NEW INSIGHTS ON THE INNER MASS DISTRIBUTION OF MASSIVE GALAXY CLUSTERS FROM A COMBINATION OF STRONG LENSING AND GALAXY KINEMATICS

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Corso di Dottorato di ricerca in Astronomia

CICLO XXXII

NEW INSIGHTS ON THE INNER MASS DISTRIBUTION OF MASSIVE GALAXY CLUSTERS FROM A COMBINATION OF STRONG LENSING AND GALAXY KINEMATICS

Tesi redatta con il contributo finanziario dell'Istituto Nazionale di Astrofisica

Coordinatore:	**Ch.mo Prof. Giampaolo Piotto**
Supervisore:	**Ch.mo Prof. Piero Rosati**
Co-supervisori:	**Prof. Giulia Rodighiero** **Dott. Amata Mercurio** **Ch.mo Prof. Alberto Franceschini**

Dottorando:
Pietro Bergamini

Academic Year 2019/2020

Abstract

Galaxy clusters are important astrophysical laboratories to study the nature of Dark Matter (DM), whose physical properties are still unknown. In particular, a detailed investigation of the mass distribution of cluster halos, by dissecting the DM and baryonic components, can provide stringent tests on the Cold Dark Matter paradigm of structure formation. Over the last decade, strong gravitational lensing has become one of the most powerful techniques to study the total mass distribution in the Universe, particularly on galaxy and cluster scale. In this context, dedicated large imaging surveys with the Hubble Space Telescope (HST) and ground-based spectroscopic campaigns on sizeable samples of massive galaxy clusters have stimulated a new generation of high-precision strong lensing models, via the identification of a large number of multiple images and cluster members.

In this thesis, we extend beyond the state of the art these recent cluster lens models, incorporating the stellar kinematics of a significant fraction of cluster galaxies, measured with the MUSE integral field spectrograph on the VLT. This study focuses on three massive clusters MACS J1206.2−0847, MACS J0416.1−2403, and Abell S1063 at $z \sim 0.4$ with HST imaging and VLT spectroscopy data of unprecedented quality. Specifically, we measured the stellar velocity dispersion of 40-60 member galaxies per cluster with MUSE, covering 4-5 magnitudes down to $m_{F160W} \simeq 21.5$. The robustness and accuracy of the velocity dispersion measurements were tested with extensive spectral simulations. We determined the normalization and slope of the galaxy Faber-Jackson relation in each cluster, and include this prior information in the cluster lens models. We find that using this novel technique, the inherent degeneracy between different mass components and possible systematics on model parameters are strongly reduced and the mass density profiles of cluster galaxies are now robustly constrained. Once re-normalized to the same absolute luminosity, our kinematic lens models predict consistent masses and sizes of sub-halos as a function of galaxy velocity dispersions. Moreover, extending previous findings, we derive consistent sub-halo mass and velocity dispersion functions for the three clusters.

By comparing the observed sub-halo mass distribution from our cluster lens models with the predictions of high-resolution N-body and hydrodynamical cosmological simulations, we find a lack of compact sub-structures in the corresponding inner regions of simulated clusters, whereas the sub-halo mass functions are found in good agreement. We still do not understand whether the

origin of these differences has to be ascribed to numerical or resolution effects in the simulations, or to some physical aspect missing in the Cold Dark Matter framework.

An additional method to investigate the mass distribution of cluster sub-halos is to exploit galaxy scale strong lensing systems (GSSLS), in which a single cluster member produces highly magnified multiple images on kpc scale around lens galaxies. We show how strong lensing modeling of GSSLS in the cluster field, in combination with spatially resolved stellar kinematics of the lens galaxies, can further constrain the structure and sizes of cluster sub-halos.

Finally, in an effort to include in our lens models the measured internal galaxy velocity dispersions and the observed scatter of the Faber-Jackson relation, we developed and made public a python module which expands the capabilities of common lens modeling tools.

The methodologies of high-precision lens modeling developed in this thesis will find important applications in large area surveys, such as the Large Synoptic Survey Telescope (LSST) and the *Euclid* satellite, when large numbers of cosmic lenses will be discovered. Applications include the exploitation of lensing clusters as powerful cosmic telescopes to investigate galaxies in the early Universe and cluster cosmography with gravitational time delay techniques.

Contents

To Maria Elena and my family.

Acknowledgments

First and foremost, I would like to thank Piero Rosati that guide and inspirited me during my PhD years. His ideas and teachings were fundamental for the results presented in this thesis and for my future career.
I owe special thanks to Amata Mercurio, Massimo Meneghetti, Claudio Grillo and Gabriel Caminha for all the critical discussions and support throughout my work. This thesis is the result of a close and fruitful collaboration among exceptional researchers, including also Francesco Calura, Carlo Giocoli, Eros Vanzella, Roberto Gilli and Adriano Agnello. I would like to thank them all for the numerous and essential advises.
Furthermore, I owe a sincere thank you to Alberto Franceschini and Giulia Rodighiero who have accompanied me during my stay in Padova.
Last, but not least, I thank Maria Elena and my family. Their encouragements, support and affection are the main ingredients of all my accomplishments.

List of publications

The results presented in this thesis are published in the following scientific articles.

Published and submitted papers

Title: Enhanced cluster lensing models with measured galaxy kinematics

Authors: Bergamini P., Rosati P, Mercurio A., Grillo C., Caminha G. B., Meneghetti M., Agnello A., Biviano A., Calura F., Giocoli C., Lombardi M., Rodighiero G. and Vanzella E.

Status and Journal: Astronomy & Astrophysics, Volume 631, id.A130, 16 pp. (Bergamini et al. 2019)

Title: Strong lensing models of eight CLASH clusters from extensive spectroscopy: accurate total mass reconstructions in the cores

Authors: Caminha G. B., Rosati P., Grillo C., Rosani G., Caputi K. I., Meneghetti M., Mercurio A., Balestra I., Bergamini P., Biviano A., Nonino M., Umetsu K., Vanzella E., Annunziatella M., Broadhurst T., Delgado-Correal C., Demarco R., Lombardi M., Maier C., Zitrin A.

Status and Journal: arXiv:1903.05103 (Caminha et al. 2019) A&A, in press

Title: Rejuvenated galaxies with very old bulges at the origin of the bending of the main sequence and of the "green valley"

Authors: Mancini C., Daddi E., Juneau S., Renzini A., Rodighiero G., Cappellari M., Rodrguez-Muoz L., Liu D., Pannella M., Baronchelli I., Franceschini A., Bergamini P., D'Eugenio C., Puglisi A.

Status and Journal: Monthly Notices of the Royal Astronomical Society, Volume 489, Issue 1, October 2019, Pages 12651290 (Mancini et al. 2019)

Title: Magnifying the Early Episodes of Star Formation: Super Star Clusters at Cosmological Distances

Authors: Vanzella E., Castellano M., Meneghetti M., Mercurio A., Caminha G. B.,

Cupani G., Calura F., Christensen L., Merlin E., Rosati P., Gronke M., Dijkstra M., Mignoli M., Gilli R., De Barros S., Caputi K., Grillo C., Balestra I., Cristiani S., Nonino M. Giallongo E., Grazian A., Pentericci L., Fontana A., Comastri A., Vignali C., Zamorani G., Brusa M., Bergamini P., Tozzi P.

Status and Journal: The Astrophysical Journal, Volume 842, Number 1 (Vanzella et al. 2017)

Title: A new gap between observations and theory revealed by cluster lensing

Authors: Meneghetti M., Davoli G., Bergamini P., Rosati P., Natarajan P., Giocoli C., Caminha G. B., Metcalf R. B., Rasia E., Borgani S., Calura F., Grillo C., Mercurio A., Vanzella E.

Status and Journal: Submitted to Science

Title: An accurate strong lensing model of Abell 2163 core

Authors: Rescigno U., Grillo C., Lombardi M., Rosati P., Caminha G. B., Meneghetti M., Mercurio A., Bergamini P., Coe D.

Status and Journal: Submitted to A&A

Title: Accurate Cluster Lensing with Hierarchical Inference: Formalism, functional tests, and public code release

Authors: Bergamini P., Agnello A., Caminha G. B.

Status and Journal: Submitted to A&A

Papers close to submission

Title: CLASH-VLT: a full dynamical reconstruction of the mass profile of RXJ 2248.7-4431 from 1 Kpc out to 4 Mpc

Authors: Sartoris B., Biviano A., Rosati P., Mercurio A., Grillo C., Caminha G. B., Ettori S., Nonino M., Bergamini P., Girardi M., Umetsu K.

Status and Journal: Close to submission to A&A

Chapter 1

Introduction

Galaxy clusters are the largest gravitationally bound structures in the Universe and they play a fundamental role both in Cosmology and Astrophysics. The enormous gravitational attraction generated by their large masses (about 10^{14} - 10^{15} solar masses) is able to trap hundreds of galaxies that move in the cluster potential wells at thousands of kilometers-per-seconds. Comparing the dynamical cluster masses derive from the virial theorem, with the luminous masses estimated considering the brightness and the number of cluster galaxies, in 1937, Zwicky found the first observational evidence of the Dark Matter (DM) (Zwicky 1937). Now we know that approximately 85-90% of the mass of galaxy clusters is in the form of DM, thus they are natural laboratories to study the macroscopic proprieties and indirectly the physics of this elusive form of matter. The remaining 10-15% of the cluster mass is made by baryons, primarily in the form of hot X-ray emitting plasma, while only the 1-2% is in stars.

Many independent studies support the hypothesis of a Cold Dark Matter (CDM, Peebles 1982; Bond et al. 1982; Blumenthal et al. 1982, 1984), made by collisionless and not relativistic particles, that drives the formation and the evolution of all the structures in the expanding Universe, from the dwarf galaxies up to the most massive clusters. This cosmological paradigm is known as the ΛCDM model. In the ΛCDM Universe, cosmic structures form in a hierarchical fashion, thus larger halos are the result of the aggregation of smaller sub-halos. In particular, massive galaxy clusters continuously accrete new galaxies over cosmic time, enriching in this way their sub-halo population, embedded into the hot, eventually virialzed intra-cluster plasma (see Fig. 1.1).

The Einstein theory of the General Relativity predicts that masses are able to bend the light rays emitted by distant sources through the curvature induced on

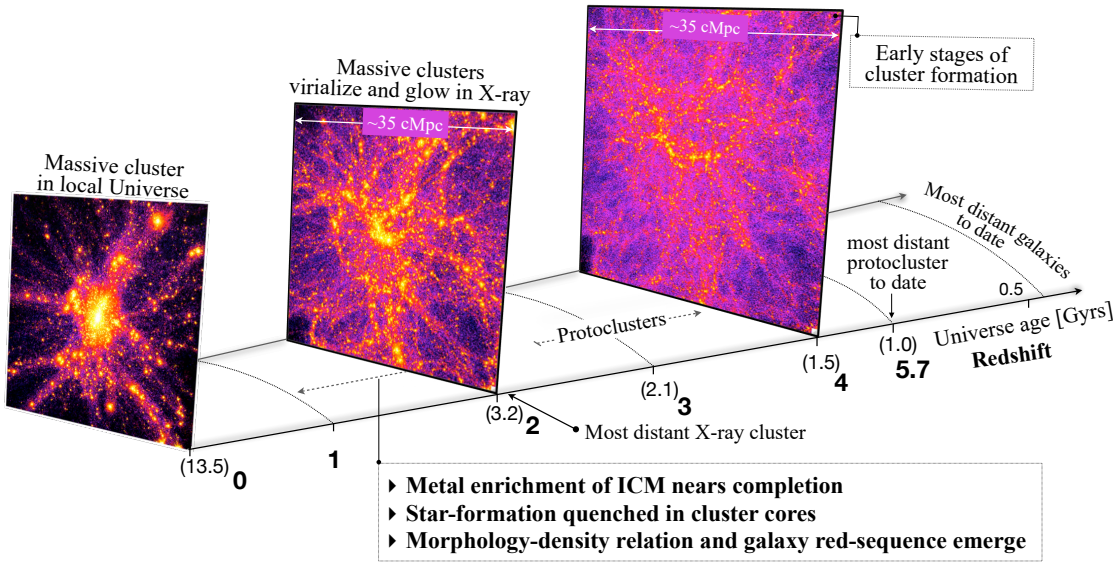


Figure 1.1: A schematic diagram of the cosmic evolution of a massive galaxy cluster. The cluster evolution is followed from the early stages of assembly, approximately 1 billion years after the Big Bang, through its progressive growth until a massive bound structure ($M \sim 10^{15}\, M_{\odot}$) is assembled in the nearby Universe, $\sim$ 10 Gyrs later. The axis labels indicate the cosmic time (age of the Universe) and corresponding redshifts. The three images represent the density maps of dark matter, extracted from cosmological hydrodynamical simulations (from Borgani & Kravtsov 2011), at three representative epochs ($z = 4,\ 2,\ 0$). When a protocluster is formed, the DM-dominated high-density peaks host star-forming galaxies, so protocluster regions can be discovered as prominent concentrations of Lyman-α emitters, back to when the Universe was 1 Gyr old (see observations in Jiang et al. 2018, covering a similar co-moving area of 35 co-moving Mpc (cMpc). (Figure from Rosati 2018)

the space-time. When the mass distributions are sufficiently compact and heavy to produce multiple images of a source this phenomenon takes the name of strong gravitational lensing.

