extreme deconvolution (Bovy et al. 2011) of galaxy color distribution.

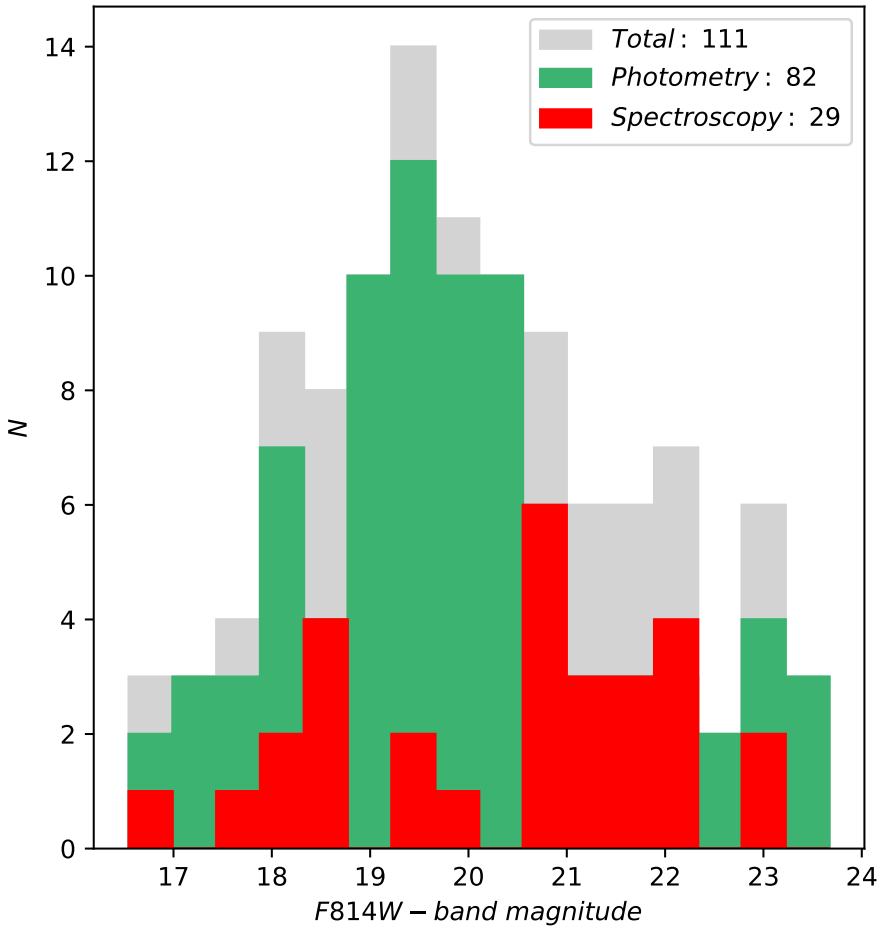


Figure 5.4: Magnitude distribution of the cluster members, having $m_{\mathrm{F814W}} < 24$ and a member probability $P > 95\%$; the whole sample is shown as a gray histogram, while spectroscopically confirmed galaxies (29) and photometrically selected members (82) are in red and green, respectively.

Specifically, we define a six-dimensional color space, within which the 7-band galaxy dataset of the whole *HST* catalog is analyzed. We perform a deconvolution of the color distribution of two populations of galaxies, namely the cluster members and the field galaxies, respectively inside and outside the redshift range used to set the membership criterion (i.e., z between 0.19 and 0.22). The deconvolution, which takes into account both the photometric errors and possible incomplete measurements (where one or more bands are missing), produces for each population a representation of the colors distri-

bution of member and field galaxies as a Gaussian Mixture Model (GMM). Since the two populations follow two different GMMs, we can calculate the membership probability for each galaxy on the basis of this probabilistic model. To this purpose, we apply Bayes'theorem, using a 50% prior that the galaxy is a member galaxy. Finally, in order to foster cluster galaxy purity over completeness, especially for the reddest sources, we choose a member probability threshold of $P > 95\%$, maximizing the inclusion of the most massive galaxies, which affect more the lens model (see also Sect. 3.3.1 in Grillo et al. (2015)). We report in Fig. 5.4 the magnitude distribution of cluster members, obtained with such a choice and with a limiting F814W magnitude of 24 mag.

With this method, 82 additional photometric members (mostly outside the MUSE FoV) were integrated in the spectroscopic catalog to constitute the final, more complete sample of 111 galaxies used in the lensing analysis.

5.3 Multiple images

The identification of lensed sources is carried out adopting different strategies: we consider all A2163 information available in the literature in light of the indications of our new data; we perform an inspection of both the *HST* images and the MUSE data-cube and subtract the contaminant light of bright cluster members, when the image candidate seems to be polluted by them; finally, we check further multiple images, as predicted from our strong lensing models (see Sect. 7.1). In what follows, we describe all these phases, integrated in the description of our first model (RUN 1), that is a initial attempt at the cluster lens model. The characteristics of all the multiple images to which we refer can be found in the final catalog (Table 5.2).

Table 5.2: Catalog of multiple images, with ID, celestial coordinates and spectroscopic redshifts with the related QFs; we also report the ID of the reference work in the literature (C18). From the top to the bottom, the horizontal lines separate, respectively, the multiple images of families F1, F2, F3, and F4.

ID	R.A. [Deg]	Decl. [Deg]	$z_{\rm sp}$	QF	C18
1a	243.95273	−6.14646	2.723	3	1.1
1b	243.95345	−6.14722	2.723	3	1.2
1c	243.95497	−6.14588	2.723	2	-
1d	243.95322	−6.13895	2.723	2	1.3
2a	243.95241	−6.14636	-	0	2.1
2b	243.95339	−6.14749	2.264	1	2.2
2c	243.95506	−6.14568	2.723	1	-
2d	243.95312	−6.13912	2.264	1	2.3
3a	243.95573	−6.14328	2.389	3	3.1
3b	243.95442	−6.14969	2.389	1	3.2
3c	243.95381	−6.14108	2.389	1	3.3
3d	243.95270	−6.14427	2.389	1	-
4a	243.94685	−6.14437	1.164	3	4.1
4b	243.94668	−6.14476	1.164	3	4.2
4c	243.94695	−6.14592	1.163	3	4.3
4d	243.94738	−6.14494	1.164	1	-

We measure the redshifts of all lensed systems selected by C18, spectroscopically confirming 8 multiple images belonging to three of the four families there identified (F1, F2, F3, F4, following the notation in C18). We then create a first model, based on a catalog which is a combination of newly and previously-detected multiple images (see RUN 1 in Table 7.1): they are 10 in total and the corresponding MUSE 1D spectra and *HST* snapshots for 7 of them are shown in Fig. 4.9. Specifically, at this stage, we do not use the images of family F2 of C18, because their redshift measurements have a low quality flag ($QF < 2$), and include families F3 and F4. Finally, F1 images were selected based on the following considerations. We identify three new candidates, visually found in the *HST* data. They are labeled as 1c, 2c and 3d, and lie in the proximity of the BCG, at an approximate distance of 5″($\sim$ 17 kpc). Since their photometry is contaminated by that of the BCG, only image 1c was classified as likely ($QF = 2$). As a result, we replace the less reliable image of F1 (from C18) with this new one (1c) and do not include 2c and 3d ($QF < 2$) in RUN 1.

We then recover a further image (4d) also for family F4: this is a very interesting multiple image system generated by the combined effect of the cluster and a cluster member gravitational potentials, a situation which is not rare in dense galaxy cluster environments (see, e.g., Grillo et al. 2014; Parry et al. 2016; Caminha et al. 2017a; Meneghetti et al., in prep.). The galaxy acting as a strong lens is a bright elliptical cluster member, around which a background source is distorted into some arclets (4a, 4b, 4c) and a more compact image (4d). This cluster member is shown in the right panel of Fig. 5.5 and can be easily identified in Fig. 5.6, where it is surrounded by the F4 lensed galaxies, reported in red.

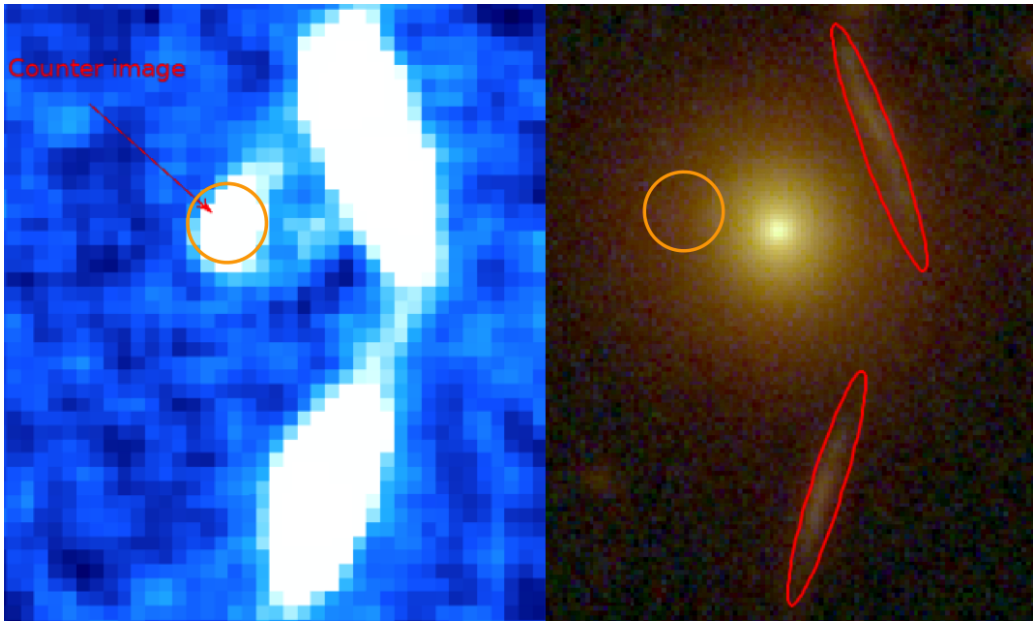


Figure 5.5: Comparison between the *HST* image (right panel) and the MUSE FoV (left panel) region around the galaxy member generating the multiple images of family F4. The new image (4d) position in marked by the orange circle, while the two arcs of F4 are within the red curves.

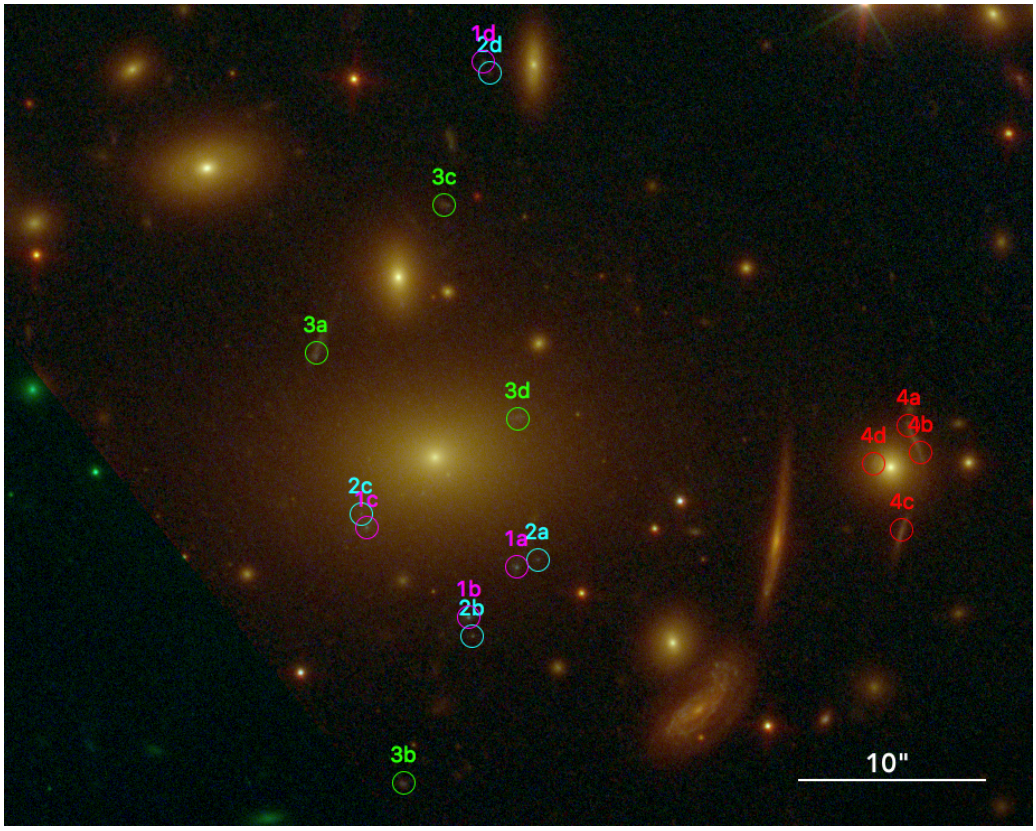


Figure 5.6: Color composite image illustrating the distribution of the multiple images included in our fiducial model.

The presence of systems of this kind helps in constraining the total mass distribution of the sub-halo associated with the cluster galaxy and, as a consequence, can improve the mass reconstruction of cluster itself. A detailed analysis of the system F4 can be found in Bergamini et al. (in prep.), where several galaxy-galaxy scale strong lensing systems in different galaxy clusters are studied. Before its identification, 4d has been thought to lie very close to this bright galaxy member, hence we have eliminated pollution effects, subtracting convenient combinations of *HST* single-band images from the original one. Unfortunately, the image 4d is not visible using *HST* data, so we try to identify it in the subtracted MUSE data-cube (successfully). In detail, through the MUSE Python Data Analysis Framework (MPDAF, proceedings of ADASS XXVI, 2016), we sum spatial pixels of the cube over the wavelength interval of the [O II] emission line doublet, which is very prominent in the spectra of the other images of the same family. From this interval, we also eliminate contamination effects by subtracting the background emission taken from two cube slices, below and above the [O II] wavelength range. We also include 4d in the first catalog used in this preliminary stage. Hence, our first model (RUN 1) consists of three lensed systems, with at least one counter image per family, having a secure redshift (they are listed in Table 7.1).

Notice that other images, excluded from RUN 1, will be considered, at a later stage, in our fiducial strong lensing model, based on its predictions and on the findings of the computational runs described in Sect. 7.1. In particular, the family F2 (a four-image system) is added to our final catalog, so that the latter consists of a total of 16 multiple images belonging to 4 families. The complete set of multiple images has a F814W observed magnitude (when present) of $m_{F814W} \simeq 24 - 29$ mag and spans a redshift range between 1.16 and 2.72. The images of the same family have, in each model, an equal redshift value, that with the highest QF. They are well represented by point-like objects (with the exception of the image 3a and the two F4 arcs) and cover the cluster core, targeted by the MUSE observations. The properties of our final catalog of 16 multiple images are reported in Table 5.2 and their positions are shown in Fig. 5.6.

CHAPTER 6

Strong lensing modeling of Abell 2163

This chapter presents an overview of the kinds of software for gravitational lens modeling (Sect. 6.1) and then focuses on that we adopt (`Lenstool`), whose description is reported in Sect. 6.2 (Jullo et al. 2007). Our method consists of a preliminary identification and differentiation of the A2163 mass components (Sect. 6.3.1) and of a successive implementation of the total mass model in the stage of run planning (Sect. 6.3.2).

6.1 Gravitational lens modeling software

Traditionally, two general categories of software are considered to model a lens via gravitational lensing: those based on *parametric* and on *non-parametric* models. The choice between the two is still an ongoing debate, since both have their own pros and cons. For the same reason, hybrid methods have been recently developed to couple their advantages, modeling the galaxies and the large-scale mass distribution through, respectively, parametric and non-parametric models (e.g., Diego et al. 2015). Before coming to a brief description, it is worth mentioning that not only is using the adjective "non-parametric" for models, in general, puzzling (all models are somehow parameterized), but it is even misleading: non-parametric lens models, in fact, are often those with the largest number of parameters. To avoid confusion, the latter are also called "free-form" models.

As the name already suggests, the parametric approach takes advantage of analytic functional forms described by a (limited) number of parameters to model the lens (e.g., the SIS or the Navarro, Frenk, and White profile) and to fit the observations. In other words, these methods use a simple parameterization of the lens mass distribution of the cluster components. Even in cases where they reach a quite high level of sophistication, parametric models are only as flexible as the functions on which they are based. The parameter space is then explored in the effort to best reproduce the multiple image observables.