Strong lensing is one of the most powerful techniques to probe the total projected mass distribution of the inner regions of galaxy clusters and therefore the DM distribution, once the baryonic mass components are independently mapped. Combining strong lensing with results of other mass traces, such as weak lensing (Umetsu et al. 1999, 2011, 2014; Schneider et al. 2000; Bartelmann & Schneider 2001; Umetsu & Broadhurst 2008; Umetsu 2010, 2013; Hoekstra et al. 2013; Gruen et al. 2013; Merten et al. 2015; Melchior et al. 2015), galaxy dynamics (Biviano et al. 2013; Stock et al. 2015) and X-ray hydrostatic analysis (Gómez et al. 2012; Ettori et al. 2013; Caminha et al. 2017b; Balestra et al. 2016), we can reconstruct the cluster mass density profile over a wide radial range, from kpc to Mpc scales

(Newman et al. 2009, 2011; Biviano et al. 2013; Balestra et al. 2016; Caminha et al. 2017b).
By characterizing the substructure of cluster cores on different scales and the mass density profile of the innermost regions, one can test the ΛCDM structure formation scenario and indirectly constrain DM physical properties, by comparing reconstructed mass maps with the latest high-resolution N-body and hydrodynamical cosmological simulations (e.g. Diemand & Moore 2011; Planelles et al. 2014; Genel et al. 2014). Indeed, large deviations from theoretical expectations would imply a different cluster formation paradigm respect to the ΛCDM predictions, or a different DM nature. In particular, the collisionless nature of DM particles could be tested in the high density cluster cores where a non-zero DM cross section might well modify the inner slope of the cluster mass profile, or the spatial segregation of the DM sub-halos with respect to the collisionless stellar component that they host.

Since the luminosity of lensed sources is magnified up to hundreds of times, galaxy clusters can also be used as powerful cosmic telescopes to study astronomical objects out to very high redshifts, well into the first billion years after the Big Bang (e.g. Zheng et al. 2012; Hashimoto et al. 2018). Gravitational lensing is particularly efficient in unveiling faint and distant, high-surface density sources. It does not only amplify but also magnifies distant galaxies. For example, recent studies exploiting cluster lensing (e.g. Vanzella et al. 2017; Karman et al. 2017) have found a rich population of low-mass (down to $10^6\,M_\odot$), low-luminosity, stellar ensembles, often with an associated strong Lyα emission, that may be the stellar systems which will soon assemble into the first galaxies. Some of those objects, with physical sizes of few tens of parsecs, may be globular cluster progenitors. The ionizing radiations emitted by these high-redshift, low-luminosity stellar ensembles may play a key role in the reionization process of the Universe.

In recent years, dedicated HST observational campaigns, such as the Cluster Lensing And Supernova survey with Hubble (CLASH, Postman et al. 2012), the Hubble Frontier Field campaign (HFF, Lotz et al. 2017) and the Reionization Lensing Cluster Survey (RELICS, Coe et al. 2019) have provided multi-band data with unprecedented quality on a sizable set of massive galaxy clusters. In parallel, extensive spectroscopic campaigns with highly-multiplexing multi-slit instruments on the ground (e.g. CLASH-VLT, Rosati et al. in prep., Balestra et al. 2016), with the MUSE integral field spectrograph on the VLT (Bacon et al. 2012), or grism spectroscopy with HST (GLASS, Treu et al. 2015, Schmidt et al. 2014), have secured redshifts for hundreds of multiply lensed images in cluster cores, as well as identified large samples of cluster galaxies. The combination of these new imaging and spectroscopic data sets have enabled the development of new high-precision

parametric strong lensing models based on 50-100 bona-fide multiple images per cluster and highly complete samples of cluster members (e.g. Richard et al. 2014, Grillo et al. 2015, Jauzac et al. 2015, Limousin et al. 2016, Kawamata et al. 2016, Lagattuta et al. 2017, Lagattuta et al. 2019, Caminha et al. 2016, Caminha et al. 2017a, Caminha et al. 2017b, Bonamigo et al. 2018, Caminha et al. 2019). These models are able to predict the multiple image positions with an average accuracy of 0.5″ or lower.
Nonetheless, the relatively large number of parameters used to describe the cluster mass distribution still suffer from strong internal degeneracies. The generally complex cluster mass distribution in parametric lens models is generally separated into a cluster-scale diffuse component, made of one or more large halos, and a clumpy distribution traced by cluster galaxies (Natarajan & Kneib 1997). The latter describes the sub-halo population of DM halos (De Lucia et al. 2004; Giocoli et al. 2010a) and hence the inner substructure of the cluster mass distribution (Springel et al. 2001b). In the effort to limit the number of free parameters, this sub-halo component is generally modeled adopting two scaling relations, which link the internal velocity dispersion and size of each halo with the luminosity of member galaxies. A fixed mass-to-light scaling is thereby assumed for all sub-halos. Since strong lensing models constrain the total projected mass within each family of multiple image positions, a certain amount of degeneracy is always present between the velocity dispersion, the size (or equivalently the profile) and the shape of each halo component of the mass distribution. Given the cross-talks among different halo components in strong lensing models, such a degeneracy can lead to systematics in the reconstruction of the substructure at different scales, for example with a transfer of mass between the clumpy and diffuse mass components. As we demonstrate in this work, such a degeneracy can be broken and significantly reduced, on the scale of the sub-halos, by using an independent measurement of the internal stellar velocity dispersion of cluster members. Thus, besides the identification of cluster members and multiple images, high-dispersion spectroscopic data provide further constraints on mass models through the measurements of galaxy stellar kinematics.

A combination of lensing and kinematic measurements has long been exploited in the study of the mass density profile of field early-type galaxies. This combination has proved to be particularly effective, since the two diagnostics complement each other (e.g., Treu & Koopmans 2004; Czoske et al. 2008; Barnabè et al. 2009), breaking the mass-anisotropy and mass-sheet (e.g., Falco et al. 1985; Meylan et al. 2006) degeneracies of the dynamical and lensing analyses, respectively. Joint strong lensing and stellar-dynamical studies have been used to determine the average logarithmic density slope of the total mass inside the Einstein radius (Treu & Koopmans 2004; Koopmans et al. 2009) and to decompose the total mass

distribution into luminous and dark components of the lens galaxies belonging to the Lenses Structure and Dynamics (LSD) and Sloan Lens ACS (SLACS) surveys. These two surveys have measured a remarkably homogeneous total (luminous and dark) mass density profile, that is consistent with an isothermal one (i.e., $\rho \propto 1/r^2$) out to a few hundreds of kiloparsecs (see Gavazzi et al. 2007; Bolton et al. 2008). A one-component isothermal model is fully characterized by the value of an effective velocity dispersion, approximated within $\lesssim 3\%$ by the value of the galaxy central stellar velocity dispersion (i.e., the velocity dispersion of the stars projected within a disk of radius $R_e/8$; see, e.g., Treu et al. 2006; Bolton et al. 2008). This result is theoretically supported by the Jeans equation for realistic stellar density distributions (e.g., Jaffe 1983; Hernquist 1990) embedded in a globally isothermal distribution (see Kochanek 1993), as observed in samples of nearby and luminous early-type galaxies (e.g., Kochanek 1994; Grillo et al. 2008).
A first attempt to include the cluster member stellar kinematics to reduce the inner mass degeneracies among cluster lens model parameters was developed by Monna et al. (2015). In particular, they demonstrated that the inclusion of this extra kinematic information is important to accurately characterize the sub-halo component of galaxy clusters, and that a meaningful scaling relation between galaxy halo truncation radii and velocity dispersions can be derived. Additional constraints on halo sizes can be obtained by modeling the surface brightness distribution of strong lensing features around individual member galaxies (see also Suyu & Halkola 2010, Eichner et al. 2013 who did not use galaxy kinematics).

In this work, we extend beyond the-state-of-the-art the current lensing models. In particular, we develop a strategy to incorporate in a self-consistent fashion the measured stellar kinematics of cluster members into the parametric strong lensing models of the clusters. To this aim, we take advantage the same data from the VLT/MUSE spectrograph used for measuring redshifts of multiply lensed images. Moreover, we use the measured galaxy stellar velocity dispersions, sometimes in the form of σ-profiles, to accurately determine the mass distribution of several galaxy scale strong lensing systems embedded in massive clusters. The robust characterization of cluster sub-halo populations that we obtain, particularly the distribution of their sizes and masses, as well as their abundance, can offer a critical test of the predictions of cosmological simulations (e.g. Limousin et al. 2009, Grillo et al. 2015, Munari et al. 2016, Natarajan et al. 2009, 2017). Such a comparison can shed light on baryonic processes shaping cluster substructure, dynamical processes leading to halo stripping in different environments, and indirectly on the nature of dark matter (Despali & Vegetti 2017; Chua et al. 2017; Nipoti et al. 2018; Niemiec et al. 2019).

The thesis is organized as follows. Chapter 2 summarizes the main physical proprieties of elliptical galaxies and in particular the equations governing their stellar dynamics. Indeed, early-type galaxies are the predominant galaxy population in massive clusters, especially in the inner regions where the strong lensing effects become important. In this chapter, we also describe the fundamental expressions characterizing the dual pseudo isothermal mass distribution used to parameterize the member galaxies in the cluster lens models. An important result is the derivation of a projection coefficient necessary to translate the cluster member measured velocity dispersions into the central velocity dispersions of their mass models.
An overview of the basic concepts of the gravitational lensing theory is in Chapter 3. Furthermore we describe the public software `LensTool` (Kneib et al. 1996; Jullo et al. 2007; Jullo & Kneib 2009) used to model the cluster mass distributions. The sub-halos in the cluster lens models suffer from a strong mass-concentration degeneracy that can be broken using galaxy stellar kinematic measurements. The reasons at the origin of this degeneracy are discussed at the end of Chapter 3.
After these two introductory chapters, in Chapter 4 we present our imaging and spectroscopic data sets and in particular the scientific instruments and the surveys that have collected the observations used in the rest of the thesis.
In Chapter 5, we show how we measure velocity dispersions of early-type galaxies from the observed spectra. To perform these measurements, we exploit the public software `pPXF` developed by Cappellari & Emsellem (2004) and upgraded as described in Cappellari (2017). In the final part of this chapter, we use extensive spectra simulations to test the `pPXF` performances when measuring the cluster member velocity dispersions. The simulations are also used to quantify realistic statistical errors on the velocity measurements and to verify the presence of systematic uncertainties.
In Chapter 6, we discuss the methodology that allows us to include the galaxy stellar kinematics into our cluster lens models, using high-quality data ofr three clusters. In particular, we derive kinematic priors for the sub-halo component of the lens models, from the measured velocity dispersion of cluster members. These priors strongly reduce inherent degeneracies of the lens models and permit an accurate description of the sub-halo population of the clusters. We also present the velocity dispersion functions for the three galaxy clusters.
In Chapter 7, we study three examples of galaxy scale strong lensing systems inside two massive galaxy clusters. The additional multiple images belonging to these systems, again combined with kinematic measurements, allow an accurate reconstruction of the mass distributions of four cluster member galaxies.
In Chapter 8, we compare the results of the three cluster lens models with kinematic priors developed in Chapter 6, with the predictions of high-resolution N-body and

Hydrodynamic cosmological simulations.
Finally, in Chapter 9, we develop a Hierarchical Bayesian inference formalism to include in cluster lens models the intrinsic scatter in the scaling relation of the cluster member population. To this aim, we build a modular wrapper of common lens modeling codes (e.g. `LensTool`), which also ensures that further constraints, e.g. from time delays, flux-ratios and shapes of background sources, can be easily included in the model inferences. A summary of the thesis results and of the future perspectives is in Chapter 10.

Chapter 2

Elliptical galaxies

This is the first of two theoretical chapters introducing the physical and mathematical concepts necessary for the development of this thesis.
The chapter focuses on the main proprieties of the elliptical galaxies, and in particular on equations governing their stellar inner kinematics and dynamics. Despite the ellipticals represent only the ∼20% of the total number of galaxies in the field, they are the dominant galaxy population in high density environments, such as clusters of galaxies (Dressler 1980, 1984; Hogg et al. 2004; Renzini 2006).

Particularly important is the Section. 2.5 where we define the dual pseudo isothermal elliptical mass density distribution. This mass profile is commonly used to describe the cluster member galaxies in cluster lens models. In Section 2.6, we show the most important scaling relations linking the main physical quantities of elliptical galaxies.

2.1 Elliptical galaxies in the Universe: The morphological classification

Galaxies in the Universe come in a lot of different shapes. While some galaxies can have a smooth elliptical light profile, others can have spiral arms departing from a central bulge, or others again an irregular morphology.
The basis of the traditional morphological classification come from Edwin Hubble that in 1926, in his paper Hubble (1926), and later inside his monograph "The Realm of the Nebulae"(Hubble 1936), divided for the first time the galaxies in what is now known as the Hubble sequence, or Hubble tuning-fork diagram (see Fig. 2.1).

The Hubble's diagram arranges galaxies into a continuous sequence of morphological types, with elliptical galaxies at the left end, spiral galaxies on the right, and irregular galaxies at the extreme right side of the diagram. The spiral galaxies are also divided in two branches, one that contains spirals with an elliptical bulge at their center, and the other containing spirals with a bar-shape bulge.
In the morphological context, elliptical galaxies are sometimes indicated as "early-type" galaxies while spirals and irregulars (the galaxies on the right end of the Hubble diagram) as "late type galaxies". This misleading nomenclature comes from the Hubble's original idea that galaxies evolve in time from the left to the right part of his diagram. However, now we know that galaxy evolution is extremely more complicated than the Hubble original picture and this nomenclature, although still used, lost its original meaning.
Nowadays, the Hubble sequence remains a fundamental tool to classify galaxies in the Universe and it was found that a lot of independent galaxy properties (such as the integrated color, the fraction of gas mass, the stellar composition, etc.) strongly correlate with the Hubble's morphological classes.