On the other hand, in the free form approach, usually, the lensing observables can be mapped onto a mesh which covers the cluster and is then converted into a pixelized mass distribution. The observables are finally connected to the lens potential with different methods. In the literature, a variety of implementations can be found (e.g., Diego et al. 2007; Merten 2016). Free form methods allow more freedom in the mass distribution than parametric models, but have also some severe disadvantages. The *perfect* fit often reached by the former is at the expense of obtaining unphysical lens mass reconstruction results. Moreover, different models may fit equally well the same data, due to the large number of free parameters with respect to the number of constraints (Jullo & Kneib 2009).

6.2 Lenstool

The program Lenstool (Jullo et al. 2007) is commonly used by the scientific community to model, in a parametric or hybrid approach (Jullo & Kneib 2009; Lefor et al. 2013), the total mass distribution of gravitational lenses: galaxies and clusters, in both strong and weak regimes. It is publicly available and the latest version can be downloaded from the Lenstool project homepage[1], where the installation instructions and the documentation files are available.

Lenstool has a terminal user interface and operates using text commands within some input files. In our case, namely for non-basic strong lensing studies of galaxy clusters, at least three files are necessary. Throughout this section, we briefly detail them, without describing the other input files of the software. For further information on Lenstool, we refer the reader to Appendix B. Moreover, we explain how to extrapolate, in practical terms, some of our results from the software output files; in order to make the text more fluid, each brief description is reported in the footnotes of Chap. 7, together with other technicalities.

Model parameter file (*.par* extension). It consists of all the basic choices and information for the model: the reference system (usually centered on the BCG), the chosen cosmological parameters, the number and the features of all the potentials used and the priors on the parameters of the corresponding mass profiles. In this file, it is also possible to set the optimization method and its characteristics (see Sect. 6.3.2), the reference to the image and the cluster member catalogs (input files) and the desired output files and their features, such as the resolution of the output images. In particular, under the section defining the potential, the profile parameters to fix and those to vary in the optimization step must be indicated. This allows the software to define the parameter-space to explore and the kinds of priors: uniform (the parameter values are equally likely, within a certain interval) or Gaussian (the most likely ones are those close to the mean). These priors are completely described setting the lower and the upper limits of the parameter values, for the former, and the mean and the standard deviation, for the latter. The parameters vary in a range of values with an increment (or *step*) defined by the user.

Image catalog file (*.cat* extension). It is the catalog of multiple images and arcs, whose positions are used by Lenstool as constraints to optimize the model, but also to predict possible counter-image locations. The file lists a few useful image descriptors in the following order: the image ID and its absolute coordinates (R.A. and Decl.) or those relative to a specific reference, the source redshift (if available[2]) and the image magnitude.

Member catalog file (*.cat* extension). This file is mandatory if the galaxy-scale halos are connected by scaling relations (Sect. 6.3.2), like in our analysis. As an alternative, each of them can be modeled separately in the model parameter file. The member catalog file sorts the galaxies by magnitude, with the most luminous source at the beginning (i.e., the BCG). It has a format similar to the image catalog and collects analogous information: the ID and the absolute coordinates of the galaxies, their magnitude and luminosity (the latter can be omitted).

[1] https://projets.lam.fr/projects/lenstool/wiki#

[2] Sometimes this quantity is unknown or has a non-secure measurements and can be included in the lensing model as a free parameter, assigning to it a null value in the image catalog file.

In `Lenstool`, the mass distribution of the lens is modeled with a parameterized mass profile and is constrained by solving the lensing equation for all multiple images and knowing the distance of the lens and that of the source. Specifically, the lens equation can be written as a function of the lensing potential, substituting the expression 1.11 of the scaled deflection angle in Eq. (1.7), so that

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \nabla_{\boldsymbol{\theta}} \psi(\boldsymbol{\theta}, \boldsymbol{p}) \ , \ \text{with} \ \psi(\boldsymbol{\theta}, \boldsymbol{p}) = \frac{2}{c} \frac{D_{LS}}{D_S} \phi(\boldsymbol{\theta}, \boldsymbol{p}) \ . \tag{6.1}$$

Here, the second equation relates the lensing and the Newtonian potential (at the image position, $\boldsymbol{\theta}$), respectively, $\psi(\boldsymbol{\theta}, \boldsymbol{p})$ and $\phi(\boldsymbol{\theta}, \boldsymbol{p})$. The latter is calculated by constraining the parameter vector $\boldsymbol{p}$ with all the multiple image positions, $\boldsymbol{\theta}$, and the lens equation; as a result, also the mass distribution of the lens is found.

The optimization of the model parameters follows two possible approaches, both implemented in `Lenstool`, as well as in several other strong lensing codes: the source plane and the image plane optimizations. In the first mode, the multiple images of each family are collectively projected onto the source plane, where they should ideally mark a single position of the same background source. However, since, in general, this is not the case, because of astrometric errors or because of the inadequacy of the lens model, there will be slightly different source positions, $\boldsymbol{\beta_i}$, to consider; they must be compared with a single position, $\boldsymbol{\beta}$ (treated as a further model parameter), in a source-plane minimization process using, for example, a χ^2 fit statistic. In the image plane optimization, the lens equation is then inverted, namely the source positions are mapped from the source back to the lens plane. Here, the positions of the predicted multiple images, $\boldsymbol{\theta_i}(\boldsymbol{\beta})$, must be compared with those of the observed ones, $\boldsymbol{\theta_i}$, through an image-plane minimization measuring the goodness of the fit.

In the source plane optimization, there is no need to invert the lens equation, hence the overall fit process is less time-consuming. This advantage is much more evident when many mass components and images are included in the model. On the other hand, optimizing in the image plane is considered more statistically sound precise (Keeton 2001), but at the expense of a longer computational time. Further details on the optimization processes implemented in `Lenstool` can be found in Jullo et al. (2007) and Jullo & Kneib (2009). There, a specific strategy is suggested to reduce the computational time during the image plane optimization. It is based on avoiding as much as possible several acceptance/rejection steps in the parameter-space exploration: the best-fit region is restricted using a preliminary source-plane optimization and then faster constrained in a following image plane optimization. In this way, this region must be just refined in the lens plane, rather than being explored from scratch. In the same paper, it is stated that this method can decrease the duration of the best-fit model convergence up to about eight times. However, since our `Lenstool` runs, with model parameters optimized on the image plane, have a reasonable duration, we decide to favor a traditional (image plane) approach (see Sect. 7.1).

6.3 Method description

6.3.1 Cluster mass decomposition

We model the total projected mass density in the core of A2163 as a combination of parametric isothermal halos. Since the cluster mass components are distributed on different scales, we distinguish these halos in two main families, on large- and small-scales. The first category accounts for all the diffuse mass components, approximately: mainly DM (Sect. 2.2), plus the hot ICM (Sect. 2.3) and a further large-scale (minor) constituent, the ICL (Sect. 2.5); on the other hand, each sub-halo consists of a small DM clump, plus the galaxy baryonic matter within it, and for this reason they are traced by the cluster members. Even if the sub-halos (mostly DM) contribute less to the total cluster mass compared to the cluster-scale component (see Fig. 7.7), to include them in the model is fundamental to reproducing the positions of the observed multiple images and to reach high-precision results (Kneib et al. 1996).

Regarding the number of large-scale halos, in general, a cluster mass model includes a single cluster-scale component (or a few, depending on the complexity of the cluster and especially on the data available), since the total number of free parameters must be confronted with the number of observables. In the very first approach to a mass model, cluster classification (see Sect. 2.1.4 and Table 2.2) directly suggests a preliminary model guess: the mass distribution of a regular-type cluster is likely characterized by the presence of a single large-scale halo, approximately centered on the only BCG; on the contrary, for a non-regular cluster, with a couple or more matter sub-clumps, two or more diffuse components should be taken into account. However, this possibilities cannot be thought of as rigid rules (especially for the latter cluster category), because further large-scale components can also exist, e.g., dense galaxy clumps, like in the case of the refined model of the cluster MACS J0416.12403 (Fig. 1.3), presented in Caminha et al. (2017a).

Nowadays, the aforementioned cluster mass decomposition is a standard of lens modeling in the parametric approach, since the arrangement of potentials at different scales provides a higher resolution for the model, preserving the advantages of using a limited number of parameters (compared to non-parametric models). On the other hand, this strategy can be extended to the hybrid approach, to include intermediate-size halos which allow sharper contrast in higher density cluster regions. An example is the study of the cluster Abell 1689 in Jullo & Kneib (2009), who develop, through the software `Lenstool`, a hybrid, multi-scale model. The latter combines a mass map grid based on radial basis functions and 60 galaxy-scale halos described by scaling relations (to be robust against over-fitting). The use of such a flexible, multi-scale method, which includes intermediate-scale halos, can lead to the improvement of the fit quality: in the case of this study, it results in halving the offset between the positions of predicted and observed multiple images, compared to previous analyses.

Before associating a specific mass profile to the mass components of A2163-A1, a first step is to consider a convenient reference system: we define the cluster center as marked by the BCG and from it we assign to every member and image a projected radius. The latter is calculated by multiplying the angular separation with the angular-diameter distance obtained by our conventions on the cosmological parameters (reported at the end of the thesis introduction). Then, we assign an isothermal mass profile to all parametric

halos. As in Bonamigo et al. (2017), for both mass components we refer to a dual Pseudo-Isothermal Elliptical (dPIE; Elíasdóttir et al. 2007) profile, chosen because it demonstrates a very good profile to reproduce the positions of the observed multiple images, with high-precision (Caminha et al. 2017a). We have already presented the dPIE profile in Sect. 1.2.3, thus, in what follows, we only report the basic quantities associated with it, namely the surface mass density, Σ, and the projected radius with ellipticity, R_ε:

$$\begin{cases} \Sigma\left(x,y\right) = \dfrac{\sigma_0^2}{2G}\,\dfrac{R_\mathrm{T}}{R_\mathrm{T}-R_\mathrm{C}}\left(\dfrac{1}{\sqrt{R_\mathrm{C}^2+R_\varepsilon^2}}-\dfrac{1}{\sqrt{R_\mathrm{T}^2+R_\varepsilon^2}}\right)\ , \\ R_\varepsilon^2 := \dfrac{x^2}{\left(1+\varepsilon\right)^2}+\dfrac{y^2}{\left(1-\varepsilon\right)^2}\ , \\ \varepsilon := \dfrac{1-q}{1+q}\ . \end{cases} \tag{6.2}$$

Here, σ_0 is the central velocity dispersion, R_T and R_C are the truncation and core radii, respectively; R_ε is the projected radius adjusted to take into account an ellipticity parameter, ε, which is defined through the minor-to-major-axis ratio, q.

The dPIE profile is also a very general and convenient choice: other profiles can be traced back to it, such as that presented in Kassiola & Kovner (1993) (the Pseudo-Isothermal Elliptical Mass Distribution, PIEMD), which is recovered when $R_T \to \infty$. We take advantage of such a possibility, stressed in Bonamigo et al. (2017), by setting the dPIE parameters according to the features of the specific mass component we want to represent. In this sense, our cluster-scale and galaxy-scale halos are modeled with slightly different isothermal profiles. Specifically, the dPIE profile has, in general, seven free parameters: the two centroid coordinates, σ_0, R_T, R_C, ϵ and the position angle, θ. For the diffuse component only, we fix the value of the truncation radius to infinity, because strong lensing data are not sensitive to this parameter, which is thus fixed at an arbitrary large value (typically 1500 kpc). We keep all the other parameters free and this implies a total of six free parameters for the cluster-scale halo. On the other hand, each cluster member is modeled with a spherical dPIE profile, having a vanishing core radius and centered on its luminosity peak, leaving a total of two free parameters.

The parameterization of each single sub-halo is not possible, due to their large number compared with that of the multiple images. In fact, the number of degree of freedom (*dof*) of a model is calculated as the difference between the total number of observables and that of the free parameters. Let us consider n_H cluster-scale halos, n_h sub-halos, n_S background sources and n_{im} multiple images (in total, namely from all the families). The number of observables (when only the image positions are considered as such) is thus $2n_{im}$, where the factor two accounts for the image spatial coordinates. Part of the total free parameters is given by the $2n_S$ coordinates of the background sources (in an image-plane optimization), while the remaining ones depend on the mass halos: if each cluster-scale component and each sub-halo have, respectively, $n_{p,H}$ and $n_{p,h}$ free parameters, the number of degrees of freedom of the model is thus

$$dof = 2n_{im} - \left(2n_S + n_H n_{p,H} + n_h n_{p,h}\right). \tag{6.3}$$

For our models (or for those with the same choices of the mass component free parameters), $n_{p,H} = 6$ and $n_{p,h} = 2$ (the truncation radius and the central velocity dispersion), so

$$dof = 2n_{im} - (2n_S + 6n_H + 2n_h)\,. \tag{6.4}$$

In the case of A2163, we consider a single large-scale halo and the maximum number of multiple images is 16, from 4 different background sources (Table 5.2). Considering these information and Eq. (6.4), it is thus easy to understand that, in order to avoid an under-constrained model, the sub-halos cannot be modeled individually. Therefore, they must be considered either globally or, at most, as an ensemble of very few (n_{set}) sub-sets: among the latter, one of them usually includes the majority of the sub-halos and the remaining contain one or few galaxy-scale halos which it is convenient to model separately (we follow this procedure in our fiducial model, see Sect. 7.2). In this way, if only two free parameters are assigned to the entire set of sub-halos (or to each galaxy-scale sub-sets), the new *dof* expression is

$$dof = 2n_{im} - (2n_S + 6n_H + 2n_{set})\,, \tag{6.5}$$

where we have substituted the number of sub-halos (n_h) with that of the sub-sets in which they are subdivided (n_{set}). This is the relation used, for example, to derive the *dof* of all the models we test[3] and reported in Table 7.1.

The above reasoning highlights the necessity to include in the model the cluster member mass contribution in an alternative fashion, making the number of constraints comparable to that of the free parameters. Following a standard procedure in cluster strong lensing analyses, we reduce the number of free parameters associated with the cluster members, assuming two scaling relations:

$$R_{\mathrm{T},i} = R_{\mathrm{T,g}} \left(\frac{L_i}{L_{\mathrm{g}}}\right)^{0.5}\,, \tag{6.6}$$

$$\sigma_{0,i} = \sigma_{0,\mathrm{g}} \left(\frac{L_i}{L_{\mathrm{g}}}\right)^{0.35}\,, \tag{6.7}$$

where $R_{\mathrm{T},i}$, $\sigma_{0,i}$ and L_i are, respectively, the values of the truncation radius, central velocity dispersion and F814W luminosity of the i-th sub-halo; $R_{\mathrm{T,g}}$, $\sigma_{0,\mathrm{g}}$ and L_{g} are the same quantities for a reference galaxy, which we identify with the BCG ($m_{F814W} \simeq 16.55$ mag). We choose these scaling relations because they reproduce the variation of the total mass to light ratio M/L with luminosity observed in early-type galaxies, known as the tilt of the fundamental plane (Faber et al. 1987; Bender et al. 1992). These equations translate into only two free parameters describing the overall properties of all the small-scale mass components, $R_{\mathrm{T,g}}$ and $\sigma_{0,\mathrm{g}}$.

[3] Except for one of them (RUN 3), in which the redshift of an image family is also left free, decreasing of a further unit the *dof* value.

6.3.2 Run setup

We infer our final mass model of A2163 through the minimization of the distances between the positions of the observed ($\boldsymbol{\theta}^{\rm obs}$, with uncertainty $\sigma^{\rm obs}$) and model-predicted ($\boldsymbol{\theta}^{\rm pred}$) multiple images. To do that, we use the function

$$\chi^2(\boldsymbol{p}) := \sum_{j=1}^{N_{\rm F}} \sum_{i=1}^{N_{{\rm Im},j}} \left(\frac{|\boldsymbol{\theta}_{i,j}^{\rm obs} - \boldsymbol{\theta}_{i,j}^{\rm pred}(\boldsymbol{p})|}{\sigma_{i,j}^{\rm obs}} \right)^2 , \tag{6.8}$$

where the subscripts i and j refer, respectively, to the multiple images, in total $N_{{\rm Im},j}$ (for the j-th family), and the corresponding family, in total $N_{\rm F}$; $\boldsymbol{p}$ is the vector grouping all the model parameters (once the cosmological ones are fixed).

As in Caminha et al. (2016), to quantify the accuracy and the precision of the model, we refer also to the root-mean-square (rms) value of the distances between the observed and model-predicted positions of the multiple images. This quantity is independent of the value of $\sigma^{\rm obs}$ and is defined as

$$\delta_{\rm rms}(\boldsymbol{p}) := \sqrt{\sum_{j=1}^{N_{\rm F}} \sum_{i=1}^{N_{{\rm Im},j}} \frac{|\boldsymbol{\theta}_{i,j}^{\rm obs} - \boldsymbol{\theta}_{i,j}^{\rm pred}(\boldsymbol{p})|^2}{N}} , \tag{6.9}$$

where N is the total number of multiple images.

As anticipated, we adopt the software `Lenstool` (Jullo et al. 2007), which implements the dPIE mass profile described by Eq. (6.2) and Bayesian Markov Chain Monte Carlo (MCMC) techniques to efficiently explore the posterior distribution of each parameter. We use them to compute the statistical errors and the correlations among the model parameters (Sect. 7.2). Moreover, we choose relatively large uniform priors, setting a conservative range of variation for the parameter values. Every run is finally conducted until convergence, using more than 10^5 points to sample the posterior probability distribution of the parameters.

At least two runs (both using image plane optimization) are conducted for each model. In a preliminary run, we tune the uncertainty on the position of the multiple images, to take into account the impact of factors influencing this quantity, such as the clumpiness of the DM distribution in the cluster, the presence of mass interlopers between lensed background sources and the observer and the limitations of parametric mass models (Jullo et al. 2010; Host 2012). Specifically, an initial uncertainty of $0''.10$ is used[4], representing about 2 pixels of the *HST* images (i.e., $0''.06$); then, to include the effects mentioned above, we increase this starting reference value in a second `Lenstool` run in order to have a minimum χ^2 value comparable with the number of *dof*.

Finally, the parameter values of the most probable model (i.e., our *fiducial model*, Sect. 7.2) are used to obtain a number of results, such as the projected, total mass profile in the cluster core.

[4] except for the positional error of image 4d, detected only in the MUSE cube; we double it ($0''.20$) because of the different spatial resolution of MUSE data compared to the *HST* images.

CHAPTER 7

Results and discussion

In this chapter, we detail a number of computational runs (Sect. 7.1), whose aim is to explore different theoretical models of the cluster mass distribution. The latter are then compared and ranked, in order to obtain a fiducial model from which we extract our final results, which are compared with the literature in Sect. 7.2. Finally we summarize our findings and draw concluding remarks and comment on the future directions that might be investigated using the results described in this thesis (Sect. 7.3).

7.1 Computational runs

We perform some preliminary runs to gain insights into the computational time required to optimize a model with `Lenstool`. In this regard, the description of our findings can be better understood introducing few considerations about the `Lenstool` optimization process.

It is implemented in a two-step fashion, through a *burn-in* and a *sampling* phase: both are based on a Bayesian MCMC method and consist in exploring the parameter space. This distinction is used to isolate the data corresponding to the initial iterations. The optimization, in fact, starts from the burn-in phase, in which all the data of the iterations at the beginning of an MCMC run are collected. Then, the burn-in period is followed by the sampling phase. At the end of each phase, two different output files are generated[1], but only that corresponding to the second phase is used to derive all the successive results. Throwing away all the initial data is a common choice in MCMC techniques and the `Lenstool` software implements this practice as explained.

Compared to the first software version (where only a maximum likelihood method was implemented), the recent Bayesian MCMC method to optimize parameters uses a different sampler for the probability density function (PDF) of the model parameters, which avoids being trapped in local minima, more efficiently than before. Specifically, `Lenstool` employs a *re-written* Bayes theorem by the following formula:

$$P(\boldsymbol{p}|D,M) = \frac{P(D|\boldsymbol{p},M)^{\lambda} P(\boldsymbol{p}|M)}{P(D|M)}\,. \tag{7.1}$$

Here, indicating with $\boldsymbol{p}$, D and M, respectively, the parameters, the observed data and the assumed model, $P(\boldsymbol{p}|D,M)$ is the posterior PDF, $P(D|\boldsymbol{p},M)$ is the likelihood of getting the data given the model parameters, $P(\boldsymbol{p}|M)$ is the parameter prior and $P(D|M)$

[1] Their names are *burning.dat* and *bayes.dat*, respectively, for the burn-in and the sampling phases.

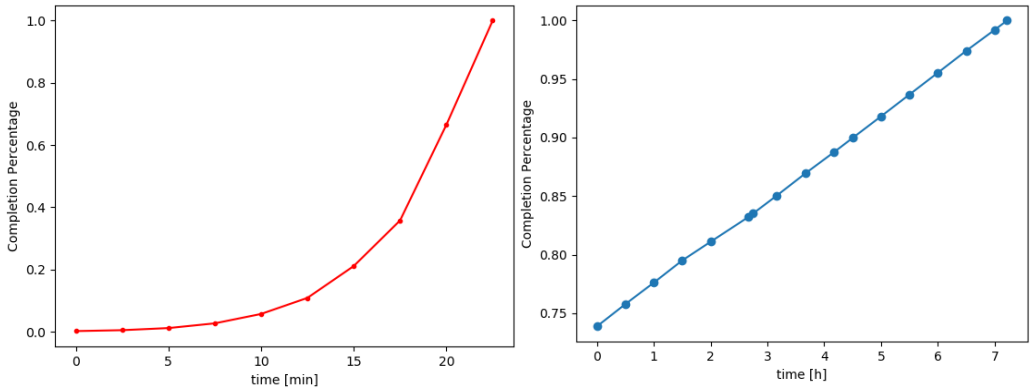


Figure 7.1: Variation of the number of iterations over the total one, as a function of the run time for the burn-in (left panel) and the sampling (right panel) phases.

is the evidence, namely the probability of getting the observed data given the assumed model. Our priors are uniform distributions, while the likelihood is related to the function of Eq. (6.8) by

$$P(D|\boldsymbol{p}, M) \propto e^{-\frac{\chi^2}{2}} \tag{7.2}$$

and the evidence can be explicitly written in the following way

$$P(D|M) = \int P(D|p', M)P(p'|M)dp' \,. \tag{7.3}$$

The parameter λ is that related to the run duration since it controls the convergence speed. The parameter space is explored, at each step, with 10 sampling points, which are selected with a variant of the Metropolis-Hasting algorithm (Metropolis et al. 1953; Hastings 1970). Each of them belongs to one of 10 interlinked Markov chains, adopted to lessen the chance to be trapped in a local minimum, in an algorithm which keeps them uncorrelated (for more details see Skilling 2004). During the different steps, the samples with the worst/best likelihood are deleted/duplicated, so that the number of running Markov chains is always 10. At the same time, the value of λ raises from 0 to 1 by a unit of

$$\delta\lambda = \frac{Rate}{\log(L_{max}) - \log(\langle L \rangle)} \,, \tag{7.4}$$

where $\langle L \rangle$ and L_{max} are the values of the mean and maximum likelihood of the 10 samples and $Rate$ is an arbitrary constant, set in the model parameter file (Sect. 6.2). When $\lambda = 0$, the denominator in Eq. (7.1) is one and the posterior distribution corresponds to the prior; on the other hand, when $\lambda = 1$, the Bayes' theorem is recovered and the posterior distribution is sampled as usually done with MCMC techniques. So, varying λ, we move progressively from the prior to the posterior distribution and the more this is gradual, the more chance of a good convergence one has.

As a consequence of the description above, the run time is a quantity strongly user-dependent. In fact, the time in which the convergence is reached is sensitive to the variation of the λ step in Eq. (7.4); $\delta\lambda$, in turn, depends on the value of $Rate$, which, thus, influences the run duration (together with the number of sampling iterations, a further

user parameter). Notice that $Rate$ also determines the precision of the optimization, namely with which resolution the parameter space is explored: higher values means fewer iterations, hence a faster run and less precision or a coarse parameter space exploration (Jullo et al. 2007). We find that, for models with member and multiple image catalogs having, respectively, about 30 and 10 sources, the burn-in and the sampling phases have approximately a duration of 30 minutes and few hours (see Fig. 7.1). We also identify a linear relation in the second phase between the percentage of performed iterations and the run time.

The value we set for the parameter $Rate$ derives from the results of Jullo et al. (2007). There, the same MCMC sampler is used to recover the model parameters of three simulated galaxy clusters, having a single cluster-scale halo and a set of galaxy-scale sub-halos. There, the $Rate$ value suggested is between 0.10 and 0.50, based on considerations about the computational time and the data reported in Fig. 7.2: here, in fact, it is shown that the uncertainty on the logarithm of the evidence is smaller for low $Rate$ values, which thus must be favored (limited to the run time). We decide to use a slow convergence speed, around the lower value of the suggested rate interval, setting $Rate = 0.05$ or 0.10. From our experience, we have found that this is a good compromise to obtain results in a reasonable amount of computation time and with accurate evidence values. Moreover, further works also suggest that a value of 0.05 leads to a good resolution (Limousin et al. 2008).

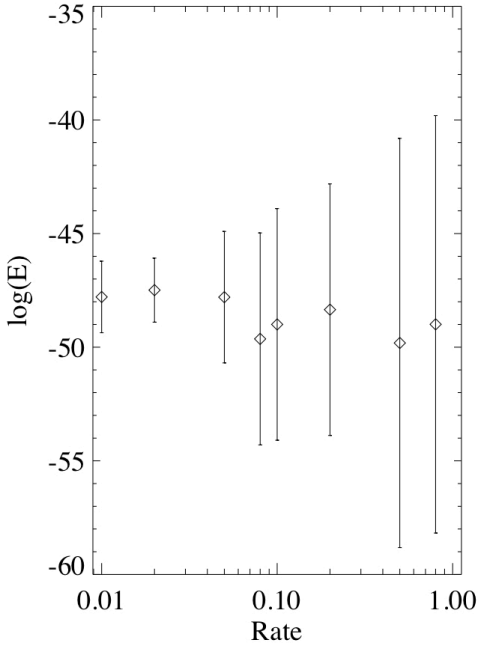


Figure 7.2: Evolution of the Evidence in function of the parameter $Rate$, regulating the convergence speed (from Jullo et al. 2007).

As previously explained (Sect. 6.3.1), to model the cluster total mass components, we fix to 6 the number of free parameters of the diffuse halo and to 2 that of each sub-set of small-scale halos. To them, we have to add the two coordinates defining the position of each multiply imaged source. On the other hand, in each model, we use the positions of a specific set of multiple images as constraint, so that the number of model *dof*, reported in Table 7.1, varies according to formula 6.5. The only exception is RUN 3, where the redshift of one multiple image family is added as a further free parameter. Note also that in RUN 5 a second small-scale halo sub-set is considered, including just an additional sub-halo.

With the method explained in Sect. 6.3, we run a collection of computational experiments, to attain a fiducial model by reducing the offset between the predicted and the observed multiple images (see Table 7.1 for a summary of our runs). We start from the model presented in Sect. 5.3 (RUN 1), in which we consider a minimal number of multiple-image systems and a reduced version of our final catalog of cluster members (only including the 29 cluster galaxies spectroscopically confirmed). We then refine the model choices of each successive run considering the results of the previous one. For ex-

ample, we include further images, properly selected on the basis of the findings of each preceding run. We now describe the analysis process from the first model to the fiducial one (corresponding to RUN 5), in which the adopted multiple images are those given in Table 5.2.

Table 7.1: Summary of run characteristics and results. For each run, we report, respectively, the ID, the list of images considered, the four quantities to calculate the model *dof* of Eq. (6.5) and two relevant statistical quantities: the initial value of the minimum χ^2 (the final one equals the number of *dof*, column 8) and the rms error, as defined in Sect. 6.3.2 (column 9).

RUN ID	images	n_{im}	n_S	n_H	$n_{set}(n_h)$	*dof*	χ^2_{in}	$\delta_{\rm rms}$
1	$1a,b,c$							
	$3a,b,c$							
	$4a,b,c,d$	10	3	1	1(29)	6	52.11	$0''.25$
2	$1a,b,c,d$							
	$3a,b,c,d$							
	$4a,b,c,d$	12	3	1	1(29)	10	72.34	$0''.26$
3	$1a,b,c,d$							
	$2a,b,c,d$							
	$3a,b,c,d$							
	$4a,b,c,d$	16	4	1	1(29)	15	76.38	$0''.24$
4	$1a,b,c,d$							
	$2a,b,c,d$							
	$3a,b,c,d$							
	$4a,b,c,d$	16	4	1	1(111)	16	76.80	$0''.24$
5	$1a,b,c,d$							
	$2a,b,c,d$							
	$3a,b,c,d$							
	$4a,b,c,d$	16	4	1	2(110+1)	14	35.06	$0''.15$

For the model of RUN 1, a final rms offset between the observed and the model-predicted positions of the multiple images of $\delta_{\rm rms} = 0''.25$ is found[2]. This model predicts further counter images for each of the families F1 and F3. In particular, as shown in Fig. 7.3 two new images are demagnified and very close to the BCG center and the remaining two are images 1d and 3d. The former corresponds to the less reliable[3] F1 image of C18. The image 3d is angularly very close to that which we have visually identified as a possible candidate and initially discarded because of its redshift quality flag ($QF = 1$).

The procedure to extrapolate from `Lenstool` the model-predicted multiple images is not a direct process since it requires two additional runs and the creation of a new input file. In fact, at the end of each run, the output file containing the information on the best model (*best.dat*) has to be properly modified and run to generate the list of all the sources (*source.dat*) corresponding to the multiple images of the input catalog, when the latter are projected onto the source plane[4]. Usually, the source-plane projected images of each family do not mark a single position for the same background source. For this

[2]The rms error on the image positions can be found, together with the value of the minimum χ^2, in the output file *chires.dat*.

[3]In that work, it is indicated as 1.3 and has a photometric redshift value ($z_{ph} = 0.34$) very different from those of the other images included in the same family (see column 4 of Table 1.3 in C18).

[4]This *best.dat* file run requires the activation of the key-word *image*, under the first identifier *runmode*.

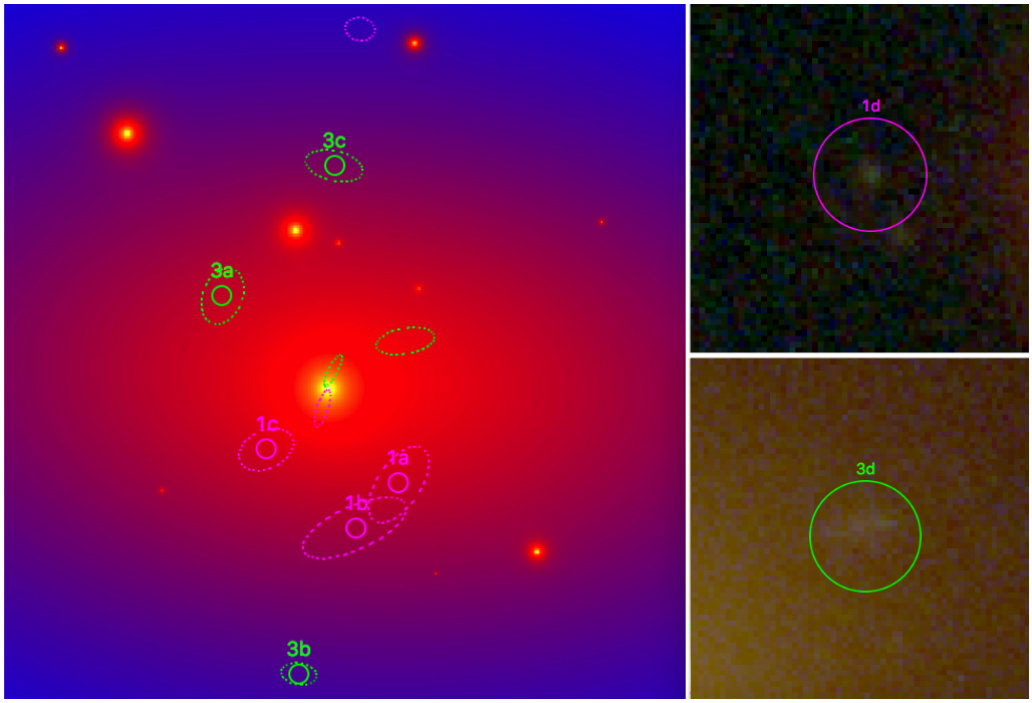


Figure 7.3: Multiple images predicted by RUN 1. In both panels, solid and dashed curves mark, respectively, the positions of the observed and model-predicted images of family F1 (in magenta) and F3 (in green). The left panel figure is a zoom on the total projected mass map of A2163-A1, derived from RUN 1 (with denser regions in bright colors). The two model-predicted images which are very close to the center of the BCG are not visible in the *HST* data, while the remaining ones (1d and 3d) have visible counterparts. The latter are reported in the two figures of the right panel.

reason, we need to average all the source positions corresponding to a single family, in order to obtain the centroid of the single source associated with that family. We collect all this new information in an input file for the software, which can now be used to possibly identify new model-predicted images in each family, through a second run[5]. Finally, we must validate these new multiple image candidates (listed in the file *image.all*) through theoretical considerations, for example checking their magnitude to understand if their (de-)magnification has an acceptable physical meaning.