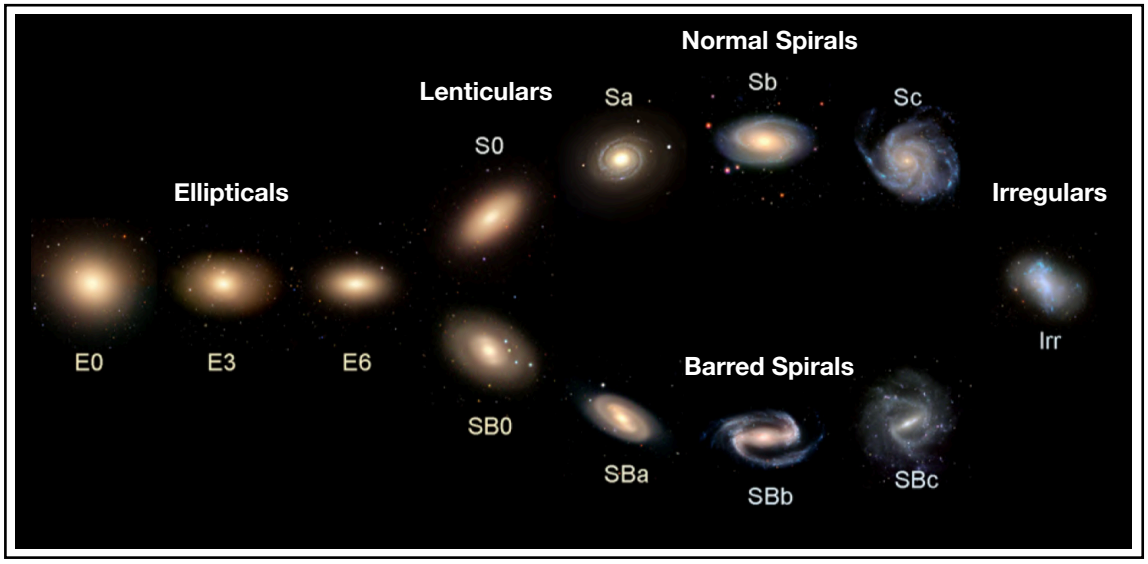


Figure 2.1: *Top*: Schematic representation of the original Hubble's tuning-fork diagram. The elliptical galaxies (early-type galaxies) are on the left, while spirals (late-type galaxies) and irregulars are on the right. (Figure from Yin Cui et al. 2014)

Here we briefly describe the four main classes in which the Hubble sequence is divided (Mo et al. 2010; Longair 2008):

- **Elliptical galaxies (E):** These galaxies have smooth and almost elliptical isophotes[1] and they lack of internal structural features in their brightness distribution. Ellipticals are usually divided in eight sub-classes from E0 to E7. The number corresponds to the closest integer to the value of $10 \cdot (1 - \frac{b}{a})$, where a and b are the lengths of the semi-major and semi-minor axis of the elliptical isophote respectively.

- **Spiral galaxies (S-SB):** They have a disk like appearance with well-defined spiral arms departing from their centers up to the external regions. In Hubble's tuning-fork diagram, spirals are divided into two branches: the normals and the barred spirals.

- **Lenticular galaxies (S0-SB0):** This class represents an intermediate stage between the ellipticals and the spirals. Thus, as ellipticals, also lenticular galaxies have a smooth light distribution without spiral arms or [HII] regions, however, as the spirals, they have a thin disk and a bulge component.

After the original Hubble's classification, a variety of other morphological schemes were introduced. In particular, in 1974, de Vaucouleurs, Sandage, Kormendy and van den Bergh extended the Hubble's original diagram to accommodate the dwarf galaxies discovered in those years. This extended morphological classification is now known as the "revised Hubble sequence of galaxies".

In addition to the morphological classification, galaxies can also be organized as a function of their spectral characteristics (see Fig. 2.2). In particular, arranging galaxies in a color-magnitude diagram we can identify two main classes: a "red sequence" composed by non-star-forming, or passive early-type galaxies with small fractions of gas and dust; and a "blue cloud" made by star-forming, gas and dust rich, late-type galaxies. An intermediate class, defined only in terms of its spectral properties, contains the so called E+A galaxies. The spectra of E+A galaxies are characterized by strong Balmer absorption lines, symptom of a significant population of A-stars (E+A means ellipticals with A-stars). The presence of these young stars proves that these galaxies underwent episodes of star formation within the last Gyr or so (i.e. the main-sequence lifetime of an A-star). However, the absence of [OII]

[1] Isophotes are surfaces of equal surface brightness

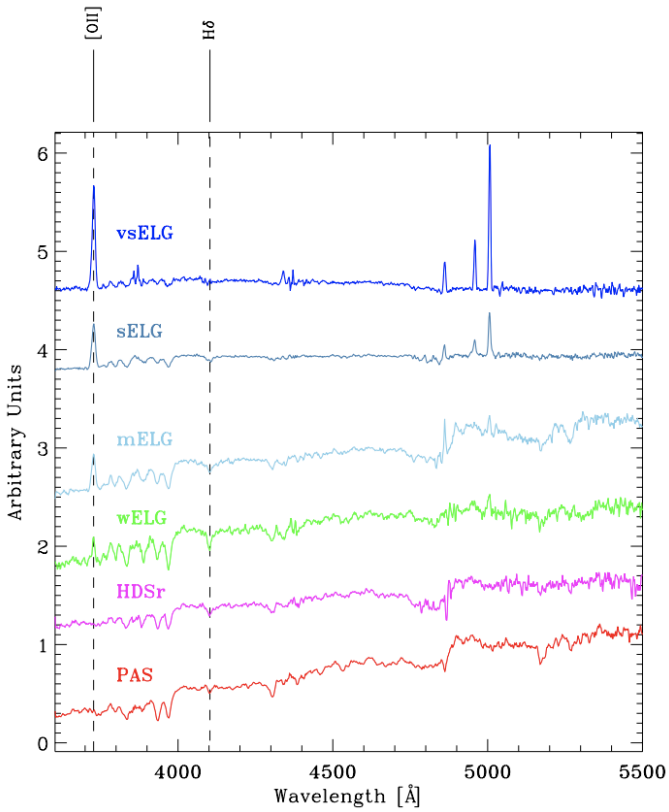


Figure 2.2: Co-added (rest-frame) spectra of cluster galaxies, belonging to the cluster MACS J1206.20847 (z=0.44), divided in six spectral classes according to the presence and strength of their Hδ absorption and [OII] emission lines. (Figure from Girardi et al. 2015)

and Hα emission lines in their spectra demonstrates that the E+A galaxies are not undergoing star formation at the epoch of the observations. They often shows the presence of a disk besides a bulge, or disturbed morphology. Thus, E+A galaxies can be interpreted as younger elliptical galaxies or post-starburst galaxies, the result for example of a relatively recent interaction of two or more nearby galaxies.

2.2 Galaxy populations in clusters

The fraction of galaxies per morphological type strongly correlates with the number density of galaxies of the surrounding environment. Thus, most of field galaxies (galaxies not belonging to groups or clusters) are spirals, while galaxies residing in dense environments, such as rich-groups or galaxy clusters, are mostly ellipticals (see Fig. 2.3 and Dressler 1980). In particular the percentage of E and S0 galaxies in clusters is about the 80-90%, while the fraction of early-type galaxies in the general

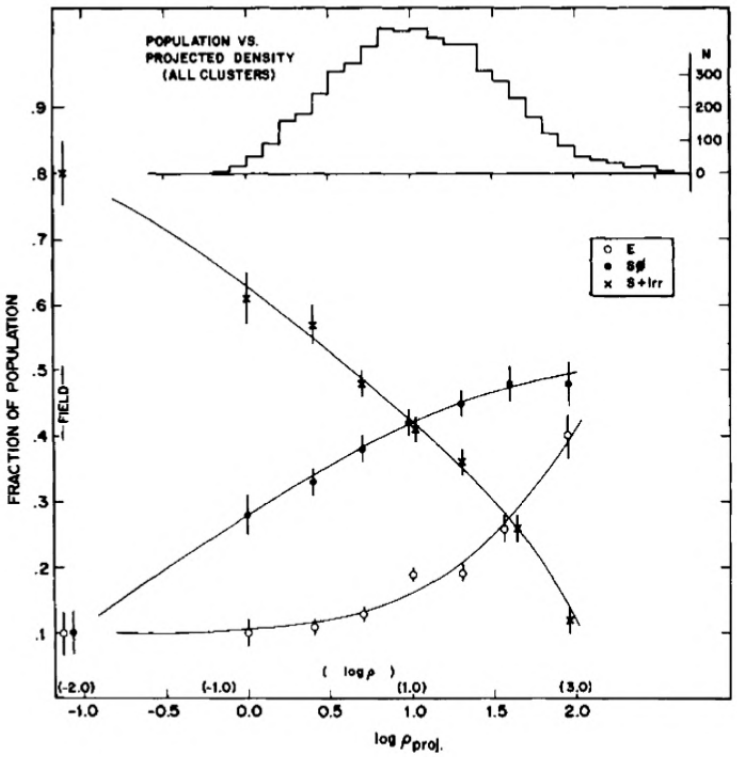


Figure 2.3: The fraction of E, S0, and S+Irr galaxies as a function of the log of the projected density, in galaxies Mpc^{-2}. The data shown are for all cluster galaxies in the sample studied by Dressler (1980) and for the field. Also shown is an estimated scale of true space density in galaxies Mpc^{-3} . The upper histogram shows the number distribution of the galaxies over the bins of projected density. (Figure from Dressler 1980)

field is only about the 20-10% (Dressler 1984; Renzini 2006). In addition to ordinary galaxies (spirals and ellipticals) clusters contain also a non negligible fraction of E+A galaxies.

At the very centre of almost all galaxy clusters sits the Brightest Cluster Galaxy (BCG) characterized by an extraordinarily diffuse and extended outer envelope. This galaxy takes the name of cD galaxy, where the D states for diffuse. BGCs are the most massive galaxies in the Universe (their typical total mass is of about $10^{12}\,\mathrm{M}_{\odot}$) and their luminosity accounts for about the 30% of the total visible light emitted by a rich cluster. These galaxies are probably the result of mergers of smaller cluster members as confirm by the presence of multiple substructures in the light distribution of the 25−50% of the BCGs (Dressler 1984).

The galaxy populations of clusters at intermediate redshifts ($z \sim 0.5$, those considered in this work) exhibit a clear trend of spectral classes, ranging from passive

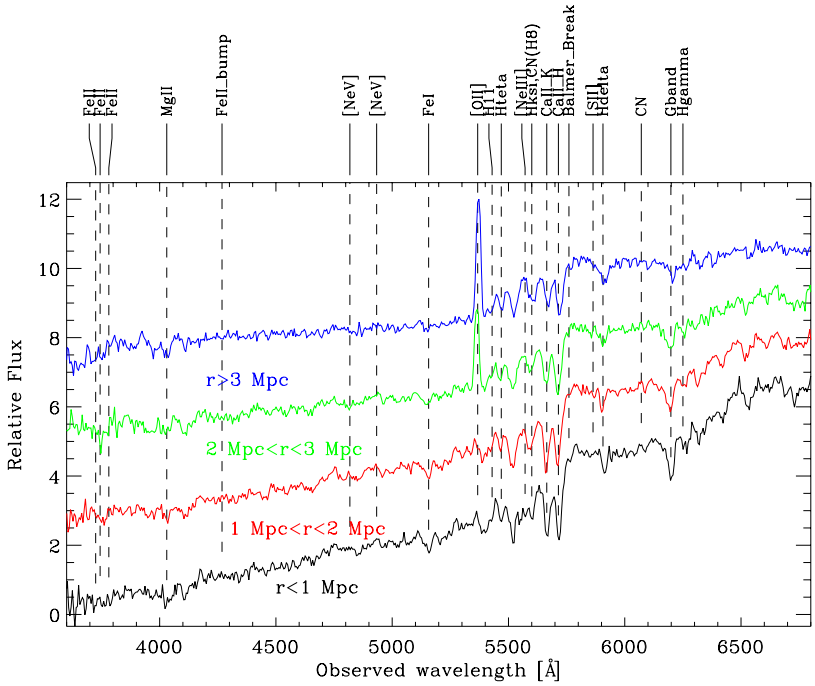


Figure 2.4: Co-added spectra of cluster galaxies in different radial bins of the cluster MACS J1206.2-0847 (z=0.44, ∼70 spectra per bin), showing the typical stellar population gradient in clusters, from high densities in the core to low densities in the outskirts. (Figure from Rosati et al. 2014)

galaxies in the cores to star-forming in the outskirts (see Fig. 2.4). This is the result of the progressive quenching of star formation in cluster cores, in the late stages of their evolution, as a result of different environmental effects (e.g. Treu et al. 2003).

Since our work is focused on the study of the cluster member galaxies, which are mostly ellipticals, in the following sections we discuss the main physical properties characterizing elliptical galaxies.

2.3 Surface brightness and luminosity of elliptical galaxies

Due to the axial symmetry of the elliptical galaxies we can define a one dimensional surface brightness profile $I(R)$ as a function of the isophotal semi-major axis length

R. This relation, found by Sérsic in 1968, is known with the name of $R^{1/n}$ profile (Mo et al. 2010; Longair 2008):

$$I(R) = I_0 \exp\left[-\beta_n \left(\frac{R}{R_e}\right)^{1/n}\right] = I_e \exp\left\{-\beta_n \left[\left(\frac{R}{R_e}\right)^{1/n} - 1\right]\right\}, \tag{2.1}$$

where I_0 is the central surface brightness of the galaxy, R_e is its effective radius[2], I_e is the surface brightness within R_e, n is the so called Sérsic's index that quantify the concentration of the profile and β_n is a constant well approximated by $\beta_n = 2n - 0.324$ for $n \geq 1$. β_n ensures that the integral of the Eq. 2.1 over the full range of R is equal to the total luminosity emitted by the galaxy.

Early surface brightness measurements found that normal elliptical galaxies have average values for n and β_n of about 4 and 7.67 respectively. The Sérsic's profile with $n = 4$ and $\beta_n = 7.67$ is known as de Vaucouleurs's profile, or $R^{(1/4)}$ law (de Vaucouleurs, 1948). More recent and accurate surface brightness measurements demonstrate that the Sérsic's index varies with the luminosity and the size of the elliptical galaxy. In particular, dwarf ellipticals have $n \sim 0.5$, while giant ellipticals have $n > 10$.
With an opportune choice for the Sérsic's index, the Eq. 2.1 can be used also to describe the surface brightness of spiral galaxies. For example, spiral disks are well approximated by a $n = 1$ Sérsic's profile, i.e. an exponential law.

The surface brightness profile in Eq. 2.1 can also be expressed in unit of $mag\, arcsec^{-2}$ (Mo et al. 2010; Schneider 2006):

$$\mu(R) = \mu_e + 1.086\, \beta_n \left[\left(\frac{R}{R_e}\right)^{1/n} - 1\right], \tag{2.2}$$

with $\mu \propto -2.5 \log(I)$.

Integrating the Eq. 2.1 we can immediately derive the total luminosity, L, emitted by the elliptical galaxy. In the simple case of a spherical system this integration reduces to:

$$L = 2\pi \int_0^\infty I(R) R dR = \frac{2\pi n \Gamma(2n)}{(\beta_n)^{2n}} I_0 R_e^2, \tag{2.3}$$

[2]radius enclosing half of the total light emitted by a source, in this case a galaxy

where $\Gamma(x) = \int_0^\infty k^{x-1}e^{-k}dk$.

Considering an isotropic light emission, the flux (F) received from a galaxy at a luminosity distance d_L is: $F = \frac{L}{4\pi d_L^2}$. Furthermore the solid angle subtended by the galaxy is $\Omega = \frac{\pi R_{gal}^2}{d_a^2}$, with d_a the angular diameter distance of the source and R_{gal} the galaxy radius. Thus, if we can neglect the cosmological difference between d_L and d_a, the mean surface brightness measured for the galaxy is simply:

$$\langle I \rangle = \frac{F}{\Omega} = \frac{L}{4\pi^2 R_{gal}^2}. \tag{2.4}$$

Since the effective radius encircle half of the total light emitted by the galaxy, the mean surface brightness within R_e is finally given by:

$$\langle I_e \rangle = \frac{L}{8\pi^2 R_e^2}. \tag{2.5}$$

In general the sizes of the brightest elliptical galaxies strongly increase with their luminosity. Thus, from the Eq. 2.5, we expect a decrease of their average surface brightness at higher luminosity. Conversely, the faint ellipticals have an almost constant effective radius and their average surface brightness increases with their luminosity. Recent observations found an increase of the average surface brightness with the luminosity for galaxies with $M_B \leq -20.5$, beyond this limit the trend is reversed (Mo et al. 2010).

2.4 Dynamics of elliptical galaxies

The stellar dynamics of the elliptical galaxies play a fundamental role in the development of this thesis. In this section we summarize the main concepts and equations governing the motion of the stars inside the ellipticals.

2.4.1 Elliptical galaxies as relaxed collisionless systems

Galaxies are gravitationally bound systems of stars, gas, dust and dark matter. If we consider an elliptical galaxy as an ensemble of particles moving in a halo of size r_h and with mean velocity dispersion $\langle \sigma_h \rangle$, we can compute the average time that these particles take to cross the whole galaxy, t_{cross}, as:

$$t_{cross} = r_h / \langle \sigma_h \rangle \,. \tag{2.6}$$

Since the average radius of a normal elliptical galaxy is of the order of $r_h = 10\,\mathrm{Kpc}$, and considering a mean stellar velocity dispersion of $\langle \sigma_h \rangle = 200\ \mathrm{km\,s^{-1}}$ we obtain: $t_{cross} = 10^8$ years. Comparing t_{cross} with the typical age of a galaxy, that is about 10^{10} years, we immediately deduce that ellipticals can be treated as relaxed systems in dynamic equilibrium under gravity (Schneider 2006; Mo et al. 2010).

Using the crossing time we can also demonstrate that elliptical galaxies are perfect examples of collisionless multi-particle systems. Indeed, a system is considered collisionless if the characteristic time between two particle collisions is much larger than t_{cross}.
Two gravitationally interacting bodies can experience two kinds of collisions. Direct collisions, such as two stars that hit each other, or indirect collisions, when the particle trajectories are curved by the gravitational interaction.
If the elliptical galaxy contains N particles of mean radius r_p, the cross-section for a direct collision is πr_p^2. Thus, the mean free path between two encounters is $\lambda = 1/(n\pi r_p^2)$, where $n = 3N/(4\pi r_h^3)$ is the particle number density. Dividing λ for the mean particle velocity we can derive an estimation of the typical time between two direct collisions, t_{direct} (Schneider 2006):

$$t_{direct} = \frac{\lambda}{\langle \sigma_h \rangle} \backsimeq \left(\frac{r_h}{r_p}\right)^2 \frac{t_{cross}}{N}.$$

Since a typical galaxy contains $N \sim 10^{10}$ stars and the radius of a star like the sun is about $6.9 \cdot 10^5\,\mathrm{km}$, we obtain a collision time of the order of 10^{21} years. This value is eleven order of magnitudes larger than the age of the Universe, therefore the direct collisions are completely negligible.

We consider now the indirect collisions. If two bodies with velocity v and mass m, undergo to a close encounter with impact parameter b[3], the gravitational interaction increases their velocity components perpendicular to the incoming directions ($\delta v_\perp$) by:

$$|\delta v_\perp| \sim a_\perp \cdot \Delta t \sim \left(\frac{Gm}{b^2}\right)\left(\frac{2b}{v}\right) = 2\frac{Gm}{bv}. \tag{2.7}$$

$a_\perp$ is the gravitational acceleration at the closest separation, while Δt is a time scale

[3]the impact parameter corresponds to the minimum distance between two approaching bodies

for the duration of the close encounter. Due to the randomic nature of the indirect collisions, the cumulative $\delta v_\perp$ has an expectation value equal to zero: $\sum_i \left\langle v_\perp^{(i)} \right\rangle = 0$. However, we have a not null mean square velocity given by (Mo et al. 2010):

$$\left\langle |\Delta v_\perp|^2(t) \right\rangle = \sum_{ij} \left\langle \delta v_\perp^{(i)} \cdot \delta v_\perp^{(j)} \right\rangle = \sum_i \left\langle |\delta v_\perp^{(i)}|^2 \right\rangle \neq 0, \tag{2.8}$$

where $\left\langle \delta v_\perp^{(i)} \cdot \delta v_\perp^{(j)} \right\rangle = 0$ for $i \neq j$ because the directions of the indirect collisions are uncorrelated. The number of close encounters that a particle experienced in a time t and with impact parameters between b and $b + db$ is: $N_e(t)\, db = 2\pi v\, n\, t\, b\, db$. Thus, we can substitute the summation in the Eq. 2.8 with the following integral over the impact parameter:

$$\left\langle |\Delta v_\perp|^2(t) \right\rangle = \int_{b_{min}}^{b_{max}} 2\pi b\, v\, n\, t \left(2\frac{Gm}{bv} \right)^2 db. \tag{2.9}$$

The gravitational system has a size r_h, thus a natural choice for the largest impact parameter is $b_{max} = r_h$. Furthermore, the Eq. 2.7 become meaningless if $\delta v_\perp$ is of the order of the particle velocities v, so we can define a minimum impact parameter considering $|\delta v_\perp| = v$, i.e. $b_{min} = 2\frac{Gm}{v^2}$. b_{min} can also be expressed in terms of r_h as $b_{min} = 2\frac{m}{M} r_h$, where we used $GM/r_h = v^2$ (see the virial theorem in Sub-section 2.4.5). Since the total mass of the system is $M = Nm$, we have $b_{max}/b_{min} \sim N$ and the Eq. 2.9 reduces to:

$$\left\langle |\Delta v_\perp|^2(t) \right\rangle = 2\pi \left(\frac{2Gm}{v} \right)^2 v\, t\, n \ln N. \tag{2.10}$$

Defining the time between two indirect collisions, t_{relax}, as the time necessary to have a perpendicular velocity equal to the particle velocity ($\left\langle |\Delta v_\perp|^2(t_{relax}) \right\rangle = v^2$), we have:

$$t_{relax} = \frac{1}{2\pi n v} \left(\frac{v^2}{2Gm} \right)^2 \frac{1}{\ln N} = \frac{1}{2\pi n v} \left(\frac{M}{2Rm} \right)^2 \frac{1}{\ln N} \sim \frac{R}{v} \frac{N}{\ln N}.$$

Assuming v equal to the velocity dispersion of the particles in the halo ,σ_h, and using the Eq. 2.6:

$$t_{relax} = t_{cross} \frac{N}{\ln N}.$$

Since in normal galaxies $N \sim 10^{10}$, we finally derive: $t_{relax} \gg t_{cross}$. This proves that ellipticals are collisionless systems with (Mo et al. 2010):

$$t_{cross} \ll t_{univ} \ll t_{relax} \ll t_{direct}. \tag{2.11}$$

The Eq. 2.11 especially holds in the inner regions of galaxies while in the outer regions (dark matter halos) $t_{cross} \sim t_{univ}$.

In previous equations we assumed a constant total number of galaxy particles over time. This is justified considering that the fractional change in the number of stars per crossing time is very small.

2.4.2 Jeans equations

A system composed by a large number of particles confined inside a smooth potential, $\Phi(\mathbf{x}, t)$, is completely described by the phase-space distribution function $f(\mathbf{x}, \mathbf{v}, t)$. This function gives the number density of particles in the phase-space $(\mathbf{x}, \mathbf{v})$ at time t. If the direct and indirect collisions between particles are negligible (see previous Sub-section), the $f(\mathbf{x}, \mathbf{v}, t)$ obeys to the following collisionless Boltzmann equation (Mo et al. 2010):

$$\frac{\partial f}{\partial t} + \sum_i v_i \frac{\partial f}{\partial x_i} - \sum_i \frac{\partial \Phi}{\partial x_i} \frac{\partial f}{\partial v_i} = 0, \tag{2.12}$$

where $i = 1, 2, 3$ denotes the three spatial directions.

Integrating this equation over the 3D velocity $\mathbf{v}$ we obtain:

$$\frac{\partial}{\partial t} \int f d^3\mathbf{v} + \sum_i \frac{\partial}{\partial x_i} \int v_i f d^3\mathbf{v} - \sum_i \frac{\partial \Phi}{\partial x_i} \int \frac{\partial f}{\partial v_i} d^3\mathbf{v} = 0, \tag{2.13}$$

where we consider a constant range of integration and v_i independent from x_i. Applying the divergence theorem and assuming a vanishing $f(\mathbf{x}, \mathbf{v}, t)$ for sufficiently large velocities, we can demonstrate that the last term in the Eq. 2.13 is equal to zero.

Since the particle spatial density is given by:

$$\nu(\mathbf{x}, t) = \int f(\mathbf{x}, \mathbf{v}, t) \mathrm{d}^3\mathbf{v}, \tag{2.14}$$

and the mean particle velocity is:

$$\langle v_i \rangle = \frac{1}{\nu} \int v_i f(\mathbf{x}, \mathbf{v}, t) \mathrm{d}^3\mathbf{v}, \tag{2.15}$$

we can substitute Eqs. 2.14 and 2.15 into Eq. 2.13 to derive the following continuity equation:

$$\frac{\partial \nu}{\partial t} + \sum_i \frac{\partial}{\partial x_i} [\nu \langle v_i \rangle] = 0. \tag{2.16}$$

Similarly, multiplying the Eq. 2.13 for v_j and integrating over $\mathbf{v}$ we have:

$$\frac{\partial}{\partial t} [\nu \langle v_j \rangle] + \sum_i \frac{\partial}{\partial x_i} [\nu \langle v_i v_j \rangle] + \nu \frac{\partial \Phi}{\partial x_j} = 0, \quad \text{for } i, j = 1, 2, 3, \tag{2.17}$$

where the term $\langle v_i v_j \rangle = \frac{1}{\nu} \int v_i v_j f(\mathbf{x}, \mathbf{v}, t) \mathrm{d}^3\mathbf{v}$ can be expressed as the sum of the two following components:

$$\langle v_j v_k \rangle = \langle v_j \rangle \langle v_k \rangle + \sigma^2_{jk}. \tag{2.18}$$

The first component accounts for the coherent streaming motion of the particles at a given position, while the second is their velocity dispersion.