Due to the previous run results, we extend the catalog of multiple images including the predicted images 1d and 3d. Thus, the RUN 2 catalog consists of three systems of four images each. RUN 2 leads to a comparable final rms error of $\delta_{\rm rms} = 0''.26$, compared to that of RUN 1, but the model has now more *dof*.

Only at this stage, we consider the four images of F2, all having a quality flag $QF < 2$: three come from the literature (C18) and the remaining one (2d) is identified by us in the *HST* data. The addition of F2 to the previous image catalog is justified by considering that the corresponding background source seems connected to that of F1, as illustrated

[5]This *best.dat* file run requires the activation of the key-word *source*, under the first identifier *runmode*.

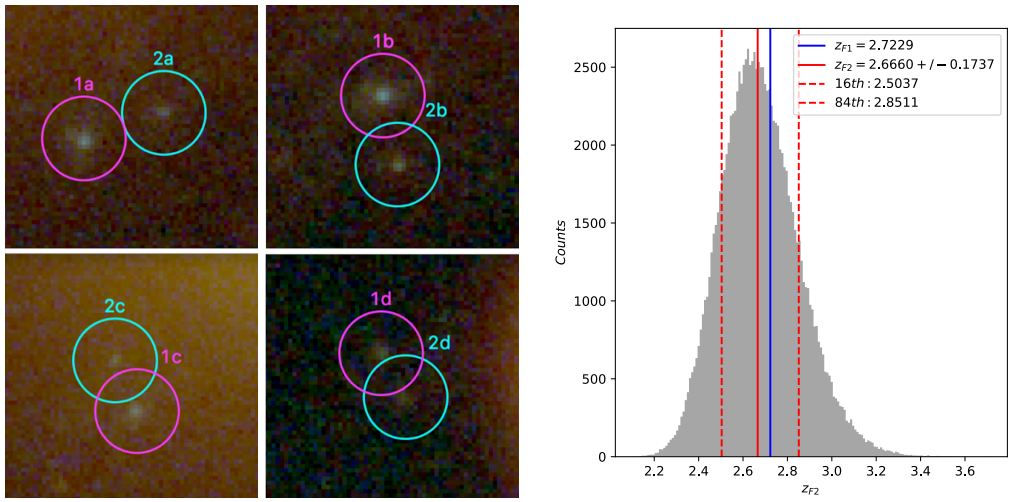


Figure 7.4: Left panel: collage of *HST* cutouts illustrating the four images of F1 and F2. Right panel: distribution of the model-predicted F2 redshift values, $z_{\rm F2}$, in RUN 3. The solid red line corresponds to the median value of $z_{\rm F2}$, while the blue one marks the measured value of F1 redshift, $z_{\rm F1}$.

in the left panel of Fig. 7.4. We find that the final rms error reduces when the F2 source redshift, $z_{\rm F2}$, is free to vary (RUN 3). However, a similar rms value is recovered when we consider an additional run, identical to RUN 3, except for $z_{\rm F2}$, which we fix to the redshift value of F1 for all the remaining runs. Moreover, the model-predicted redshift, $z_{\rm F2} = 2.666 \pm 0.174$, has a value consistent with that measured for family F1, with a difference of $\Delta z \sim 0.057$ (right panel of Fig. 7.4).

For the last two runs (RUN 4 and RUN 5), we consider an extended cluster member catalog, including both spectroscopic and photometric members (as detailed in Sect. 5.2). Notice that for these runs, the computational time increases, since the length of the member list has a considerable effect on `Lenstool` run-time, particularly if, like in our case, the optimization is in the image plane (Sec. 6.2). Compared to RUN 3, in RUN 4 $\delta_{\rm rms}$ does not improve significantly and the number of free parameters is smaller. The BCG velocity dispersion value is $\sigma_{0,\rm g} \simeq 400$ km s^{-1}. We seek to explain this quite high value, by analyzing how the scaling relations (6.6)-(6.7) work in practice. We find that the values of the parameters of the scaling relations are driven by the galaxy around which the family F4 is observed, consequently the high velocity dispersion value of the BCG derives from the assumption that the same relations hold for all the cluster members. A comparison between our value of the BCG velocity dispersion with that derived by other strong lensing studies is not possible here, due to the the lack of such information in the literature. Based on these considerations, in RUN 5, we free from the scaling relations the galaxy contributing in the formation of F4 arcs and model it with a spherical isothermal mass profile, with two additional free parameters, the central velocity dispersion, $\sigma_{0,\rm g2}$, and the truncation radius, $R_{\rm T,g2}$.

We refer to RUN 5 (the *fiducial model*) to illustrate our findings since it provides the best results, in the sense explained in what follows. The MCMC method of `Lenstool` follows a Bayesian approach (Sect. 7.1): the sampler does not aim at finding the model best-fit parameters (although the χ^2 is calculated), but the fiducial ones, namely we use the information on the distribution of the parameter values, where the posterior PDF is higher. For this reason, we present our *best-fit* parameters in terms of a median value (and the 16% and 84% quantiles around it) from the marginalized posterior distributions. The latter are shown in the 2D histograms along the diagonal of Figs. 7.5 and 7.6.

Thus, to conclude that RUN 5 corresponds to the fiducial model, we compare all the models (RUN 1-RUN 5) considering not a frequentist goodness-of-fit test, but an excellent tool for model selection, i.e., the Bayesian evidence[6]. This quantity is used to rank models as a quantitative measure of their complexity. Models with more free parameters can fit the data better, but they have also an extra complexity, which should be avoided whenever a simpler model properly describes the observations[7]. The relative probabilities of competing theoretical models are then evaluated and not only in light of the data, but also based on any prior information available. Finally, the better balance between quality of fit and predictivity defines the fiducial model. The evidence, thus, favors models with the smallest number of parameters and with a prior PDF as close to the posterior PDF as possible. As a comparison, the reduced χ^2 is only a rough approximation to the evidence analysis and provides an absolute goodness-of-fit estimator (when data error estimates are accurate).

7.2 Comparing the fiducial model with the literature

Our fiducial model predicts very accurately the positions of the multiple-image systems, with an rms error, defined in Eq. (6.9), of $\delta_{\rm rms} \simeq 0''.15$., i.e., approximately 2.5 *HST* pixels. The (median) values of the parameters and their errors can be found in Table 7.2 and Figs. 7.5 and 7.6.

Table 7.2: Median values and confidence level (CL) uncertainties of the parameters for the lens model RUN 5 of A2163. Centers are relative to the BCG (R.A. = 243.9539405° and Decl. = -6.1448406°).

Parameter	Median	68% CL	95% CL	99.7% CL
$x_{\rm h}$ [$''$]	1.98	$^{+0.34}_{-0.30}$	$^{+0.74}_{-0.56}$	$^{+1.19}_{-0.81}$
$y_{\rm h}$ [$''$]	0.18	$^{+0.15}_{-0.16}$	$^{+0.29}_{-0.33}$	$^{+0.42}_{-0.50}$
$\epsilon_{\rm h}$	0.61	$^{+0.04}_{-0.04}$	$^{+0.07}_{-0.07}$	$^{+0.11}_{-0.11}$
$\theta_{\rm h}$ [$^\circ$]	178.5	$^{+0.7}_{-0.6}$	$^{+1.5}_{-1.1}$	$^{+2.4}_{-1.6}$
$R_{\rm C,h}$ [kpc]	36.0	$^{+2.8}_{-2.5}$	$^{+5.8}_{-4.8}$	$^{+8.9}_{-6.8}$
$\sigma_{0,\rm h}$ [km s^{-1}]	959	$^{+26}_{-26}$	$^{+52}_{-50}$	$^{+77}_{-71}$
$\sigma_{0,\rm g2}$ [km s^{-1}]	316	$^{+36}_{-43}$	$^{+63}_{-79}$	$^{+89}_{-97}$
$R_{\rm T,g2}$ [kpc]	2.28	$^{+0.98}_{-0.47}$	$^{+2.53}_{-0.61}$	$^{+4.93}_{-0.62}$
$\sigma_{0,\rm g}$ [km s^{-1}]	342	$^{+20}_{-20}$	$^{+39}_{-40}$	$^{+58}_{-62}$
$R_{\rm T,g}$ [kpc]	42.3	$^{+18.0}_{-12.0}$	$^{+50.5}_{-20.4}$	$^{+87.5}_{-26.1}$

[6]The logarithm of the evidence is provided by the `Lenstool` file *best.dat* at the end of every run

[7]In Jullo et al. (2007), in fact, the evidence is described as an *Occam's razor: "All things being equal, the simplest solution tends to be the best one."*

The halo ellipticity reported in Table 7.2 is defined like in the `Lenstool` potential we adopt: $\epsilon_h = (1-q^2)/(1+q^2)$, so that the minor-to-major axis ratio we find is $q \sim 0.49$. On the other hand, we correct the output velocity dispersion values from `Lenstool` for the constant factor suggested in the literature (e.g., Appendix C.1 in Bergamini et al. 2019).

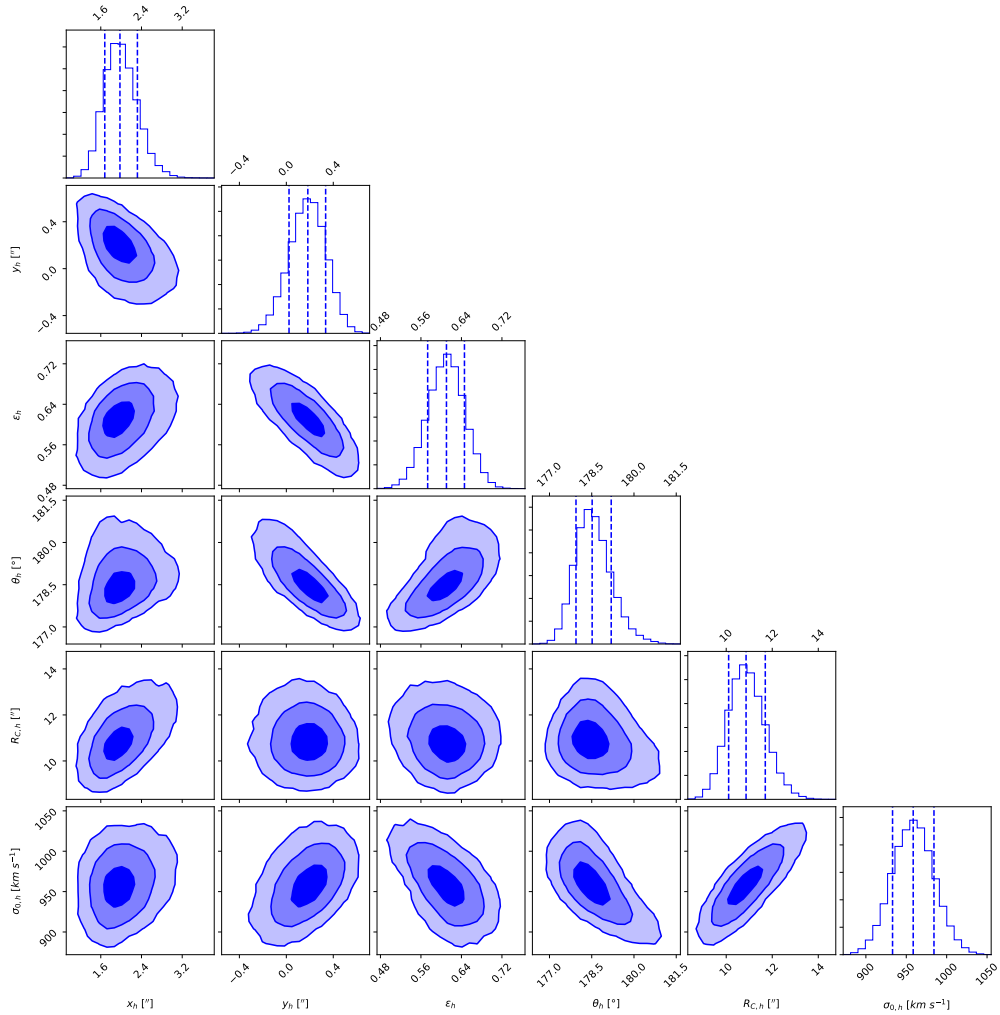


Figure 7.5: Posterior distributions of the parameter values of the cluster-scale mass component. Blue contours correspond to the 1, 2 and 3 σ confidence levels of a Gaussian distribution, while vertical blue dashed lines in the histograms are the 16th, 50th and 84th percentiles. Corner plots were created using the `corner.py` module by Foreman-Mackey (2016).

Our fiducial model shows some degeneracies (usually occuring in parameterized models). They are illustrated in Fig. 7.5, for the diffuse halo, and Fig. 7.6, for the BCG (left panel) and for the cluster member with *HST* ID 4104 (right panel). It is the galaxy around which the multiple images of F4 are observed, which in RUN 5 does not follow the scaling relations (6.6)-(6.7) (see Sect. 7.1). Galaxy 4104 shows a degeneracy between the truncation radius and the central velocity dispersion and has a very compact halo, with a maximum posterior density at the low R_T edge of our prior range.

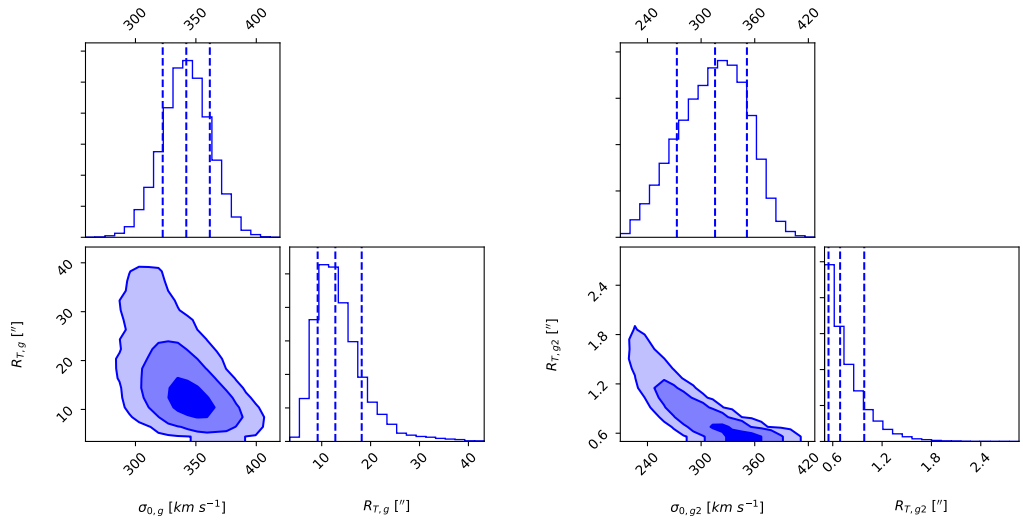


Figure 7.6: As in Fig. 7.5. Left panel: BCG parameters, which are linked to the small-scale sub-halo mass component, through the member galaxy scaling relations (6.6)-(6.7). Right panel: free parameters of the galaxy which contributes to the formation of F4 images.

We remark that recent results, obtained by combining strong lensing and velocity dispersion measurements of the members of different galaxy clusters, seem to favor very compact total mass profiles, e.g., Bergamini et al. (2019) (panel d of their Fig. 4). In that work, the sub-halos constrained by the use of cluster member kinematics have, as a result, a strongly reduced (inherent) degeneracy between the central velocity dispersion and the truncation radius. For the latter, an empirical formula is obtained, which can predict values as small as the one we find for the sub-halo of galaxy 4104 (considering the scatter in that relation). Small values of the truncation radius are also found for three over eight clusters in Caminha et al. (2019): $2.1^{+2.0}_{-1.2}$ kpc, 7^{+15}_{-6} kpc and 8^{+2}_{-2} kpc, respectively for the clusters RX J2129, MACS J1931 and MACS J2129. Moreover, some constraints on the small halo size of a satellite galaxy are obtained by modeling the surface brightness distribution of the extended multiple images of a strongly lensed source (see Suyu & Halkola 2010, who found a truncation radius of about 6.0 ± 2.5 kpc for a galaxy with a velocity dispersion value of 127 km s^{-1}). A plausible physical interpretation (or a concurrent cause) of the small sub-halo of galaxy 4104 is the presence of strong tidal fields. This seems reasonable because the closer the galaxy is to the cluster center, the smaller is the halo size, which is a behavior confirmed by Natarajan et al. (2009), who investigated the galaxy cluster Cl 0024+16, and by Limousin et al. (2009), in their hydrodynamical N-body simulations. In this respect, galaxy 4104 is very close to the BCG and at about the boundary of the region where the lens is critical.