Subtracting from Eq. 2.17 the product of the continuity equation (Eq. 2.16) with $\langle v_j \rangle$, we finally derive the Jeans' equation for our particle fluid (Mo et al. 2010):

$$\frac{\partial \langle v_j \rangle}{\partial t} + \sum_i \langle v_i \rangle \frac{\partial \langle v_j \rangle}{\partial x_i} = -\frac{1}{\nu} \sum_i \frac{\partial \left(n \sigma^2_{ij} \right)}{\partial x_i} - \frac{\partial \Phi}{\partial x_j}, \quad \text{for } i, j = 1, 2, 3. \tag{2.19}$$

The local quantity $n\sigma^2_{ij}$, called stress tensor, quantifies the particle random movements around their local mean motion. This term is equivalent to an an-isotropic pressure for the particle fluid. Since the stress tensor is symmetric, we can simplify the jeans' equations using a reference frame in which this tensor is diagonal. The ellipsoid whose principal axes have the directions of the orthogonal eigenvectors of σ^2_{ij} and lengths proportional to its eigenvalues is called *velocity ellipsoid*.

The Jeans' equations form a system of three equations in nine unknown variables. Three are the components of the streaming motion $\langle v_i \rangle$, and the other six are the

independent components of the rank two symmetric stress tensor. Since the number of equations is lower than the number of unknowns we need additional assumptions on the form of the stress tensor to solve the Jeans' equations.

The Jeans' equations assume an especially useful form when they are expressed in the spherical coordinates r, θ and ϕ. In particular the radial jeans' equation is given by:

$$\nu\frac{\partial\langle v_r\rangle}{\partial t}+\nu\left(\langle v_r\rangle\frac{\partial\langle v_r\rangle}{\partial r}+\frac{\langle v_\theta\rangle}{r}\frac{\partial\langle v_r\rangle}{\partial\theta}+\frac{\langle v_\phi\rangle}{r\sin\theta}\frac{\partial\langle v_r\rangle}{\partial\phi}\right)+\frac{\partial}{\partial r}\left(\nu\sigma_{rr}^2\right)+\frac{1}{r}\frac{\partial}{\partial\theta}\left(\nu\sigma_{r\theta}^2\right)$$
$$+\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\left(\nu\sigma_{r\phi}^2\right)+\frac{\nu}{r}\left[2\sigma_{rr}^2-\left(\sigma_{\theta\theta}^2+\sigma_{\phi\phi}^2+\langle v_\theta\rangle^2+\langle v_\phi\rangle^2\right)+\sigma_{r\theta}^2\cot\theta\right]=-\nu\frac{\partial\Phi}{\partial r}, \tag{2.20}$$

where $v_r=\frac{dr}{dt}$, $v_\theta=r\frac{d\theta}{dt}$, $v_\phi=r\sin\theta\frac{d\phi}{dt}$.

In the case of a steady-state system in hydrodynamic equilibrium, we have: $\frac{\partial\langle v_r\rangle}{\partial t}=0$ and $\langle v_r\rangle=0$ (negligible radial flows). Moreover, if we consider a spherically symmetric system: $\langle v_\theta\rangle=\langle v_\phi\rangle=0$, $\sigma_{r\theta}^2=\sigma_{r\phi}^2=\sigma_{\theta\phi}^2=0$, $\langle v_\theta^2\rangle=\langle v_\phi^2\rangle$, $\sigma_{rr}^2=\langle v_r^2\rangle$, $\sigma_{\theta\theta}^2=\langle v_\theta^2\rangle$ and $\sigma_{\phi\phi}^2=\langle v_\phi^2\rangle$ (see Eq. 2.18).
Under these conditions the Eq. 2.20 reduces to the simpler form:

$$\frac{d(\nu\langle v_r^2\rangle)}{dr}+2\frac{\nu}{r}\left[\langle v_r^2\rangle-\langle v_\theta^2\rangle\right]=-\nu\frac{d\Phi}{dr} \tag{2.21}$$

2.4.3 Projection of physical quantities: The Abel's transform

Most of the relevant quantities in astrophysics (e.g. the surface brightness, the mass density distributions and the velocity dispersion profiles) are observed in projection on a plane perpendicular to the line-of-sight (hereafter LoS). A useful mathematical technique to achieve this projection from the 3D to the 2D space is the Abel's transform.
Suppose we have a spherically symmetric function $f(r)$, with $r=\sqrt{x^2+y^2+z^2}$, describing the 3D variation of a physical quantity of a system. If we are observing this system with a LoS along the x direction, the projection of $f(r)$ on the yz-plane, $F(R)$, is given by the Abel's transform:

$$F(R)=\int_{-\infty}^{+\infty}f(r)dx=2\int_0^{\infty}f(r)dx=2\int_R^{\infty}\frac{f(r)rdr}{\sqrt{r^2-R^2}}, \tag{2.22}$$

In this equation we consider the parity of $f(r)$ and we define $R = \sqrt{y^2 + z^2}$ (this formula can be easily generalized to the more general case of an axially symmetric system).
If $f(r)$ goes to zero faster than $1/r$, we can also define the inverse Abel's transform as:

$$f(r) = -\frac{1}{\pi} \int_r^\infty \frac{dF(R)}{dR} \frac{dR}{\sqrt{R^2 - r^2}}. \tag{2.23}$$

The Eq. 2.23 permits an analytical deprojection of the physical quantities from the 2D to the 3D space. Thus, we can derive for example the three-dimensional space mass density of a galaxy, $\rho(r)$, from its projected mass density, $\Sigma(R)$, or the three-dimensional luminosity density, $\nu(r)$, from the observed two dimensional surface brightness, $I(R)$. Unfortunately, a direct application of the Abel's equations on the real observations is extremely difficult due to the presence of the noise. To overcome this problem we can firstly fit the observed projected profiles with analytic functions and then apply the inverse Abel's transform only to these functions. Another possibility consists in projecting, through the Eq. 2.22, 3D functional forms with free parameters. The free parameter values are then optimized comparing the projected models with the observed profiles.
Note that the inverse Abel's transform can only be used on axially symmetric systems. For all the other systems, such as the triaxial systems, there is not a unique solution to the inversion problem and we need ad-hoc assumptions on the 2D (or 3D) distributions to obtain a result.

2.4.4 Galaxy velocity dispersions and comparison with the observed quantities

In Section 2.4.2 we derived the radial jeans' equation for a spherically symmetric system in a condition of steady-state hydrodynamic equilibrium (see Eq. 2.20). Now, we compute from this equation an expression for the aperture-average LoS stellar velocity dispersion of an elliptical galaxy. This useful quantity is directly measurable from the spectroscopic observations (see Section 5.2).

If the system that we are considering is a galaxy made of stars with similar luminosity, the particle spatial density (ν) in Eq. 2.20 can be identified, up to a constant, with the stellar luminosity density. Similarly, the particle velocities correspond to stellar velocities, while Φ is the gravitational potential of the galaxy.

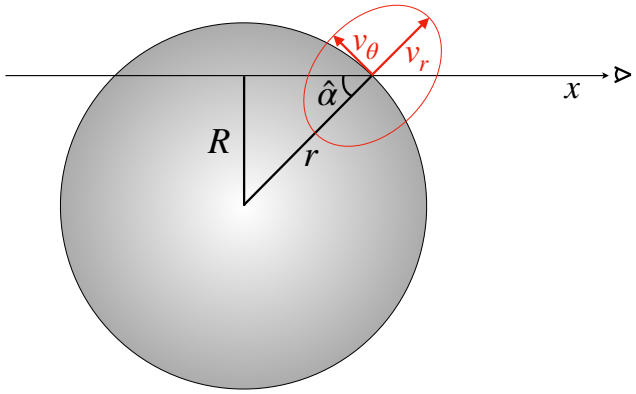


Figure 2.5: Schematic representation of a spherically symmetric elliptical galaxy. The velocity components are in red. The line-of-sight is along the x axis.

Defining the pressure anisotropy parameter (An & Evans 2011):

$$\beta(r) = 1 - \frac{\langle v_\theta^2 \rangle}{\langle v_r^2 \rangle}, \tag{2.24}$$

we can rewrite the Eq. 2.20 as:

$$\frac{d\left(\nu \langle v_r^2 \rangle\right)}{dr} + \frac{2\beta(r)\nu \langle v_r^2 \rangle}{r} = -\nu \frac{d\Phi}{dr}. \tag{2.25}$$

Integrating this equation over r, we find an equation for the radial velocity dispersion $\langle v_r^2 \rangle = \sigma_{rr}^2$ (Agnello et al. 2014, van der Marel 1994):

$$\langle v_r^2 \rangle = \frac{G}{\nu(r)} \int_r^\infty \frac{m(r')\nu(r')}{r'^2} J_\beta(r, r') dr', \tag{2.26}$$

where we substituted the expression for the gravitational potential, $\frac{d\Phi}{dr} = \frac{Gm(r)}{r}$, and we defined $J_\beta(r, r') = \exp\left[\int_r^{r'} \frac{2\beta(s)}{s} ds\right]$.
From the Eq. 2.25 we can also derive an expression for the circular velocity $v_c(r)$ (Binney & Mamon 1982):

$$v_c^2 = \frac{Gm(r)}{r} = -\langle v_r^2 \rangle \left(\frac{d \ln \nu}{d \ln r} + \frac{d \ln \langle v_r^2 \rangle}{d \ln r} + 2\beta \right). \tag{2.27}$$

The LoS component of the stellar velocity dispersion at a distance r from the galactic centre is: $\left\langle (v_r \cos\alpha + v_\theta \sin\alpha)^2 \right\rangle = \langle v_r^2 \rangle \cos^2\alpha + \langle v_\theta^2 \rangle \sin^2\alpha$ (see Fig. 2.5), and the LoS stellar velocity dispersion profile is simply the luminosity weighted mean of this value along the LoS:

$$\sigma_p^2(R) = \frac{1}{I(R)} \int_{-\infty}^{+\infty} \left(\left\langle v_r^2 \right\rangle \cos^2 \alpha + \left\langle v_\theta^2 \right\rangle \sin^2 \alpha \right) \nu dx = \\ = \frac{2}{I(R)} \int_R^\infty \left(1 - \beta \frac{R^2}{r^2} \right) \frac{\nu \left\langle v_r^2 \right\rangle r}{\sqrt{r^2 - R^2}} dr. \quad (2.28)$$

To compute the last term we use the Abel' transform in Eq. 2.22, while the surface brightness $I(R)$ is defined as the 2D projection of the luminosity density:

$$I(R) = 2 \int_R^\infty \frac{\nu(r) r}{\sqrt{r^2 - R^2}} dR. \quad (2.29)$$

Replacing in Eq. 2.28 the expression for the radial velocity dispersion $\langle v_r^2 \rangle$ (Eq. 2.26), exchanging the order of integration and integrating by parts, we obtain the following final equation for the LoS stellar velocity dispersion profile (Agnello et al. 2014):

$$\sigma_p^2(R) = \frac{2G}{I(R)} \int_R^\infty \frac{\nu(r) m(r)}{r^2} \left(\sqrt{r^2 - R^2} + k_\beta(R, r) \right) dr, \quad (2.30)$$

where:

$$k_\beta(R, x) = \int_R^x \frac{(2r^2 - 3R^2)\, \beta(r) J_\beta(r, x)}{r \sqrt{r^2 - R^2}} dr. \quad (2.31)$$

In reality the LoS stellar velocity dispersion is not a directly observable quantity. Indeed, the galaxy spectra from which we measure the stellar velocity dispersions (see Section 5.2) are extracted within apertures. Thus, what we obtain from the observations is the luminosity average of σ_p inside an aperture of radius R.
In particular, if the spectra are extracted within slits, R coincides with the slit widths; if we are using a fiber spectrograph, R is at last equal to the fiber size; while for an integral field spectrograph, it is the radius of the aperture chosen for the spectral extraction (and the smaller R coincides with the pixel size).
The mathematical expression for the luminosity-average LoS stellar velocity dispersion inside an aperture of radius R (Agnello et al. 2014) is given by:

$$\sigma_{ap}^2(R) = \frac{2\pi}{L(R)} \int_0^R R' I(R') \sigma_p^2(R') dR' \quad (2.32)$$

,

where we define the total 2D luminosity inside the aperture as:

$$L(R) = 2\pi \int_0^R R' I(R') dR'. \tag{2.33}$$

Obviously, we have to consider that in real observations all these quantities are blurred by the point-spread function (PSF) of the instruments. For the earth base spectrographs the PSF value is strongly dominated by the atmosphere seeing conditions while the impact of the diffraction effects, due to the finite telescope apertures, are totally negligible.

2.4.5 The virial theorem

The virial theorem is one of the most fundamental theorems in astrophysics. We can derive its tensor formulation from the Eq. 2.17 as describe in Mo et al. (2010):

$$\frac{1}{2}\frac{\mathrm{d}^2 I_{jk}}{\mathrm{d}t^2} = 2K_{jk} + W_{jk} + \Sigma_{jk}. \tag{2.34}$$

$I_{jk} = \int \rho x_j x_k \mathrm{d}^3\mathbf{x}$ is the moment of inertia tensor, $K_{jk} = \frac{1}{2}\int \rho \langle v_j v_k \rangle \mathrm{d}^3\mathbf{x}$ is the kinetic energy tensor, $W_{jk} = -\int \rho x_k \frac{\partial \Phi}{\partial x_j} \mathrm{d}^3\mathbf{x}$ is the Chandrasekhar potential energy tensor and $\Sigma_{jk} = -\sum_i \int x_k \rho \langle v_j v_i \rangle \mathrm{d}S_i$ is a surface pressure term.

Computing the trace of the Eq. 2.34, we obtain the scalar formulation of the virial theorem:

$$\frac{1}{2}\frac{\mathrm{d}^2 I}{\mathrm{d}t^2} = 2K + W + \Sigma, \tag{2.35}$$

where:

$$I = \mathrm{Tr}(I_{ij}) = \int \rho\, r^2 \mathrm{d}^3\mathbf{x},$$

corresponds to the total moment of inertia of the system;

$$K = \mathrm{Tr}(K_{ij}) = \frac{1}{2}\int \rho \left\langle v^2 \right\rangle \mathrm{d}^3\mathbf{x},$$

is the total kinetic energy;

$$W = \mathrm{Tr}(W_{ij}) = -\int \rho \mathbf{x} \cdot \nabla\Phi \mathrm{d}^3\mathbf{x},$$

is the total potential energy; and

$$\Sigma = \mathrm{Tr}(\Sigma_{ij}) = -\int \rho \left\langle v^2 \right\rangle \mathbf{x} \cdot \mathrm{d}\mathbf{S},$$

quantifies the work done by external pressure forces on the system. Note that W coincides with the gravitational energy of the system only if all the masses outside the surface S can be ignored in the computation of the potential.

The virial theorem, in its scalar formulation, can be used to obtain a first order approximation of the total mass of an elliptical galaxy. Under the realistic assumption that all the external surface pressure forces exerted by the surrounding Universe on our galaxy are negligible, we have $\Sigma = 0$, i.e. the stability of the system is completely determined by its kinetic and potential energy. In particular, if the kinetic energy of the system is larger than its potential energy the system expands, while if the kinetic energy is smaller the system contracts. Since the kinetic energy in galaxies is almost conserved, we conclude that ellipticals are nearly static systems. In a static system $\frac{\mathrm{d}^2 I}{\mathrm{d}t^2} = 0$, and the virial theorem becomes:

$$E = -K = \frac{W}{2}, \tag{2.36}$$

where $E = K + W$ is total energy and the kinetic energy is:

$$K = \frac{1}{2}\sum m_i \sigma_i^2 = \frac{1}{2} M_{tot} \left\langle \sigma^2 \right\rangle = \frac{3}{2} M_{tot} \left\langle \sigma_{los}^2 \right\rangle .$$

M_{tot} is the total mass of the system and $\langle \sigma^2 \rangle$ a mass weighted velocity dispersion of its particles. In the last step we assumed isotropic particle velocities, i.e.:

$$\left\langle \sigma^2 \right\rangle = 3 \left\langle \sigma_{los}^2 \right\rangle , \tag{2.37}$$

where σ_{los} is the LoS velocity dispersion.

To derive the potential energy term W in Eq. 2.36 we consider the galaxy as spherical. In this case $|W| = Gm^2/r_{gal}$, where $r_{gal} = 2m^2 \left(\sum_{i \neq j} \frac{m_i m_j}{r_{ij}} \right)^{-1}$ is the 3D gravitational radius.

Replacing the expressions for T and W in the Eq. 2.36, we finally obtain an estimation of the total mass of the galaxy:

$$M_{tot} = 3 \frac{\left\langle \sigma_{los}^2 \right\rangle r_{gal}}{G}. \tag{2.38}$$

This equation can be expressed in term of the projected gravitational radius $R_{gal} = 2m^2 \left(\sum_{i \neq j} \frac{m_i m_j}{R_{ij}} \right)^{-1} = \frac{2}{\pi} r_{gal}$:

$$M_{tot} = \frac{3}{2} \frac{\pi \left\langle \sigma_{los}^2 \right\rangle R_{gal}}{G}, \tag{2.39}$$

where both $\langle \sigma^2_{los} \rangle$ and R_{gal} are directly observable physical quantities.

Despite the simplistic assumptions of isotropy and sphericity, the Eq. 2.39 can be used to compute an approximated mass for a galaxy. Moreover, this equation can be easily extended to every kind of gravitationally bound systems, such as a the galaxy clusters.

2.4.6 Elliptical galaxy mass density profiles

In this sub-section we derive three widely used models for the mass density distribution of elliptical galaxies, i.e.: the singular isothermal sphere, the pseudo-isothermal mass distribution and the dual pseudo-isothermal mass distribution.

The Singular Isothermal Sphere (SIS) model relies on the following three assumptions (Schneider 2006):

1. The velocity distribution of massive particles (baryonic and dark matter particles) is locally described by a Maxwell distribution, i.e. particles are in thermal equilibrium.
2. The particle velocity dispersion (temperature of the particle-gas) does not depend on the radius, i.e. we assume an isothermal distribution of particles.
3. The system is in dynamic equilibrium.

These three requests define a set of models, known as isothermal spheres models, characterized by just one free parameter.

A system of N massive particles is in dynamic equilibrium if the attractive gravitational force acting on a mass element $\rho(r)\mathrm{d}V$, at a distance r from the system center, is perfectly balanced by the pressure gradient at the same point, i.e.:

$$\nabla P = \frac{\mathrm{d}P}{\mathrm{d}r} = -\rho(r)\frac{Gm(r)}{r^2}. \tag{2.40}$$

ρ(r) is the mass density profile of the system, while $m(r)$ is the total mass within the radius r:

$$m(r) = \int_0^r 4\pi r'^2 \rho(r')\mathrm{d}r'.$$

Differentiating the Eq. 2.40 respect to r we obtain:

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{r^2}{\rho(r)}\frac{\mathrm{d}P}{\mathrm{d}r}\right) + 4\pi G r^2 \rho(r) = 0. \tag{2.41}$$

This equation takes the name of Lane-Emden equation.

If we define the average particle mass $\langle m \rangle$, we can express the mass density profile as $\rho(r) = \langle m \rangle \, n(r)$, where $n(r)$ is the particle number density.
For a system in thermal equilibrium we have the following relation between the temperature and the mean square velocity of the particles: $\frac{3}{2}k_B T = \frac{\langle m \rangle}{2}\langle v^2 \rangle$. Instead, the pressure can be expressed through the perfect gas law: $P = \rho(r) k_B T / \langle m \rangle = n(r) k_B T$.
The particle-gas is locally described by a Maxwell distribution, thus we have isotropic particles velocities, i.e.: $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \sigma_0^2$, where we introduce the 1D LoS velocity dispersion σ_0. Since $\langle v^2 \rangle = \sigma_x^2 + \sigma_y^2 + \sigma_z^2$, we also have: $\sigma_0^2 = \frac{\langle v^2 \rangle}{3}$.
Replacing all these equations in Eq. 2.40 we can express the pressure gradient as:

$$\frac{\mathrm{d}P}{\mathrm{d}r} = \frac{k_B T}{\langle m \rangle}\frac{\mathrm{d}\rho(r)}{\mathrm{d}r} = \frac{\langle v^2 \rangle}{3}\frac{\mathrm{d}\rho(r)}{\mathrm{d}r} = \sigma_0^2 \frac{\mathrm{d}\rho(r)}{\mathrm{d}r}.$$

While, from the Eq. 2.41 we finally obtain:

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{\sigma_0^2 r^2}{\rho(r)}\frac{\mathrm{d}\rho(r)}{\mathrm{d}r}\right) + 4\pi G r^2 \rho(r) = 0. \tag{2.42}$$

This is a non-linear differential equation that completely characterizes the isothermal sphere models. The analytical solution to this equation is the SIS profile defined by the following mass density relation:

$$\rho_S(r) = \frac{\sigma_0^2}{2\pi G r^2}. \tag{2.43}$$

Thus, the single free parameter of the SIS profile is the 1D isotropic velocity dispersion σ_0.

Despite the SIS profile is widely used in Astrophysics, it suffers for two main issues that make this model unphysical. Firstly, the SIS mass density distribution diverges in the limit of $r \to 0$, and secondly its total mass ($m(r) = \frac{2\sigma_0^2}{G} r$) is unbounded ($m(r) \to \infty$) for $r \to \infty$.

To avoid the singularity of the SIS profile at $r = 0$, Kassiola & Kovner (1993) introduced a variant to the SIS called Pseudo-Isothermal Elliptical Mass Distribution

(PIEMD). This profile, in its circular form (ellipticity $e = 0$), is defined through an extra free parameter known as core radius, r_{core}. In particular, the PIEMD mass density is given by:

$$\rho_P(r) = \frac{\sigma_0^2}{2\pi G}\frac{1}{r^2 + r_c^2}. \tag{2.44}$$

For $r = 0$ the PIEMD profile has a finite central density given by:

$$\rho_{P,0} = \frac{\sigma_0^2}{2\pi G r_c^2}.$$

Computing the derivative of the Eq. 2.44 respect to r, at $r = 0$ we have:

$$\frac{d\rho_P(r)}{dr} = -\frac{\sigma_0^2}{\pi G}\frac{r}{(r^2 + r_c^2)^2}\bigg|_{r=0} = 0.$$

Thus, the PIEMD mass density distribution has a flat core with central density $\rho_{c,0}$. In the limit of $r \gg r_c$ the PIEMD behaves like a pure SIS profile and its total mass is still divergent.

To solve also this second issue, Elíasdóttir et al. (2007) and Limousin et al. (2005) defined the dual Pseudo-Isothermal Elliptical mass distribution (dPIE). Respect to the PIEMD, this mass profile has an extra parameter r_{cut} defining a truncation radius for the mass distribution. Considering a spherically symmetric mass distribution, the dPIE mass density is given by:

$$\rho_d(r) = \frac{\rho_{d,0}}{\left(1 + \frac{r^2}{r_{core}^2}\right)\left(1 + \frac{r^2}{r_{cut}^2}\right)}, \quad \text{with} \quad r_{cut} > r_{core}, \tag{2.45}$$

where

$$\rho_{d,0} = \frac{\sigma_0^2}{2\pi G}\left(\frac{r_{cut} + r_{core}}{r_{core}^2 r_{cut}}\right). \tag{2.46}$$

In the inner region the dPIE has a flat core with central density $\rho_{d,0}$. For $r_{core} < r < r_{cut}$ it behaves like a SIS profile ($\rho_d(r) \propto r^{-2}$). While for $r \gg r_{cut}$ the density drop as $\rho_d(r) \propto r^{-4}$.
In the limit of $r \ll r_{cut}$ the dPIE approximates a PIEMD profile.

2.5 SIS and dPIE: Main physical quantities

In this section we derive the main physical quantities characterizing SIS and dPIE profiles, while PIEMD expressions can be obtained from those of the dPIE in the limit of $r_{cut} \to \infty$.
Despite the dPIE mass distribution is widely used to describe member galaxies in cluster lens models, we found several inconsistencies and errors in the computation of its equations in the literature. Here, following the Appendix C of Bergamini et al. (2019), we show a rigorous derivation of the dPIE most important physical quantities.

Singular Isothermal Sphere

Computing the Abel's transform of the Eq. 2.43 we can derive an expression for the surface mass density of the SIS profile:

$$\Sigma_S(R) = 2\int_R^\infty \frac{\rho_S(r)r}{\sqrt{r^2 - R^2}}dr = \frac{\sigma_0^2}{2GR}. \tag{2.47}$$

Integrating the Eq. 2.43 over a spherical volume of radius r, we obtain an equation for the total SIS mass within this volume, $m(r)$:

$$m_S(r) = 4\pi\int_0^r \rho_S(r')r'^2dr' = \frac{2\sigma_0^2}{G}r. \tag{2.48}$$

Similarly, integrating the surface density $\sigma_S(R)$ over an aperture of radius R, we derive the total mass encircled by the aperture, $M(R)$:

$$M_S(R) = 2\pi\int_0^R \Sigma_S(R')R'dR' = \frac{\pi\sigma_0^2}{G}R. \tag{2.49}$$

In the limit of $r \to \infty$ ($R \to \infty$) the mass $m(r) \to \infty$ ($M(R) \to \infty$), i.e. the SIS has a divergent total mass (see Section 2.4.6).
Finally we compute the expressions for the circular velocity and velocity dispersions of the SIS profile. Using the Eq. 2.27 we obtain a value for the circular velocity, v_c, given by:

$$v_{c,S} = \sqrt{\frac{Gm_S(r)}{r}} = \sqrt{2}\sigma_0. \tag{2.50}$$

To derive the radial velocity dispersion profile ($\sigma_S(r)$), the LoS projected velocity dispersion ($\sigma_p(R)$) and the aperture average velocity dispersion ($\sigma_{ap}(R)$), we exploit the Eqs. 2.26, 2.28 and 2.32 respectively. In particular, if we assume a luminosity density profile identical to the SIS mass density, $\nu(r) = \rho(r)$, we obtain:

$$\sigma_S^2(r) = \frac{G}{\rho_S(r)} \int_r^\infty \frac{m_S(r')\rho_S(r')}{r'^2} dr' = \sigma_0^2, \tag{2.51}$$

$$\sigma_{p,S}^2(R) = \frac{2}{\Sigma_S(R)} \int_R^\infty \frac{\rho_S(r)\sigma_S^2(r)r}{\sqrt{r^2 - R^2}} dr = \sigma_0^2, \tag{2.52}$$

$$\sigma_{ap,S}^2(R) = \frac{2\pi}{M_S(R)} \int_0^R R'\Sigma_S(R')\sigma_{p,S}^2(R')dR' = \sigma_0^2. \tag{2.53}$$

Thus, $\sigma_S(r) = \sigma_p(R) = \sigma_{ap}(R) = \sigma_0$.