We believe that only refined dynamical models of group/cluster members will clarify whether such small truncation radii are preferred for galaxies residing in overdense environments. Tests of this kind, which will be very interesting, go beyond the scope of the present thesis and can be performed in future works. There, a similar approach to that in Suyu & Halkola (2010) can be followed to constrain the three-dimensional tidal radius: the Bayesian inference allows one to overcome the restrictions of observations and

lens modeling, which only provide the projected two-dimensional distance between the primary lens and the sub-halo. The probability density of the three-dimensional tidal radius of the sub-halo given the projected distance can then be used to test the plausibility of the model-predicted size of the sub-halo. Moreover, understanding the dynamics of this particular galaxy during the merging processes between the A2163 central sub-clumps can give a crucial contribution in defining a conclusive explanation. Finally, a further interpretation of the small halo size originates from the SIDM model: collisional DM halo of elliptical galaxies in clusters can evaporate due to the high temperatures of the hot ICM (Gnedin & Ostriker 2001). In this scenario, the halo particles can act either as a fluid or as a rarefied gas, depending on their value of the mean free path $\lambda \simeq 1/[\rho\sigma/m]$, where ρ is the halo density and σ/m is the SIDM cross-section per particle mass. In the fluid regime, $\lambda \ll r$, the heat diffuses on the conduction time scale to the cool region (the sub-halo), while, for $\lambda > r$, the particle interaction is an elastic scattering.

Regarding the smooth halo, we find that it is flattened and elongated towards the A2163-A2 south-west direction, with a major axis length $a \sim 2b$. Moreover, we measure a small (projected) distance of $\sim 2''.0$ between the position of the diffuse component and the center of the BCG (see Fig. 7.5), with the latter towards east. We describe in Sect. 7.3 the relations between the characteristics of the cluster-scale halo and the possible merging history of A2163, as well as the consistency with the SIDM model predictions.

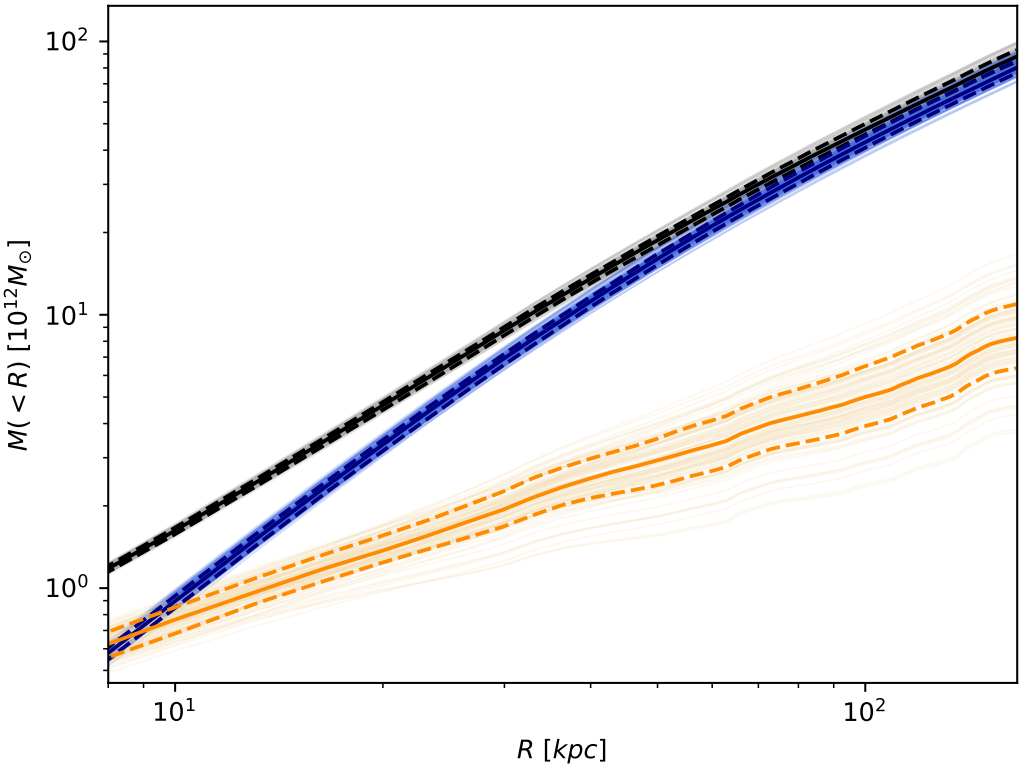


Figure 7.7: Cumulative projected mass profiles from our fiducial model; black, blue and orange curves represent the total, the smooth and the clumpy components, respectively. Solid and dashed lines trace the median and 16th - 84th percentiles, while the light ones complete the sub-sample extracted from the final MCMC chains.

In general, compared to the results by C18, we find parameter values with smaller errors, but they are consistent overall. We remark though that the inclusion in the models of the spectroscopic redshift values for the strongly lensed sources alleviates the parameter value degeneracies and thus significantly reduces the statistical error on the cumulative total mass profile of the cluster's core. In fact, our extrapolated value at 300 kpc, $M(< 300 \text{ kpc}) = 1.43^{+0.07}_{-0.06} \times 10^{14}\,\mathrm{M}_{\odot}$, is consistent, within the errors, with that found by C18, $M(< 300 \text{ kpc}) = (1.6 \pm 0.3) \times 10^{14}\,\mathrm{M}_{\odot}$, but its error is about an order of magnitude smaller then that in C18.

In our new strong lensing model of A2163-A1, the cumulative projected total mass profile is calculated from concentric annuli, centered at the BCG position. The profile is shown in Fig. 7.7, where we report three sets of curves with different colors to distinguish between the total mass profile (in black) and the mass contribution from two isolated cluster mass components: the diffuse halo (in blue) and the sub-halos associated with cluster members (in orange). Solid and dashed lines identify the median and 16th − 84th percentiles, respectively, while the light ones show a sub-sample in the final MCMC chains. The total mass profile confirms that in A2163-A1 the cluster halo is traced by its total light distribution.

In order to compare our total mass profile to those of different clusters found in the literature, we rescale the projected, total mass and the projected radius of A2163, respectively, with $M_{200} = 1.01 \times 10^{15}\,\mathrm{M}_{\odot}$ and $R_{200} = 2$ Mpc. These values are derived from independent measurements of weak lensing (Soucail 2012), which constrains the outskirts of galaxy clusters, from the external regions until about 200 kpc. We compare in Fig. 7.8 our projected, total mass profile with the results of Caminha et al. (2019). They perform a strong lensing analysis of eight CLASH galaxy clusters, using spectroscopic information from MUSE, which is also complemented with CLASH-VLT redshift measurements. Their sample has clusters with redshifts $z = 0.23 - 0.59$ and spans a range of weak lensing masses of $M = [5 - 35] \times 10^{14}\,\mathrm{M}_{\odot}$.

In Fig. 7.8, which illustrates this comparison, the legend refer to seven clusters. They are a combination of those in Caminha et al. (2019) having the best lensing models and three others from the literature, which are also precisely modeled, thanks to a relatively large number of images, in particular in the region $R/R_{200} < 10^{-1}$. Thr clusters in this sample have an unimodal or a bimodal mass distribution, suggesting very different dynamical states. For the sake of clarity, we decide to report in Fig. 7.8 only some points of the projected, total mass profile of A2163, which has been already presented in Fig. 7.7. We find that the shape of the projected total mass of A2163 is remarkably similar to the profile traced by all these clusters. This is particularly relevant considering the complexity of A2163 as described in Chap. 3. Finally, Caminha et al. 2019 find a value of $0.13 \times M_{200}$ for the mean projected total mass of the seven clusters listed in the legend of Fig. 7.8. It is calculated within 10% of R_{200} and has a small scatter of 5%. This value is compatible with that of the A2163 projected total mass, measured at the same rescaled radius, namely $[0.11 \pm 0.01] \times M_{200}$. This result for A2163, as well as the trend of the cluster sample in Caminha et al. 2019, is consistent with the predictions by Diemer & Kravtsov (2014). According to them, once the mass profiles of DM halos are expressed in units of spherical overdensity radii defined with respect to the critical (or mean) density of the Universe (especially R_{200}), they reveal a self-similar behaviour in their inner (or outer) structure.

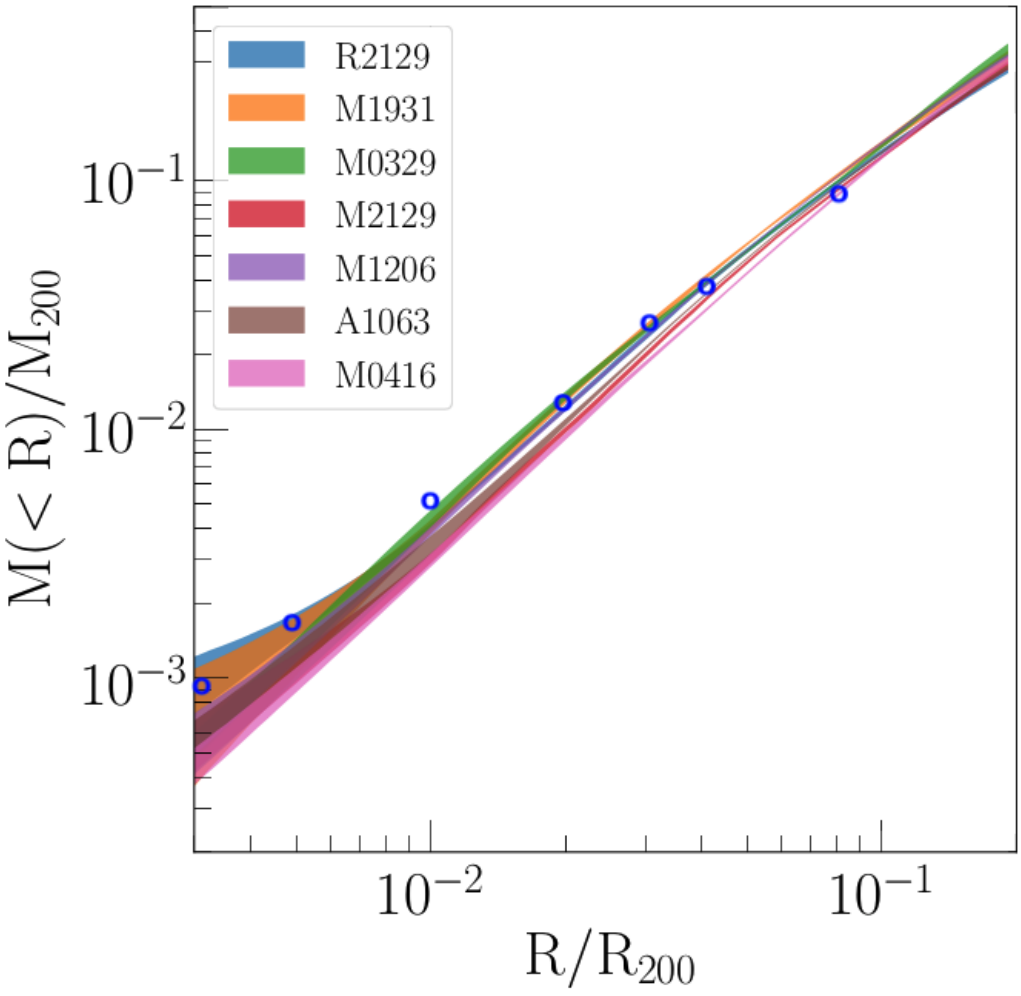


Figure 7.8: Projected total mass profiles of a sample of clusters described in Caminha et al. (2019) and references therein (colored curves); the blue circles mark some points of the projected total mass profiles of A2163 as derived by our fiducial model. All the curves are rescaled by the quantities M_{200} and R_{200} from independent weak lensing measurements.

From the total mass profile derived by our fiducial model, we find that the contribution to the total mass of the cluster-scale and the galaxy-scale components are $\sim 90\%$ and $\sim 10\%$, at $R > 100$ kpc, respectively. Nevertheless, we remark that including mass substructures in a strong lensing model is fundamental to approach a detailed cluster mass distribution for different reasons: first, to reproduce accurately the observed positions of the multiple images (Kneib et al. 1996; Meneghetti et al. 2007, 2017), then, to understand the effective lens efficiency in the presence of a large numbers of perturbers and, finally, to avoid the introduction of systematic effects (Jullo et al. 2007).

Another aspect to take into account to compare our results to those in the literature is the shape assigned to the sub-halos. In our models, we adopt simple circular total mass profiles to model the cluster members, namely the values of the ellipticity and position angle of the sub-halos are not fixed to those measured from the light distribution of the cluster galaxies, like in C18. This is because in the literature there is no conclusive evidence that the distribution of the stellar component of a galaxy can trace well its total mass distribution at all distances from its center, especially where the DM component dominates. Moreover, other strong lensing studies in galaxy clusters did not find any evidence supporting the inclusion of the ellipticity values of luminous cluster members: to represent their total mass distribution in this way does not provide a better fit to the data. For example, in Meneghetti et al. (2017), detailed comparisons between different modeling choices are made for the same (simulated) clusters and it is shown that the CATS team, which models cluster members with circular total mass profiles (see their Table 3), scores extremely well for all metrics considered (see their Fig. 27). In Treu et al. (2016), different models of a HFF cluster were compared. Among them, in the models by Grillo and Sharon, the galaxy-scale halos are, respectively, spherical (see their Sect. 4.2.1) and elliptical (with the same ellipticity as the light distribution, Sect. 4.4.1). Despite this difference, their global convergence maps are quite similar.

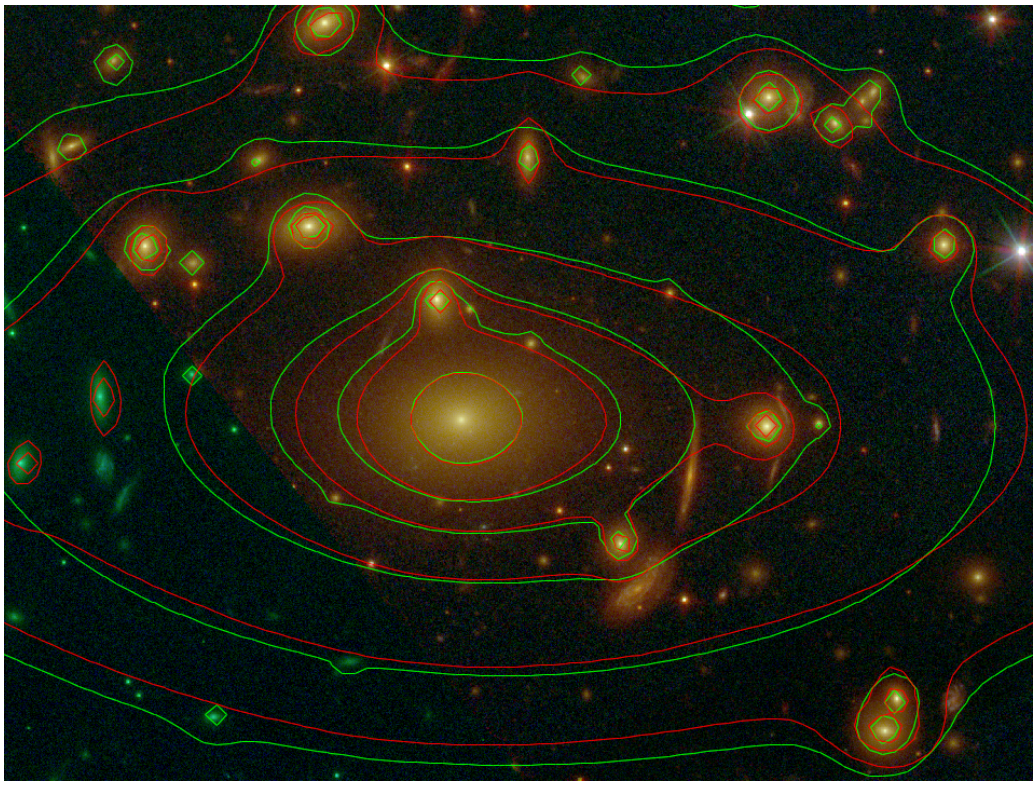


Figure 7.9: Composite image of A2163-A1, with overlays of the total surface mass density distribution. Green and red contours refer, respectively, to this work and C18. Contour line values are $[0.75\,,1.00\,,1.50\,,2.00\,,2.50\,,3.50] \times 10^9\,\mathrm{M}_\odot\,\mathrm{kpc}^{-2}$.