Dual Pseudo-Isothermal spherical mass distribution

The same quantities derived for the SIS profile are now computed for the spherical dPIE.
From the Eq. 2.45 and 2.46 we obtain the surface mass density (Limousin et al. 2005):

$$\Sigma_d(R) = 2\int_R^\infty \frac{\rho_d(r)r}{\sqrt{r^2 - R^2}} dr = \frac{\sigma_0^2}{2G} \frac{r_{cut}}{r_{cut} - r_{core}} \left(\frac{1}{\sqrt{r_{core}^2 + R^2}} - \frac{1}{\sqrt{r_{cut}^2 + R^2}} \right), \tag{2.54}$$

while for the 3D and surface mass distributions (Elíasdóttir et al. 2007, Limousin et al. 2005):

$$\begin{aligned} m_d(r) = 4\pi \int_0^r \rho_d(r')r'^2 dr' = \\ = \frac{2\sigma_0^2}{G} \frac{r_{cut}}{r_{cut} - r_{core}} \left[r_{cut} \arctan\left(\frac{r}{r_{cut}}\right) - r_{core} \arctan\left(\frac{r}{r_{core}}\right) \right] \end{aligned} \tag{2.55}$$

$$M_d(R) = 2\pi \int_0^R \Sigma_d(R') R' dR' = \\ = \frac{\pi\sigma_0^2}{G} \frac{r_{cut}}{r_{cut} - r_{core}} \left(\sqrt{r_{core}^2 + R^2} - r_{core} - \sqrt{r_{cut}^2 + R^2} + r_{cut} \right) \quad (2.56)$$

The finite total mass of the dPIE can be computed from the Eq. 2.55 (Eq. 2.56) assuming $r \to \infty$ ($R \to \infty$):

$$M_{tot,d} = \frac{\pi\sigma_0^2 r_{cut}}{G}. \quad (2.57)$$

In the limit of $r_{core}/r_{cut} \to 0$, a sphere of radius $r = r_{cut}$ encloses about half of the total 3D mass of the dPIE. While the 60 (90)% of the total projected mass is encircled by an aperture with radius $R = r_{cut}$ ($R = 5r_{cut}$) (see Fig. 2.6).
From Eqs. 2.56 and 2.57 we can also derive an expression for the half-mass projected radius $R_{M_{tot,p}/2}$, i.e the radius of the aperture that encloses half of the total projected mass of the spherical dPIE (Elíasdóttir et al. 2007):

$$R_{M_{tot,p}/2} = \frac{3}{2} \sqrt{r_{core}^2 + \frac{10}{3} r_{core} r_{cut} + r_{cut}^2} \, . \quad (2.58)$$

As we did for the SIS model, now we compute the kinematic quantities for the spherical dPIE, i.e. the circular velocity ($v_{c,d}(r)$), the velocity dispersion profile ($\sigma_d(r)$), the LoS projected velocity dispersion ($\sigma_{p,d}(r)$) and the aperture average LoS velocity dispersion ($\sigma_{ap,d}(R)$). In particular, the circular velocity is given by:

$$v_{c,d}^2(r) = \sqrt{\frac{G m_d(r)}{r}} = \\ = 2\sigma_0^2 \frac{r_{cut}^2}{r_{cut} - r_{core}} \frac{1}{r} \left[\arctan\left(\frac{r}{r_{cut}}\right) - \frac{r_{core}}{r_{cut}} \arctan\left(\frac{r}{r_{core}}\right) \right]. \quad (2.59)$$

In the limit of $r_{cur} \gg r_{core}$, the maximum of this equation approaches the SIS circular velocity value in Eq. 2.50, i.e. $\sqrt{2}\sigma_0$ (see Fig. 2.7)

To derive the equations for $\sigma_d(r)$, $\sigma_{p,d}(r)$ and $\sigma_{ap,d}(R)$, we assume a luminosity density ν that varies with the same functional form of the 3D dPIE density $\rho_d(r)$. Unlike the SIS profile, none of these quantities has an analytical expression and the integrals in Eqs. 2.26, 2.28 and 2.32 need to be computed numerically.

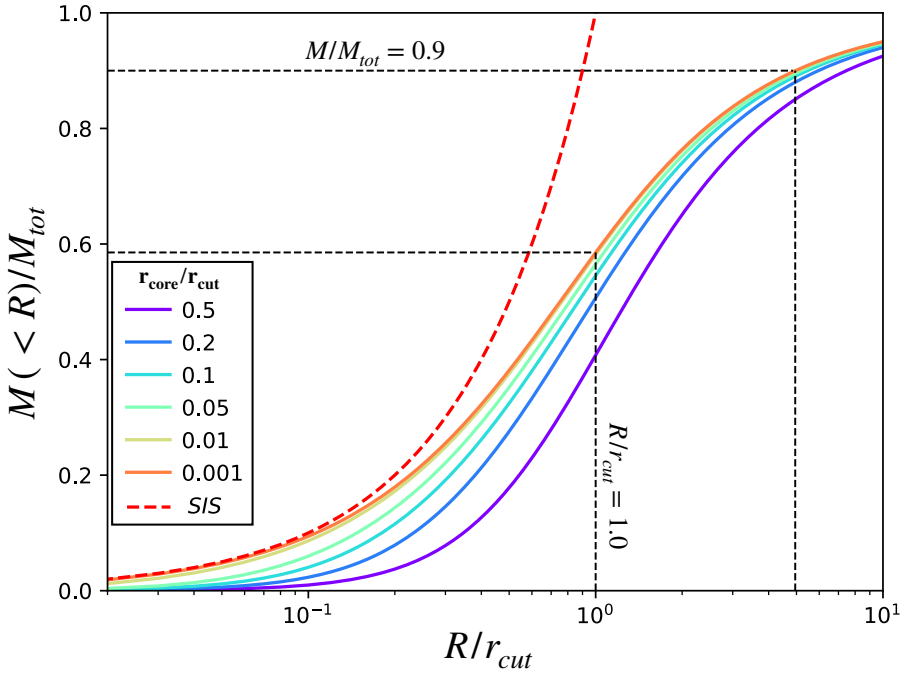


Figure 2.6: Fraction of projected mass over the total mass as a function of aperture radii (in units of r_{cut}) for dPIE profiles with different r_{core}/r_{cut} values. For small r_{core}/r_{cut}, $\sim 60\%$ of the total mass is contained within r_{cut}, while the 90% is contained within $5r_{cut}$. (Figure from Bergamini et al. 2019)

For the radial velocity profile we have:

$$\sigma_d^2(r) = \frac{G}{\rho_d(r)} \int_r^\infty \frac{m_d(r')\rho_d(r')}{r'^2} dr' =$$
$$= \sigma_0^2 \frac{2r_{cut}}{r_{cut} - r_{core}} \left(1 + \frac{r^2}{r_{core}^2}\right) \left(1 + \frac{r^2}{r_{cut}^2}\right) \int_r^\infty \frac{r_{cut} \arctan\left(\frac{r'}{r_{cut}}\right) - r_{core} \arctan\left(\frac{r'}{r_{core}}\right)}{r^2 \left(1 + \frac{r'^2}{r_{core^2}}\right) \left(1 + \frac{r'^2}{r_{cut}^2}\right)} dr'. \tag{2.60}$$

Using the Eq. 2.30, the LoS projected velocity dispersion is given by:

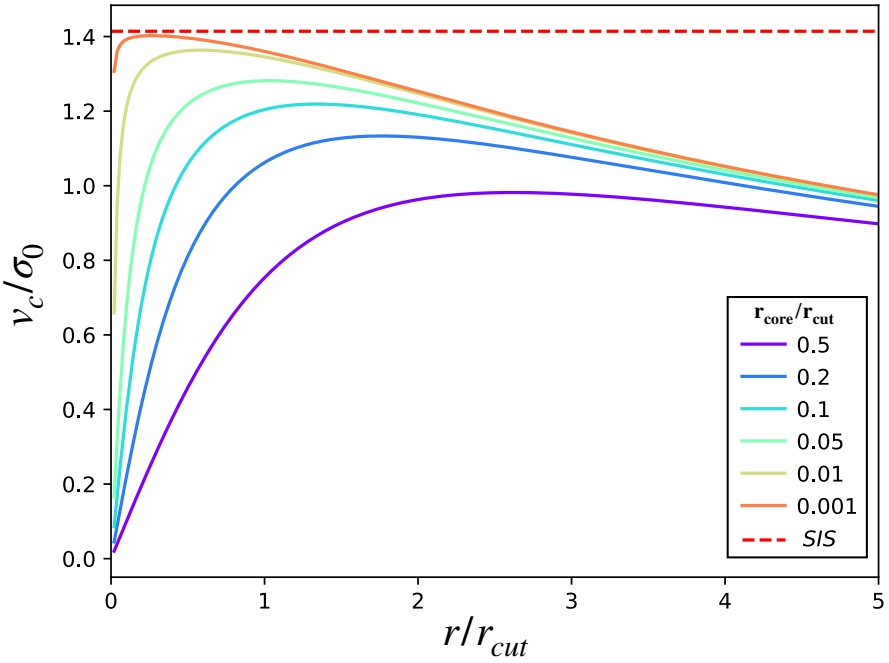


Figure 2.7: Circular velocity as a function of the distance from the galaxy centre (in units of r_{cut}) for dPIE profiles with different r_{core}/r_{cut} values. For small r_{core}/r_{cut}, the maximum of the circular velocity approaches the SIS value of $\sqrt{2}\sigma_0$ (see Eq. 2.50).

$$\sigma^2_{p,d}(R) = \frac{2G}{\Sigma(R)} \int_R^\infty \frac{\rho(r)m(r)}{r^2} \sqrt{r^2 - R^2} dr = \\ = \sigma_0^2 \frac{4}{\pi} \frac{(r_{cut} + r_{core})}{r^2_{core} r_{cut}} \left(\frac{1}{\sqrt{r^2_{core} + R^2}} - \frac{1}{\sqrt{r^2_{cut} + R^2}} \right)^{-1} \\ \cdot \int_r^\infty \frac{r_{cut} \arctan\left(\frac{r}{r_{cut}}\right) - r_{core} \arctan\left(\frac{r}{r_{core}}\right)}{r^2 \left(1 + \frac{r^2}{r_{core^2}}\right) \left(1 + \frac{r^2}{r^2_{cut}}\right) (r^2 - R^2)^{-\frac{1}{2}}} dr. \quad (2.61)$$

Finally, from Eq. 2.32, the aperture averaged LoS velocity dispersion for a spherical dPIE profile is:

$$\sigma^2_{ap,d}(R) = \frac{2\pi}{M_d(R)} \int_0^R R' \Sigma_d(R') \sigma^2_{p,d}(R') dR \equiv \frac{2}{3} \sigma_0^2 c_p^2(R). \quad (2.62)$$

In the last equation we define the projection coefficient, $c_p(R)$, as:

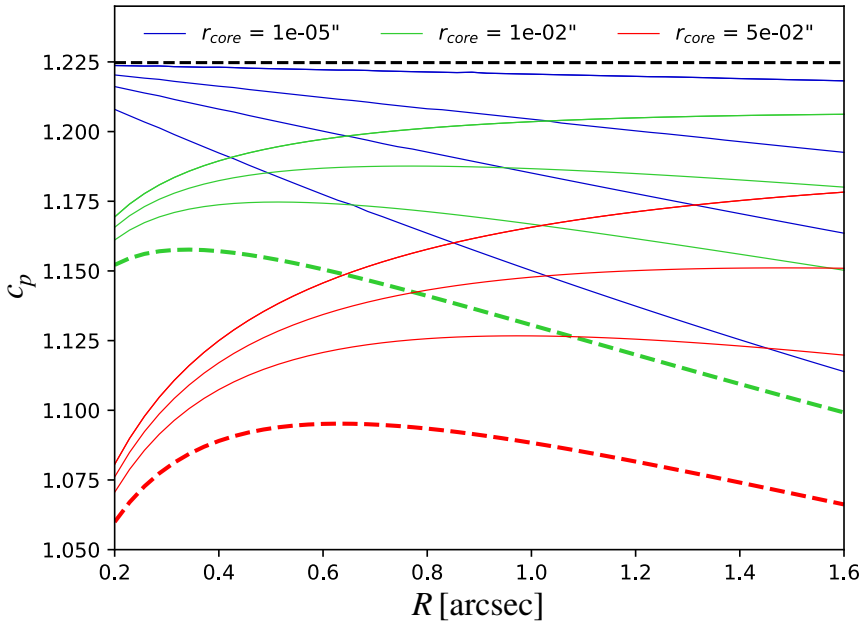


Figure 2.8: Projection coefficient c_p as a function of aperture radius R in arcseconds. Curves associated to the same r_{core} are grouped with the same color, with $r_{cut} = 5,\ 10,\ 20,\ 100$ arcsec, from bottom to top. The thick dashed red and green lines correspond to typical values of r_{core} and r_{cut} derived from our lensing+dynamics modeling of clusters (see Section 6). The dashed black horizontal line corresponds to $\sqrt{3/2}$. (Figure from Bergamini et al. 2019)

$$c_p^2(R) = \frac{6}{\pi}\frac{r_{core}+r_{cut}}{r_{core}^2 r_{cut}}\left(\sqrt{r_{core}^2+R^2}-r_{core}-\sqrt{r_{cut}^2+R^2}+r_{cut}\right)^{-1} \cdot \int_0^R R' \int_{R'}^{\infty} \frac{r_{cut}\arctan\left(\frac{r}{r_{cut}}\right)-r_{core}\arctan\left(\frac{r}{r_{core}}\right)}{(1+r^2/r_{core}^2)(1+r^2/r_{cut}^2)}\frac{\sqrt{r^2-R'^2}}{r^2}drdR'. \quad (2.63)$$

The reason why, in Eq. 2.62, we leave the factor 2/3 outside from the definition of c_p will be clear in the following chapter.
Given a spherical dPIE, with core radius r_{core} and truncation radius r_{cut}, and considering an aperture radius R, we can numerically compute the projection coefficient $c_p(R)$. This coefficient can be used to determine the aperture average LoS velocity dispersion, $\sigma_{ap,d}(R)$, form the central dPIE velocity dispersion, σ_0, and vice-versa.
In Fig. 2.8, we show the projection coefficient as a function of aperture radius R, for different values of r_{core} and r_{cut}. The dashed black line indicates the asynthotic value of $c_p = \sqrt{3/2}$ corresponding to the limit $r_{core} \to 0$ and $r_{cut} \to \infty$, where the dPIE

reduces to a SIS.