In addition, not only the ellipticity, but also the alignment between the position angle of a galaxy and its DM halo may be different and/or scale dependent. In particular, the position angle associated with the DM distribution can (highly) differs from that of the observable counterpart, e.g., van Uitert et al. (2012). This orientation misalignment is even found when the DM halo of some galaxies has a similar ellipticity compared to their stars (e.g., Jauzac et al. 2018). For all these reasons, in our current study we have preferred to model cluster member halos with simpler circular total mass profiles, instead of forcing the values of the ellipticity and the position angle to be the same as those of galaxy members. Our choice seems to produce similar results for galaxy members, compared to C18 strong lensing models of A2163-A1, where the sub-halos have ellipticities tied to the luminous distributions. This is shown in Fig. 7.9, where we report the contour lines of the surface mass density as derived in this work (in green) and in C18 (in red).

We find an overall agreement, with only some minor variations which can be explained in terms of the different cluster member selection and sub-halo total mass modeling. Unfortunately, a direct comparison with further total mass maps in the literature is not possible, since the needed information are not provided (e.g., in Soucail 2012, the mass map contour levels derived from weak lensing analyses have no precise quoted values and, in any case, the mass map resolution is too low). The only exception is indeed C18, whose convergence map is available on MAST.

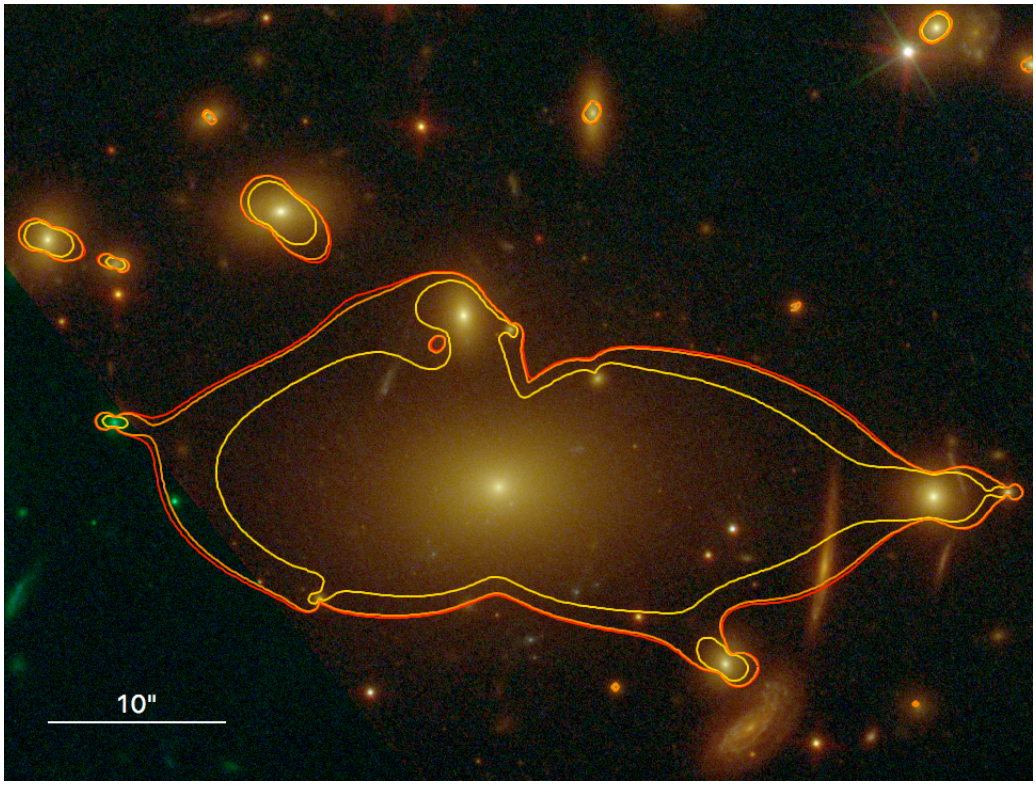


Figure 7.10: Critical lines of the fiducial model, for sources at redshifts 1.164 (gold), 2.389 (orange) and 2.723 (red), superposed on a color composite image of A2163-A1.

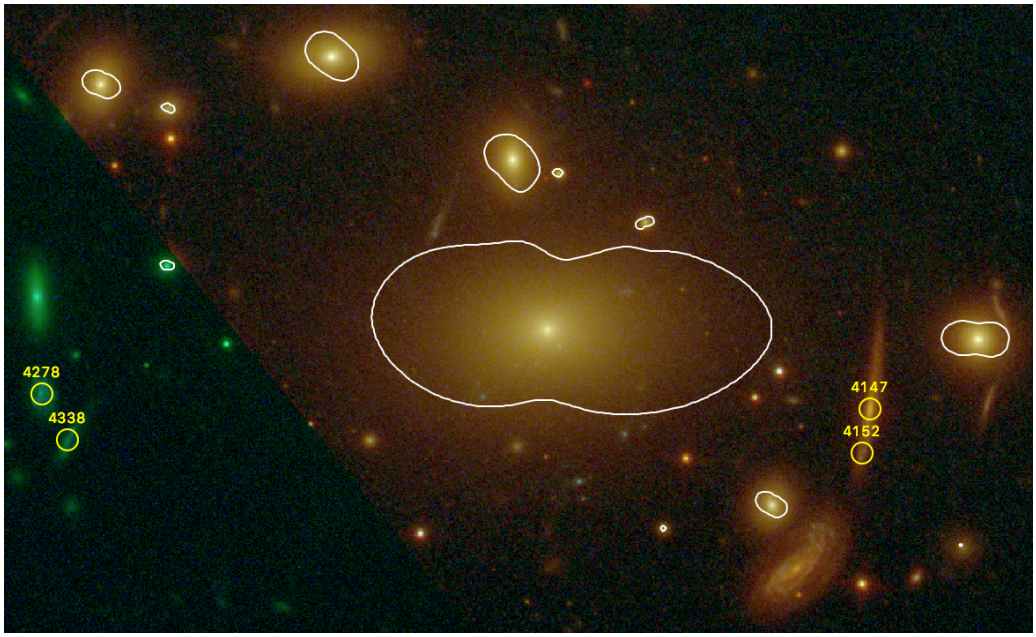


Figure 7.11: Critical lines (in white) for some galaxies (in yellow), having redshifts of about $z \sim 0.665$. This value is compatible with that of the arc-shape galaxy (4147)..

`Lenstool` also allows one to calculate the critical lines[8]. Beyond setting the redshifts, it is possible to define the area where the critical lines are searched for, to choose one of two algorithms and to set some other characteristics of the the latter. We show the critical lines in Fig. 7.10, in a region within the MUSE FoV and for the three redshift values of the four multiple image systems used in the fiducial model: $z = 2.723$, for the families F1 and F2, $z = 2.389$ for the family F3 and $z = 1.164$, for F4. We opt for the `Lenstool` marching squares algorithm which generates contour lines for a 2D scalar field (see Appendix A of Jullo et al. 2007). The scalar field region to consider is decided by the user, by setting the size of an imaginary grid. Specifically, the algorithm divides this user-defined grid in four small squares and, according to their size and their amplification values in the center and the four corners, further divides (or not) each square in four additional small squares[9]. As long as the square size is greater than a chosen limit, it is automatically divided until a lower size, also set by the user. The algorithm considers the amplification signs in the smallest squares to determine the parts of the critical line that pass through each square. Finally, these parts are fused into the the critical lines' contour.

The critical lines of Fig. 7.10 "pass through" a stretched galaxy, located east of the BCG. In Fig. 7.11, the luminous, central peak of this arc-shape galaxy is marked by the circle with *HST* ID 4147, while the white curves are the critical lines corresponding to its (low) redshift of $z_{4147} = 0.6646$. We exclude the possibility that this arc could be part of a multiple image family together with other sources with similar redshifts.

[8]The procedure is analogous to that to derive the model-predicted images, namely the *best.dat* file must be run, after adding a secondary identifier, *cline*, which implements the different functions described in the text. Finally, the output ASCII files *ci.dat* and *ce.dat* can be used to display the critical lines.

[9]In the case of a rectangular grid, the greater value between the width and the height of the field is chosen as the square size.

7.3 Conclusions and future directions

In this thesis, we have concentrated on the innermost region of the cluster A2163. We have presented our redshift measurements for all the member galaxies in the cluster core and within our MUSE FoV, compiling a pure spectroscopic catalog. Then, we have extended it to a sub-sample of member candidates selected on the basis of their photometric information. We have also reported the discovery of new multiple images and, exploiting our new MUSE data, have confirmed spectroscopically the majority of those already found in the literature. With such a solid dataset, we have built five new strong lensing models, using *HST* positions of multiple-image systems as constraints, have ranked them and then we have determined an accurate projected total mass profile in the core of A2163. Finally, we have decomposed the profile in two cluster- and galaxy-scale components, clarifying which parameters of the diffuse mass distribution are in favor of a possible cluster merging scenario.

The main results of this thesis can be summarized as follows.

1. **New spectroscopic catalogs**. We have measured the spectra of about 230 sources in the A2163 core and spectroscopically confirmed 8 multiple images, 35 foreground and background sources (Appendix A) and 29 cluster members (Table 5.1).

2. **Discovery of multiple images**. We have discovered 4 new multiple images and presented a final image catalog consisting of 16 multiple images from 4 different sources, with redshifts from 1.16 up to 2.72 (Table 5.2).

3. **Precise cluster mass model**. We have compared the predicted positions of the multiple images with those of their observational counterparts and found, for our fiducial model, an rms of $\delta_{\rm rms} = 0''.15$.

4. **Cluster projected total mass profile**. From our fiducial model, we have measured the cluster cumulative projected total mass profile very precisely and have found that $M(< 100 \text{ kpc}) = 4.75^{+0.23}_{-0.20} \times 10^{13}\,\mathrm{M}_{\odot}$, which is consistent with the value with significantly larger statistical errors found in the literature.

5. **Mass components and their distribution**. After testing different lensing models, we have concluded that the projected total mass distribution of A2163-A1 can be well represented by a diffuse component with a dPIE profile, a galaxy-scale spherical halo and a population of 110 sub-halos, with a total M/L ratio increasing with luminosity, as observed in early-type galaxies (tilt of the fundamental plane).

6. **Confirmation of DM halos self-similarity.** The cluster projected total mass profile, rescaled by M_{200} and R_{200}, is in agreement with those found for other clusters, even for very different cluster dynamical states, confirming the self-similarity of cluster-size halos predicted by cosmological simulations.

7. **Hints on the cluster merger history**. The position and the elongated shape of the diffuse halo of A2163-A1 in the direction of A2163-A2 support a post-merging scenario for these two sub-clumps in the cluster core (see below for a detailed description).

8. **SIDM consistency check**. The relative positions and the alignment of A2163 central mass components found in our fiducial model could be consistent with a SIDM scenario for merging sub-clusters (see below for a detailed explanation).

A number of scientific projects can be improved with results 1-6, such as the RELICS program, already introduced in Sect. 4.1.3, where we have detailed the data used, the sample selection strategy and its primary and ancillary goals. Other research projects can be devised as a natural extension of the results of points 7 and 8, e.g., the conclusive scenario for A2163 core merging processes and a possible test on the DM self-interaction. Below, we sketch these possibilities in light of the considerations and the results of the present thesis.

The discovery of high-redshift galaxies to constrain better their luminosity function is one of the goals of the RELICS program. Regarding this aim, we report the detection of few galaxies at a redshift of $z \sim 3-5$ (Appendix A). This is a minor contribution since the primary targets of this project are very high-redshift sources, in order to better understand the epoch of reionization and, thus, the early galaxy formation and evolution. Nevertheless, our spectroscopic catalog of A2163 multiple images and galaxy members is crucial to accurately model this galaxy cluster. This, in turn, furthers the RELICS purpose of determining precision magnification maps, which are necessary to interpret the measurements of background galaxies at high redshifts. For this reason, the RELICS group will remodel each lens when spectroscopic redshifts are available. Concerning A2163, the project preliminary results, presented in C18, include the first version of the strong lensing model of this cluster, but there only photometric redshifts are used as priors in the modeling process. As reported in Fig. 4.1, when no spectroscopically confirmed multiple image is introduced in a lens model, the systematic error of the magnification could reach up to 20%. Our more accurate cluster model, based on new spectroscopic measurements, results in an excellent matching between the predicted and observed positions of the multiple images and has model parameters with an error an order of magnitude lower than those in C18, so it will centrally contribute to enhance the magnification map precision. RELICS will benefit from the discovery of our new multiple images and the redshifts measured for them and for all the other ones. Finally, studying the spatial distribution of the mass components of RELICS clusters is also thought to be a starting point for DM particle constraints (see Coe et al. 2019 and Sect. 4.1.3). In this sense, if the configuration we find of the diffuse halo and of the luminous matter in A2163 will be confirmed, our analysis can also serve in this respect (see below).

Regarding the merging scenario of A2163 reported in the literature and summarized at the end of Sect. 3.2, it lacks a confirmation from strong lensing analyses. With the present thesis we bridge this gap for the processes occurring in the main, central cluster structure (A2163-A). From our fiducial model, we conclude that the distance between the center of A2163-A1 halo and the cluster BCG is practically negligible compared to the distance between their (nearly common) positions and the X-ray main peak (in Maurogordato et al. 2008). The latter is far from being coincident with the A2163-A1 center, which disfavors a pre-merging phase. Moreover, the shape of the diffuse halo of A2163-A1 is elongated in the direction of A2163-A2. Our explanation is that, after the merger, the halo of A2163-A1 has relaxed to an elliptical shape that retains information about the direction of the merger, thus pointing to the BCG of A2163-A2. For this reason, a precise measure of the ellipticity (and of the position angle) is crucial. In this sense, it is relevant that our values of these parameters have an error about an order of magnitude smaller than previously found (C18).

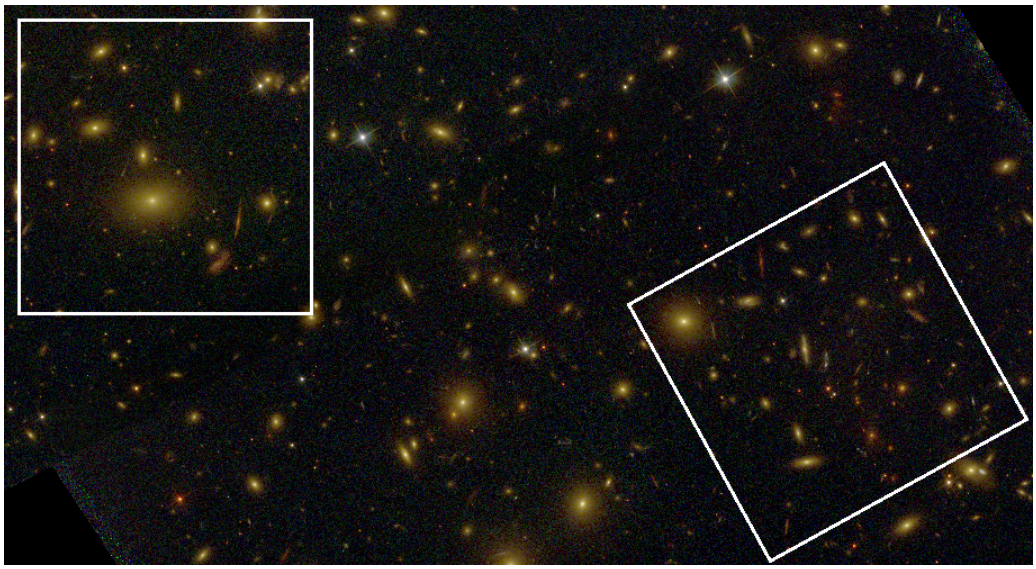


Figure 7.12: MUSE pointings in the core of A2163 (white boxes), in the north-est and south-west regions, respectively, on the left and right part of the figure; the length of each FoV side is $\sim 1'$.

Concerning the western sub-clump, A2163-A2, it has not been yet modeled using strong lensing techniques. This can be the objective of interesting future work, based on the available MUSE data of the A2163 south-west. The instrument's FoV covering this additional cluster region corresponds to the squared box (in white) reported on the right part of Fig. 7.12, together with the one used in this thesis (on the left). A strong lensing model of this system would be extremely promising: the constraints on the south-west cluster region coming from such a study and our results on the north-east clump would have the potential to allow an in-depth and conclusive analysis about the merging history of the cluster core. On larger scales, a confirmation of the overall merging scenario is not possible with strong lensing only. Besides, the presence of the *bridge* of faint galaxies along the north-south axis, which would connect the cluster core with the sub-clump A2163-B (Sect. 3.2), cannot be observed: this structure is supposed to be in a region well outside the MUSE FoV we focus on, whose size is one arc minute.

As discussed in Sect. 7.2, in our fiducial model, all sub-halos have positions we fix to those of the cluster galaxies. On the contrary, the coordinates of the large-scale halo center are free parameters and can vary within practically the entire region defined by the MUSE FoV, in other words this halo (mostly of DM) could deviate from the paradigm for which the mass distribution follows the light distribution. Its center is then found within $\sim 2''$ away from the BCG. Models which separate the mass contributions of the galaxies, the diffuse DM and the hot gas need to be explored, as they might mitigate the offset between these two mass components. In this direction, a relevant and new approach has been presented for the cluster MACS J0416.1-2403: in Bonamigo et al. (2017), the total mass and hot gas distributions have been separated and, in Annunziatella et al. (2017), the further subtraction of the stellar component and the consequent decoupling of the DM distribution has led to a complete mass decomposition. If these models will confirm a significant distance between the stellar and the DM distribution centers, we should not exclude the self-interaction of DM as a possible explanation for it[10].

[10] We have briefly introduced this DM category in Sect. 4.1.3, within the context of RELICS ancillary goals.

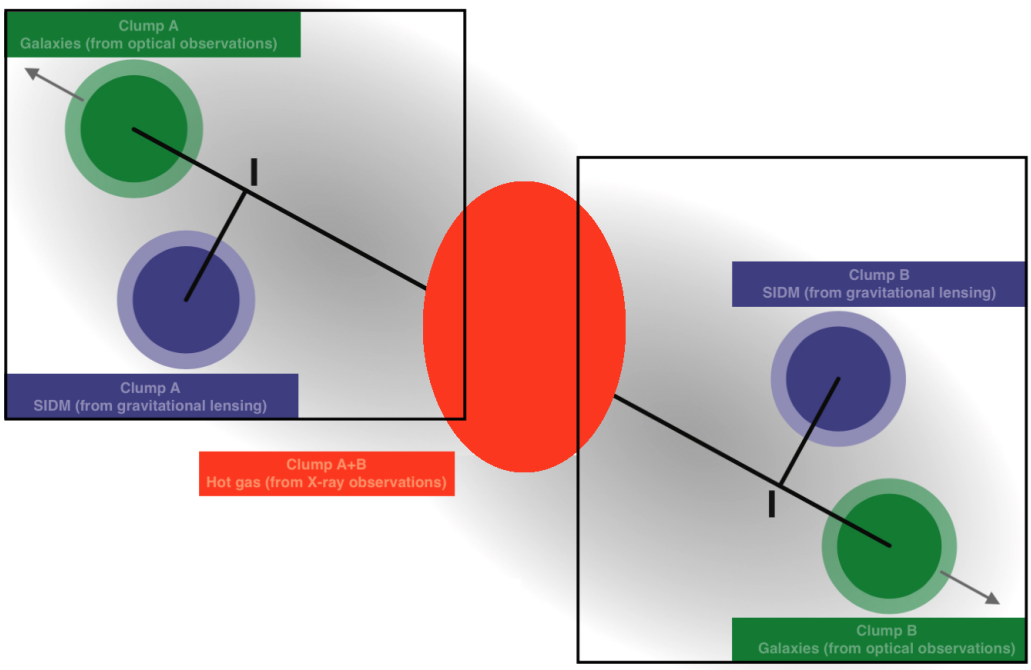


Figure 7.13: Spatial configuration of the three mass components of two merging sub-clumps (A and B) in a SIDM scenario: the stellar distribution (in green), usually centered at the BCG position, and those of the SIDM (in violet) and of the gas (in red) halos. The two arrows point to the motion direction of the sub-clumps after the collision.

SIDM, if real, would play an important role in solving some outstanding issues with standard cosmology, such as, for example, the *missing satellite problem* (Klypin et al. 1999). We remember that the latter deals with dwarf systems orbiting a common-size galaxy and consists in that cosmological simulations exhibit more satellites than the observed counts. The self-interacting and the (standard) collisionless nature of DM particles can be discriminated by focusing on massive clusters where merging processes have already occurred. Under the standard cosmological model assumptions, when two cluster sub-clumps collide, the galaxies of each of them "pass through each other", like the DM component, because they have a collisionless behavior. On the other hand, the gas of each clump does collide, forming shock waves and decelerating. On the contrary, if the DM interacts with itself, it will also slightly separate from the stellar distribution. For this reason, the crucial quantity to consider is the offset between the center of the cluster-scale DM halo and that of the stellar matter distribution, which must be non-negligible in order to confirm this DM model. However, the relative positions between the three mass components must also be consistent with the spatial configuration representative of such a DM scenario, as illustrated in Fig. 7.13. There, their predicted arrangement after the collision is indicated: an X-ray peak, from the gas (in red), should be located in the barycenter of the two merging systems, due to the gas collisional behavior; then, from here towards two opposite directions, there should be the centers of the two SIDM halos associated to each sub-clump (in violet) and, to follow, the luminous stellar component (in green).

As anticipated, we find an offset, in projection, of $\sim 2''$ (or ~ 7 kpc) between the BCG center and that of the diffuse halo. The relative positions and the alignment of the mass components of A2163 predicted by our fiducial model are also consistent with the predictions of a SIDM scenario for merging sub-clusters: the main X-ray peak, identified by Maurogordato et al. (2008), is between the east and the west sub-structures in the cluster core; then, toward the east, the diffuse mass component is found between the central gas and the BCG of A2163-A1. Moreover, the mass content of the cluster-scale halo is dominated by DM, thus the total mass derived by gravitational lensing is sometimes used as an approximation for the DM distribution, especially for merging systems in which the gas is substantially removed from the DM potential wells of the two colliding clumps (e.g., Molnar 2016 and references therein). Nevertheless, as of today, no merging scenario can explain all observed features of the A2163 core (described in Chap. 3), thus, in a conservative approach, our findings must be considered only as a preliminary result in characterizing the possible self-interaction of DM.

Clearly, the case of A2163 urges further investigations, such as detailed numerical simulations, to disentangle its complex phenomenology and find out whether it can be explained by the ΛCDM predictions or a non-standard DM paradigm is instead required.

APPENDIX A

Catalog of foreground and background sources

Catalog of foreground and background sources, with *HST* IDs, celestial coordinates and spectroscopic redshifts with QFs. We exclude sources with $QF < 1$ because they are too faint and/or noisy and do not show clear spectroscopic features. ID 9000 is assigned to a source which we identify and that is not present in the *HST* catalog.

ID	R.A. [Deg]	Decl. [Deg]	$z_{\rm sp}$	QF
3192	243.94458	−6.13500	0.0	3
3710	243.94535	−6.14001	0.0	3
3356	243.94752	−6.13804	0.0	3
4430	243.94896	−6.14884	0.0	3
3818	243.95029	−6.14547	0.0	3
3819	243.95066	−6.14589	0.0	3
3817	243.95176	−6.14685	0.0	3
3594	243.95516	−6.13921	0.0	3
3391	243.95563	−6.13696	0.0	3
4545	243.95596	−6.14804	0.0	3
3431	243.95760	−6.13750	0.0	3
4219	243.95904	−6.14505	0.0	3
3902	243.95992	−6.14182	0.0	3
3802	243.96003	−6.13957	0.0	3
3886	243.96081	−6.14225	0.0	3
4791	243.96117	−6.15096	0.0	3
5256	243.96141	−6.15066	0.0	3
3240	243.95248	−6.13561	0.0	2
3726	243.95371	−6.14011	0.3344	3
3371	243.95047	−6.13713	0.3728	1
3868	243.95238	−6.14197	0.4075	3
3242	243.95538	−6.13565	0.4399	3
3576	243.94649	−6.13783	0.5940	3
4338	243.96156	−6.14658	0.6640	3
4147	243.94883	−6.14609	0.6646	3
4152	243.94896	−6.14678	0.6657	3
4278	243.96196	−6.14585	0.6658	3
4463	243.95000	−6.14856	0.7284	3
9000	243.95826	−6.14134	0.8249	3
4492	243.96150	−6.14761	0.8673	1
4644	243.94810	−6.14873	0.9787	3
3284	243.96128	−6.13606	1.0406	1
3568	243.94898	−6.13870	1.1664	3
3580	243.95434	−6.13708	1.3725	3
3285	243.95490	−6.13686	1.3740	3
3222	243.95373	−6.13571	1.3775	3
4713	243.95345	−6.14944	3.1660	1
3854	243.94453	−6.14187	4.5837	2
3693	243.94649	−6.13971	4.9901	2

APPENDIX B

Notes on Lenstool

First written by Jean-Paul Kneib in his Ph.D. thesis (1993) and now (September 2019) released in its 7.1.1 version, the program Lenstool is publicly available on the Lenstool project homepage[1], together with the installation instructions, the documentation files and a list of papers using the program.

Lenstool source code is compatible with Linux and OS environments, with dependencies on WCSTOOLS, CFITSIO (at least version 2.510), GSL and PGPLOT (optional). Problems in the installation process can be overcome with standard libraries and the most common usage difficulties can be faced by a (not-yet complete) documentation. To fully exploit all the functions implemented in such a complex and comprehensive gravitational lens modeling software, to know how to use the distributed tools and to solve more complex issues, one possibility is to take advantage of online information: the forum discussions (like those in the Lenstool Google group) or the wiki section on the software website, which is more updated than the latest version of the user manual (hence to prefer to the latter). However, the authors also give a set of example files to gain familiarity with the Lenstool options and commands.

This software has a terminal user interface (or Character-based User Interface, ChUI) and can be used listing text commands within few input files. The file text structure is organized in two kinds of keywords, *primary* and *secondary identifiers*: the former are 15 keyword groups, under which the latter are stated, on each file line, together with a certain number of numerical or string values (the parameters of the secondary identifiers). In other words, a specific set of second identifiers corresponds to each first identifier and is always followed by its parameter values. Not all the first identifiers nor their secondary identifiers are mandatory, hence the complexity of the file describing the lens model depends on the specific case under analysis.

[1] https://projets.lam.fr/projects/lenstool/wiki#

Bibliography

A. B. Congdon, C. R. K. 2018, Principles of Gravitational Lensing. Light Deflection as a Probe of Astrophysics and Cosmology (Springer)

Abell, G. 1958, Astrophysical Journal Supplement, 3, 211

Abell, G. O. 1975, Stars and Stellar Systems IX: Galaxies and the Universe (Chicago: University of Chicago), 601

Abell, G. O., Corwin, H. G. J., & Olowin, R. P. 1989, ApJS, 70, 1

Adam, R., Aghanim, N., Ashdown, M., et al. 2016, *Astron. Astrophys.*, 596, A108

Ade, P. A. R., Aghanim, N., Arnaud, M., et al. 2016a, *Astron. Astrophys.*, 594, A13

Ade, P. A. R., Aghanim, N., Arnaud, M., et al. 2016b, *Astron. Astrophys.*, 594, A27

Akamatsu, H., Mizuno, M., Ota, N., et al. 2018, *Astron. Astrophys.*, 618, A74

Allen, S. W., Schmidt, R. W., & Fabian, A. C. 2001, *Mon. Not. R. Astron. Soc.*, 328, L37

Allgood, B., Flores, R. A., Primack, J. R., et al. 2006, *Mon. Not. R. Astron. Soc.*, 367, 1781

Annunziatella, M., Bonamigo, M., Grillo, C., et al. 2017, *Astrophys. J.*, 851, 81

Arnaud, M., Hughes, J. P., Forman, W., et al. 1992, *Astrophys. J.*, 390, 345

Bacon, R., Accardo, M., Adjali, L., et al. 2012, The Messenger, 147, 4

Bacon, R., Brinchmann, J., Richard, J., et al. 2015, *Astron. Astrophys.*, 575, A75

Bahcall, N. A. 1977a, Annual Review of Astronomy and Astrophysics, 15, 505

Bahcall, N. A. 1977b, *Astrophys. J.*, 218, L93

Bahcall, N. A. 1977c, *Astrophys. J.*, 217, L77

Balestra, I., Mercurio, A., Sartoris, B., et al. 2016, ApJS, 224, 33

Balogh, M. L., Pearce, F. R., Bower, R. G., & Kay, S. T. 2001, *Mon. Not. R. Astron. Soc.*, 326, 1228

Bartelmann, M. & Meneghetti, M. 2004, *Mon. Not. R. Astron. Soc.*, 418, 413

Bautz, L. P. & Morgan, W. W. 1970, The Astrophysical Journal, 162, L149

Bender, R., Burstein, D., & Faber, S. M. 1992, *Astrophys. J.*, 399, 462

Bergamini, P., Rosati, P., Mercurio, A., et al. 2019, *Astron. Astrophys.*, 6, 461

Bertin, E. & Arnouts, S. 1996, Astron. Astrophys. Suppl. Ser., 117, 393

Blandford, R. D. & Kochanek, C. 1987, Astrophys. J., 321, 658

Bonamigo, M., Grillo, C., Ettori, S., et al. 2017, *Astrophys. J.*, 842, 132

Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, Apj, 379, 440

Botteon, A., Gastaldello, F., & Brunetti, G. 2018, *Mon. Not. R. Astron. Soc.*, 476, 5591

Bourdin, H., Arnaud, M., Mazzotta, P., et al. 2011, *Astron. Astrophys.*, 527, A21

Bouwens, R. J., Bradley, L., Zitrin, A., et al. 2014, *Astrophys. J.*, 795, 126

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, *Astrophys. J.*, 803, 34
Bovy, J., Hogg, D. W., & Roweis, S. T. 2011, The Annals of Applied Statistics, 5, No. 2B, 1657
Bradač, M., Erben, T., Schneider, P., et al. 2005a, *Astron. Astrophys.*, 437, 49
Bradač, M., Schneider, P., Lombardi, M., & Erben, T. 2005b, *Astron. Astrophys.*, 437, 39
Bradley, L. D. & CLASH Team. 2014, American Astronomical Society, AAS Meeting 223, id.139.03
Burenin, R. A., Vikhlinin, A., Hornstrup, A., et al. 2007, The Astrophysical Journal Supplement Series, 172, 561
Butcher, H. & Oemler, A. J. 1978, *Astrophys. J.*, 219, 18
Caminha, G. B., Grillo, C., Rosati, P., et al. 2016, *Astron. Astrophys.*, 587, A80
Caminha, G. B., Grillo, C., Rosati, P., et al. 2017a, *Astron. Astrophys.*, 600, 20
Caminha, G. B., Grillo, C., Rosati, P., et al. 2017b, *Astron. Astrophys.*, 607, 21
Caminha, G. B., Rosati, G., Grillo, C., et al. 2019, *Astron. Astrophys.*, 632, A36
Capak, P., Abraham, R. G., Ellis, R. S., et al. 2007, The Astrophysical Journal Supplement Series, 172, 284
Cavaliere, A. & Fusco-Femiano, R. 1978, *Astron. Astrophys.*, 70, 677
Cavaliere, A. G., Gursky, H., & Tucker, W. H. 1971, Nature, 231, 437
Cerny, C., Sharon, K., Andrade-Santos, F., et al. 2018, *Astrophys. J.*, 859, 159
Chen, W., Kelly, P. L., Diego, J. M., et al. 2019, *Astrophys. J.*, 881
Chirivì, G., Suyu, S. H., Grillo, C., et al. 2018, *Astron. Astrophys.*, 614, A8
Clemens, M. S., Bressan, A., Nikolic, B., et al. 2006, *Mon. Not. R. Astron. Soc.*, 370, 702
Coe, D., Salmon, B., Bradač, M., et al. 2019, arXiv: 1903.02002v1
Coe, D., Zitrin, A., Carrasco, M., et al. 2013, *Astrophys. J.*, 762, 32
Dalal, N. & Keeton, C. R. 2003, ArXiv: 0312072
de Blok, W. J. G. 2009, Advances in Astronomy, article id. 789293
Diego, J. M., Broadhurst, T., Benitez, N., Lim, J., & Lam, D. 2015, *Mon. Not. R. Astron. Soc.*, 449, 588
Diego, J. M., Tegmark, M., Protopapas, P., & Sandvik, H. B. 2007, *Mon. Not. R. Astron. Soc.*, 375, 958
Diemer, B. & Kravtsov, A. V. 2014, *Astrophys. J.*, 789, 1
D'Onghia, E. & Navarro, J. F. 2007, *Mon. Not. R. Astron. Soc.*, 380, L58
Draine, B. T. & Salpeter, E. E. 1979, *Astrophys. J.*, 231, 77
Dyson, F. W., Eddington, A. S., & Davidson, C. R. 1920, Mem. Roy. Astron. Soc., 62, 291
Einasto, J. 1965, Trudy Astrofizicheskogo Instituta Alma-Ata, 5, 87
Einstein, A. 1915, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), 844
Elbaz, D., Arnaud, M., & Boehringer, H. 1995, *Astron. Astrophys.*, 293, 337–346
Elíasdóttir, Á., Limousin, M., Richard, J., et al. 2007, arXiv:0710.5636
Ettori, S., De Grandi, S., & Molendi, S. 2002, *Astron. Astrophys.*, 391, 841
Faber, S. M., Dressler, A., Davies, R. L., et al. 1987, Faber S.M. (eds) Nearly Normal Galaxies. Santa Cruz Summer Workshops in Astronomy and Astrophysics. Springer, New York, NY, 175–183
Farouki, R. & Shapiro, S. L. 1981, *Astrophys. J.*, 243, 34
Fedeli, C., Bartelmann, M., & Moscardini, L. 2012, Journal of Cosmology and Astroparticle Physics, 2012
Feretti, L., Fusco-Femiano, R., Giovannini, G., & Govoni, F. 2001, *Astron. Astrophys.*, 373,

106–112
Ford, H. C., Clampin, M., Hartig, G., et al. 2003, Future EUV/UV and Visible Space Astrophysics Missions and Instrumentation, 4854, 81
Foreman-Mackey, D. 2016, Journal of Open Source Software, 1(2), 24
Gallagher, J. S. I. & Ostriker, J. P. 1972, *Astrophys. J.*, 77, 288
Garilli, B., Fumana, M., Franzetti, P., et al. 2010, ASTRONOMICAL SOCIETY OF THE PACIFIC, 122, 827–838
Garilli, G., Paioro, L., Scodeggio, M., et al. 2012, The Astronomical Society of the Pacific, 124, 1232
Gnedin, O. Y. & Ostriker, J. P. 2001, *Astrophys. J.*, 561, 61
Govoni, F., Markevitch, M., Vikhlinin, A., et al. 2004, The Astrophysical Journal, 605, 695–708
Graham, A. W., Merritt, D., Moore, B., Diemand, J., & Terzić, B. 2006, Astrophysical Journal Supplement, 132, 2701
Graur, O., Rodney, S. A., Maoz, D., et al. 2013, *Astrophys. J.*, 783, 28
Grillo, C., Gobat, R., Presotto, V., et al. 2014, *Astrophys. J.*, 786, 11
Grillo, C., Karman, W., Suyu, S. H., et al. 2016, *Astrophys. J.*, 822, 78
Grillo, C., Suyu, S. H., Rosati, P., et al. 2015, *Astron. Astrophys.*, 800, 38
Gunn, J. E. & Gott, J. R. 1972, Astrophysical Journal, 176, 1
Ha, J.-H., Ryu, D., & Kang, H. 2018, *Astrophys. J.*, 857, 26
Harrison, E. R. 1974, *Astrophys. J.*, 191, L51
Harvey, D., Massey, R., Kitching, T., Taylor, A., & Tittley, E. 2015, Science, 347, 1462
Hastings, W. 1970, Biometrika, 57, 97
Hattori, S., Ota, N., Zhang, Y.-Y., Akamatsu, H., & Finoguenov, A. 2017, Astronomical Society of Japan, 69, 39
Herbig, T. & Birkinshaw, M. 1994, American Astronomical Society, 26, 1403
Hetznecker, H. & Burkert, A. 2006, *Mon. Not. R. Astron. Soc.*, 370, 1905
Host, O. 2012, *Mon. Not. R. Astron. Soc.*, 420, L18
Jauzac, M., Harvey, D., & Massey, R. 2018, *Mon. Not. R. Astron. Soc.*, 477, 4046
Jee, M. J., Dawson, K. S., Hoekstra, H., et al. 2011, *Astrophys. J.*, 737, 59
Jullo, E. & Kneib, J.-P. 2009, *Mon. Not. R. Astron. Soc.*, 395, 1319
Jullo, E., Kneib, J.-P., Limousin, M., et al. 2007, New Journal of Physics, 9, 447
Jullo, E., Natarajan, P., Kneib, J. P., et al. 2010, Science, 329, 924
Kaiser, N. 1986, *Mon. Not. R. Astron. Soc.*, 222, 323
Kassiola, A. & Kovner, I. 1993, *Astrophys. J.*, 417, 450
Kaurov, A. A., Dai, L., Venumadhav, T., Miralda-Escudé, J., & Frye, B. e. a. 2019, *Astrophys. J.*, 880
Keeton, C. R. 2001, ArXiv: 0102340
Kelly, P. L., Diego, J. M., & Weiner, B. J. 2018, Nature, 2, 334
Kimble, R. A., MacKenty, J. W., O'Connell, R. W., & Townsend, J. A. 2008, Space Telescopes and Instrumentation 2008: Optical, Infrared, and Millimeter, 7010, 70101E
Klypin, A., Kravtsov, A. V., Valenzuela, O., & Prada, F. 1999, *Astrophys. J.*, 522, 89
Kneib, J.-P., Ellis, R. S., Smail, I., Couch, W. J., & Sharples, R. M. 1996, *Astrophys. J.*, 471, 643
Kovner, I. 1987, Astrophys. J., 312, 22
Lacey, C. & Cole, S. 1993, MNRAS, 262, 627
Lefor, A. T., Futamase, T., & Akhlaghi, M. 2013, New Astronomy Reviews, 57, 1

Limber, D. N. 1959, *Astrophys. J.*, 130, 414
Limousin, M., Kneib, J.-P., & Natarajan, P. 2005, MNRAS, 356, 309
Limousin, M., Richard, J., Kneib, J.-P., et al. 2008, *Astron. Astrophys.*, 489, 23
Limousin, M., Sommer-Larsen, J., Natarajan, P., & Milvang-Jensen, B. 2009, *Astrophys. J.*, 696, 1771
Lin, Y.-T., Stanford, S. A., Eisenhardt, P. R. M., et al. 2012, The Astrophysical Journal Letters, 745, L3
Livermore, R. C., Finkelstein, S. L., & Lotz, J. M. 2017, *Astrophys. J.*, 835, 113
Lotz, J. M., Koekemoer, A., Coe, D., et al. 2017, *Astron. Astrophys.*, 837, 97
Lotz, J. M., Koekemoer, A., Coe, D., et al. 2016, ArXiv: 1605.06567
Malhotra, S., Rhoads, J. E., & Turner, E. L. 1997, *Mon. Not. R. Astron. Soc.*, 288, 138
Maller, A. H., Dekel, A., & Somerville, R. 2002, *Mon. Not. R. Astron. Soc.*, 329, 423
Malumuth, E. M. & Kirshner, R. P. 1981, *Astrophys. J.*, 251, 508
Markevitch, M., Gonzalez, A. H., Clowe, D., et al. 2004, *Astrophys. J.*, 606, 819
Markevitch, M., Mushotzky, R., Inoue, H., et al. 1996, Astrophysical Journal, 456, 437
Markevitch, M. & Vikhlinin, A. 2001, Astrophysical Journal, 563, 95–102
Markevitch, M. & Vikhlinin, A. 2007, Physics Reports, 443, 1
Markevitch, M., Yamashita, K., Furuzawa, A., & Tawara, Y. 1994, Astrophysical Journal, 436, L71
Maurogordato, S., Cappi, A., Ferrari, C., et al. 2008, *Astron. Astrophys.*, 481, 593–613
Mazzotta, P., Moscardini, L., & Tormen, G. 2004, *Mon. Not. R. Astron. Soc.*, 354, 10
McCarthy, I. G., Frenk, C. S., Font, A. S., et al. 2008, *Mon. Not. R. Astron. Soc.*, 383, 593
McGee, S. L. & Balogh, M. L. 2010, *Mon. Not. R. Astron. Soc.*, 405, 2069
McLure, R. J., Pentericci, L., Cimatti, A., et al. 2018, *Mon. Not. R. Astron. Soc.*, 479, 25
Melin, J.-B., Bartlett, J. G., Cai, Z.-Y., et al. 2018, *Astron. Astrophys.*, 617, A75
Meneghetti, M., Bartelmann, M., Jenkins, A., & Frenk, C. 2007, *Mon. Not. R. Astron. Soc.*, 381, 171
Meneghetti, M., Fedeli, C., Pace, F., Gottlöber, S., & Yepes, G. 2010, *Astron. Astrophys.*, 519, A90
Meneghetti, M., Natarajan, P., & Coe, D. 2017, *Mon. Not. R. Astron. Soc.*, 472, 3177–3216
Meneghetti, M., Rasia, E., Merten, J., et al. 2009, *Astron. Astrophys.*, 514, A93
Merten, J. 2016, *Mon. Not. R. Astron. Soc.*, 461, 2328
Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., & Teller, E. 1953, Chemical Physics, 21(6), 1087
Molnar, S. M. 2016, Frontiers in Astronomy and Space Science, 522, 89
Moore, B., Ghigna, S., Governato, F., et al. 1999, Astrophysical Journal Letters, 524(1), L19–L22
Morandi, A., Ettori, S., & Moscardini, L. 2007, *Mon. Not. R. Astron. Soc.*, 379, 518–534
Morandi, A., Pedersen, K., & Limousin, M. 2010, *Astrophys. J.*, 713, 491
Morgan, W. W. & White, R. A. 1975, *Astrophys. J.*, 199, 545
Munari, E., Grillo, C., De Lucia, G., Biviano, A., & Annunziatella, M. e. a. 2016, The Astrophysical Journal Letters, 827, L5
Murante, G., Giovalli, M., Gerhard, O., et al. 2007, *Mon. Not. R. Astron. Soc.*, 377, 2
Narayan, R. & Bartelmann, M. 1995, *Formation of Structure in the Universe, 1995 Jerusalem Winter School* (Cambridge University Press)
Natarajan, P., Kneib, J., Smail, I., et al. 2009, *Astrophys. J.*, 693, 970
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, *Astrophys. J.*, 462, 563

Nesvadba, N. P. H., Lehnert, M. D., Eisenhauer, F., et al. 2006, *Astrophys. J.*, 650, 693
Newman, A. B., Treu, T., Ellis, R. S., & Sand, D. J. 2013a, *Astrophys. J.*, 765, 25
Newman, A. B., Treu, T., Ellis, R. S., Sand, D. J., & Meneghetti, M. 2011, The Astrophysical Journal Letters, 728, L39
Newman, A. B., Treu, T., Ellis, R. S., et al. 2013b, *Astrophys. J.*, 765, 24
Nord, M., Basu, K., Pacaud, F., Ade, P. A. R., & Bender, A. N. 2009, *Astron. Astrophys.*, 506, 623
Oemler, A. J. 1974, *Astrophys. J.*, 194, 1
Ogrean, G. A., van Weeren, R. J., Jones, C., Clarke, T. E., & Sayers, J. e. a. 2015, *Astrophys. J.*, 812, 153
Oguri, M. & Blandford, R. D. 2008, *Mon. Not. R. Astron. Soc.*, 392, 930
Oguri, M., Hennawi, J. F., Gladders, M. D., et al. 2009, *Astrophys. J.*, 699, 1038
Okabe, N., Bourdin, H., Mazzotta, P., & Maurogordato, S. 2011, New Journal of Physics, 741, 116
Ostriker, J. P. & Hausman, M. A. 1977, *Astrophys. J.*, 217, L125
Ostriker, J. P. & Tremaine, S. D. 1975, *Astrophys. J.*, 202, L113
Ota, N., Nagayoshi, K., Pratt, G. W., et al. 2014, *Astron. Astrophys.*, 562, A60
Parry, W. G., Grillo, C., Mercurio, A., et al. 2016, *Mon. Not. R. Astron. Soc.*, 458, 1493–1503
Peebles, P. J. E. & Ratra, B. 2003, Reviews of Modern Physics, 75, 559
Peterson, J. R. & Fabian, A. C. 2006, Physics Reports, 427, 1
Postman, M., Coe, D., Benitez, N., et al. 2012, *Astrophys. J.*, 199, 25
Postman, M., Franx, M., Cross, N. J. G., et al. 2005, *Astrophys. J.*, 623, 721
Pratt, G. W. & Arnaud, M. 2002, *Astron. Astrophys.*, 394, 375
Quintana, H. & Lawrie, D. G. 1982, *Astrophys. J.*, 87, 1
Quintana, H. & Melnick, J. 1982, Astronomical Journal, 87, 972
Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, The Astrophysical Journal Letters, 802, L19
Rodney, S. A., Balestra, I., Bradač, M., et al. 2018, Nature, 2, 324
Rood, H. J. & Sastry, G. N. 1971, Publ. Astron. Soc. Pacific, 83, 313
Rosati, P., Balestra, I., Grillo, C., & Mercurio, A. e. a. 2014, The Messenger, 158, 48–53
Rosati, P., Borgani, S., & Norman, C. 2002, Annual Review of Astronomy and Astrophysics, 40, 539
Rudd, D. H., Zentner, A. R., & Kravtsov, A. V. 2008, *Astrophys. J.*, 672, 19
Salmon, B., Coe, D., Bradley, L., et al. 2017, ArXiv: 1710.08930
Salmon, B., Coe, D., Bradley, L., et al. 2018, ApJL, 864, L22
Salzano, V., Rodney, S. A., Sendra, I., et al. 2013, *Astron. Astrophys.*, 557, A64
Schneider, P. 2015, Extragalactic Astronomy and Cosmology - An introduction (Springer)
Schneider, P., Ehlers, J., & Falco, E. E. 1992, *Gravitational Lenses* (Springer)
Sereno, M. & Zitrin, A. 2012, *Mon. Not. R. Astron. Soc.*, 419, 3280
Simionescu, A., Roediger, E., Nulsen, P. E. J., et al. 2009, *Astron. Astrophys.*, 495, 721
Skilling, J. 2004, URL: http://www.inference.phy.cam.ac.uk/bayesys/
Soldner, J. 1804, Berliner Astron. Jahrb., 161
Somerville, R. S. & Kolatt, T. S. 1999, MNRAS, 305, 1
Soto, K. T., Lilly, S. J., Bacon, R., Richard, J., & Conseil, S. 2016, *Mon. Not. R. Astron. Soc.*, 458, 3210
Soucail, G. 2012, *Astron. Astrophys.*, 540, 61

Spergel, D. N. & Steinhardt, P. J. 2000, Physical Review Letters, 84, 3760
Spitzer, L. J. 1946, reprinted with commentary in Astronomy Quarterly, 1990, 7, 131
Springel, V. & Farrar, G. 2007, *Mon. Not. R. Astron. Soc.*, 380, 911
Sunyaev, R. A. & Zel'dovich, Y. B. 1970, Astrophysics and Space Science, 7, 3
Suyu, S. H. & Halkola, A. 2010, *Astron. Astrophys.*, 524, A94
Thölken, S., Reiprich, T. H., Sommer, M. W., & Ota, N. 2018, *Astron. Astrophys.*, 619, A68
Tokunaga, A. T. & Vacca, W. D. 2005, The Astronomical Society of the Pacific, 117, 830
Toomre, A. & Toomre, J. 1972, *Astrophys. J.*, 178, 623
Treu, T., Brammer, G., Diego, J. M., et al. 2016, *Astrophys. J.*, 817, 60
Tsapras, Y. 2018, Geosciences, 8, 365
Umetsu, K., Medezinski, E., Nonino, M., et al. 2014, *Astrophys. J.*, 795, 163
Umetsu, K., Zitrin, A., Gruen, D., et al. 2016, *Astrophys. J.*, 821, 116
van Albada, G. B. 1960, Bulletin of the Astronomical Institutes of the Netherlands, 15, 165
van Uitert, E., Hoekstra, H., Schrabback, T., et al. 2012, *Astrophys. J.*, 545, A71
Vikhlinin, A., Kravtsov, A., Forman, W., et al. 2006, *Astrophys. J.*, 640, 691
Vitvitska, M., Klypin, A. A., Kravtsov, A. V., et al. 2002, *Astrophys. J.*, 581, 799
Weilbacher, P. M., Streicher, O., Urrutia, T., et al. 2014, Astronomical Data Analysis Software and Systems XXIII ASP Conference Series, 485, 451
Werner, N., Finoguenov, A., Kaastra, J. S., et al. 2008, *Astron. Astrophys.*, 482, L29
White, S. D. M. 1984, *Astrophys. J.*, 286, 38
Whitmore, B. C. & Gilmore, D. M. 1991, *Astrophys. J.*, 367, 64
Xue, Y.-J. E. & Wu, X.-P. 2000, *Astrophys. J.*, 538, 65
Zabludoff, A. I. & Mulchaey, J. S. 1998, *Astrophys. J.*, 496, 39
Zheng, W., Postman, M., & van der Wel, A. 2012, Nature, 489, 406
Zitrin, A., Moustakas, J., Bradley, L., et al. 2012, The Astrophysical Journal Letters, 747, L9
Zwicky, F. 1933, Helvetica Physica Acta, 6, 110
Zwicky, F., Herzog, E., Wild, P., Karpowicz, M., & Kowal, C. 1961-1968, Pasadena: California Institute of Technology, 1-6