In the case of an annular aperture with internal radius R^{in} and external radius R^{ex}, we can compute the aperture average LoS velocity dispersion, within the annulus, through the following generalization of the Eq. 2.62:

$$\sigma^2_{ring,d}(R) = \frac{2\pi}{M_d(R)} \int_{R^{in}}^{R^{ex}} R' \Sigma_d(R') \sigma^2_{p,d}(R') dR. \tag{2.64}$$

This can be used when spatially resolved velocity dispersion profiles are available.

2.6 Scaling relations for elliptical galaxies

Systematic studies of galaxies proprieties in the past years led to discover several empirical scaling relations between their kinematic and photometric parameters. Recently, a physical justification for the majority of these laws were found and they became powerful tools to study the formation and the evolution of the galaxies in the Universe
In this section, the main scaling relations for the elliptical galaxies will be discussed.

Faber-Jackson relation

In 1976 Sandra Faber and Roger Jackson discovered a strong correlation between the luminosity and the central velocity dispersion of elliptical galaxies (Faber & Jackson 1976). This relation, known as 'Faber-Jackson relation' (hereafter FJ), is expressed by the following equation:

$$L \propto \sigma_0^{\delta}, \tag{2.65}$$

where the slope δ has a value between 3 and 5. Usually a $\delta = 4$ is assumed.

The physical justification of this relation comes from the virial theorem. Assuming a constant mass-to-light ratio ($M \propto L$) for elliptical galaxies, the Eq. 2.38 reduces to $L \propto \frac{\langle \sigma^2_{los} \rangle r_{gal}}{G}$. Since the mean surface brightness of a galaxy is $\langle I \rangle = \frac{L}{4\pi^2 r^2_{gal}}$ (see Eq. 2.4), we can rewrite the previous equation as: $L \propto \frac{\langle \sigma^2_{los} \rangle^4}{4\pi G^2 \langle I \rangle}$. Finally, considering an isothermal profile for the ellipticals, we can make the substitution $\langle \sigma^2_{los} \rangle = \sigma_0$ (see Eq. 2.52). Thus, the FJ scaling relation is proved.

Since the FJ relation predicts the intrinsic luminosity of a galaxy from a measure of its central velocity dispersion, it can be used to obtain an estimation of the galaxy distance.

Kormendy's relation

In 1977 John Kormendy discovered a relation between the length of the effective radius of an elliptical galaxy and its surface brightness within this radius.

$$\langle I \rangle_e \propto R_e^{\epsilon}, \quad \text{where} \quad \epsilon \sim -1.2. \tag{2.66}$$

Replacing in Eq. 2.5 the Kormendy's relation we obtain:

$$L \propto \langle I \rangle_e^{-2/3} \tag{2.67}$$

Thus, the Kormendy's relation predicts that ellipticals with higher luminosity have larger effective radii, but a lower mean surface brightness within R_e.

Fundamental Plane

A less scattered description for observed elliptical galaxies is obtainable considering the effective radius, the central velocity dispersion and the mean surface brightness within R_e, together in the following three parameters relation:

$$R_e = c\, \sigma_0^a \, \langle I \rangle_e^b , \quad \text{or in log} \quad \log R_e = b \, \langle \mu \rangle_e + a \log \sigma_0 + c. \tag{2.68}$$

The values of the two slopes a and b and the constant c depend on the considered photometric band.
This equation defines a plane in the three-parameters space (R_e, σ_0, $\langle I \rangle_e$) known as fundamental plane (hereafter FP, Djorgovski & Davis 1987; Dressler et al. 1987). Thus, the FJ and the Kormendy relations are simply the different projections of the FP on the different two-parameters spaces.

A simple physical justification of the FP can be obtained from the virial theorem (Treu 1998). The virial mass of a galaxy within its effective radius is $M_e \propto \sigma_0^2 R_e$ (see Eq. 2.38). Since the total luminosity within R_e is $L_e \propto \langle I \rangle_e R_e^2$ (Eq. 2.5), we obtain:

$$R_e \propto \frac{L_e}{M_e} \frac{\sigma_0^2}{\langle I \rangle_e}. \tag{2.69}$$

Studying a sample of cluster member elliptical galaxies belonging to the Virgo cluster, Bender et al. (1992) found the following values for the slopes in Eq. 2.68: $a = 1.4$ and $b = -0.85$. Assuming these slopes and a mass-to-light ratio that scales as $\frac{M}{L} \sim L^{0.2}$, we can immediately derive the theoretical FP in Eq. 2.69.

The availability of high quality photometric and spectroscopic data permitted to test the validity, and the evolution, of the FP relation over a wide redshift range, both for field and clusters galaxies (e.g. Treu et al. 2001; Fernández Lorenzo et al. 2011; Jørgensen et al. 2006; Holden et al. 2005).

$\mathbf{D_n}$-σ relation

The D_n-σ scaling law relates the central velocity dispersion of an elliptical galaxy with the physical quantity D_n. D_n represents the length of the diameter of an ellipse within which the average surface brightness $\langle I \rangle_n$ assumes a value of 20.75 mag/arcsec^2 in the B-band. In particular Dressler et al. (1987), using the measured parameters of 97 elliptical galaxies of six reach clusters, found:

$$D_n \propto \sigma_0^{1.33}. \tag{2.70}$$

This relation exhibits a lower intrinsic scatter ($\sim 15\%$) than the Faber-Jackson's relation.

Chapter 3

Gravitational Lensing

One of the main predictions of the Einstein theory of the General Relativity (hereafter GR) is the bending of the light rays caused by the presence of nearby masses.
The Fermat's principle states that photons always travel along path of stationary optical length with respect to variations of the path, i.e. they follow the so called Space-Time (ST) geodesics. In an empty region of ST (Minkowski ST) a geodesic is simply a straight line and as consequence photons move on straight paths. Conversely, if the region hosts a mass distribution the ST becomes curved and light rays are deflected. The bending of the light rays, emitted by a distant source, caused by mass distributions along the LoS is known as gravitational lensing, while these mass distributions are called gravitational lenses.

In this chapter we summarize the main equations governing the gravitational lensing theory. Moreover, we describe several mass distribution models widely used to parameterize the gravitational lenses and especially galaxies in cluster lens models. In the last section we present the public software `LensTool` used, in the following chapters, to develop the galaxy cluster lens models.

3.1 Lensing refraction index and time delays

A key concept of the Einstein theory of GR is the ST infinitesimal interval, or line element, ds^2:

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu. \tag{3.1}$$

The tensor $g_{\mu\nu}$ corresponds to the ST metric, while dx^μ and dx^μ are the infinitesimal displacements along the four ST directions. In this equation we assume the Einstein notation for the summations.
Most of the gravitational lensing phenomena (except if the lenses are extremely dense objects such as super-massive black holes, Event Horizon Telescope Collaboration et al. 2019) can be described in the weak-field limit approximation. In this approximation the ST metric $g_{\mu\nu}$ can be expressed as (Carroll 1997):

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \quad \text{with} \quad |h_{\mu\nu}| << 1, \tag{3.2}$$

where $\eta_{\mu\nu}$ is the Minkowski ST metric ($diag[-c^2, 1, 1, 1]$) and $h_{\mu\nu}$ is a small perturbation on the flat ST. If we assume static gravitational fields (i.e. fields that vary much more slowly than the speed of light c) and mass distributions made by dust [1], the perturbation $h_{\mu\nu}$ takes the diagonal form $h_{\mu\nu}{=}diag[\frac{-2\phi}{c^2}, \frac{-2\phi}{c^2}, \frac{-2\phi}{c^2}, \frac{-2\phi}{c^2}]$, where ϕ is the gravitational potential. Replacing this expression in Eq. 3.2, the ST interval can be written as:

$$ds^2 = -\left(1 + \frac{2\phi}{c^2}\right) c^2 dt^2 + \left(1 - \frac{2\phi}{c^2}\right) d\mathbf{x}^2. \tag{3.3}$$

Photons propagate along null line elements ($ds^2 = 0$, null proper time), covering a distance $|d\mathbf{x}|$ in a coordinate time dt. Thus, from Eq. 3.3 we can immediately derive the speed of light c' (as measured by an external observer) in a region filled by a gravitational field (Meneghetti 2019):

$$c' = \frac{|d\mathbf{x}|}{dt} = c\sqrt{\frac{1 + \frac{2\phi}{c^2}}{1 - \frac{2\phi}{c^2}}} \simeq c\left(1 + \frac{2\phi}{c^2}\right). \tag{3.4}$$

In the last step we assumed $\frac{2\phi}{c^2} << 1$.
Using this equation we can finally derive an expression for the index of refraction, n, of a curve ST region:

$$n = \frac{c}{c'} = \frac{1}{\left(1 + \frac{2\phi}{c^2}\right)} \simeq 1 - \frac{2\phi}{c^2}. \tag{3.5}$$

[1] in GR we call dust a pressureless fluid that well approximate most of the massive bodies in the Universe such as stars, galaxies, dark matter, etc.

Since $\phi \leq 0$, we always have $n \geq 1$ and the speed of light is always reduced by the presence of a gravitational field. In particular, c' is smaller where the gravitational field is stronger.
The time delay suffered by a light ray due to the presence of a gravitational field is known as Shapiro delay and it is given by the formula:

$$\Delta t = \int \frac{\mathrm{d}l}{c'} - \int \frac{\mathrm{d}l}{c} = \int (n-1)\, dl = -\frac{2}{c^3}\int \phi \mathrm{d}l. \tag{3.6}$$

3.2 Deflection angle

The Fermat's principle states that light rays always propagate along path of stationary optical length with respect to variations of the path. Thus, the photon light-paths, $\mathbf{S}$, have to satisfy the following relation (Meneghetti 2019):

$$\delta \int_A^B n\left(\mathbf{S}(l)\right) dl = 0. \tag{3.7}$$

Using simple arguments we can demonstrate that all the solutions to this equation satisfy also the Euler equation:

$$\frac{d}{d\lambda}\left(n\mathbf{e}\right) - \boldsymbol{\nabla} n = 0, \tag{3.8}$$

where $\mathbf{e}$ is the tangent unit vector along the light path and λ is a suitable curve parameter chosen such that $\frac{d\mathbf{S}}{d\lambda} \equiv \mathbf{e}$. From the Euler equation we have:

$$\frac{d\mathbf{e}}{d\lambda} = \frac{1}{n}\left[\boldsymbol{\nabla} n - \mathbf{e}\left(\boldsymbol{\nabla} n \cdot \mathbf{e}\right)\right] = \frac{1}{n}\boldsymbol{\nabla}_\perp \ln n \sim -\frac{2}{c^2}\boldsymbol{\nabla}_\perp \phi. \tag{3.9}$$

In the last step we substitute the expression for the refraction index (Eq. 3.5), and we make the approximation $\ln n \sim \frac{2\phi}{c^2}$. The amount of bending of a light ray, due to the gravitational potential, is quantify by the deflection angle vector $\hat{\boldsymbol{\alpha}}$ (see Fig. 3.1) given by the integral of $-\frac{d\mathbf{e}}{d\lambda}$ along the photon path:

$$\hat{\boldsymbol{\alpha}} = \frac{2}{c^2}\int_{\lambda_A}^{\lambda_B} \boldsymbol{\nabla}_\perp \phi \, d\lambda \tag{3.10}$$

Considering a light ray that propagates along the z axis and assuming small deviation of the photon trajectories, we can substitute the integration variable λ

in Eq. 3.10 with z. Moreover, considering a distance between the observer and the source, D_{OS}, much grater than the region filled by the gravitational potential (at $z = 0$) we can write:

$$\hat{\boldsymbol{\alpha}} = \frac{2}{c^2} \int_{-\infty}^{\infty} \boldsymbol{\nabla}_{\perp} \phi \, dz \tag{3.11}$$

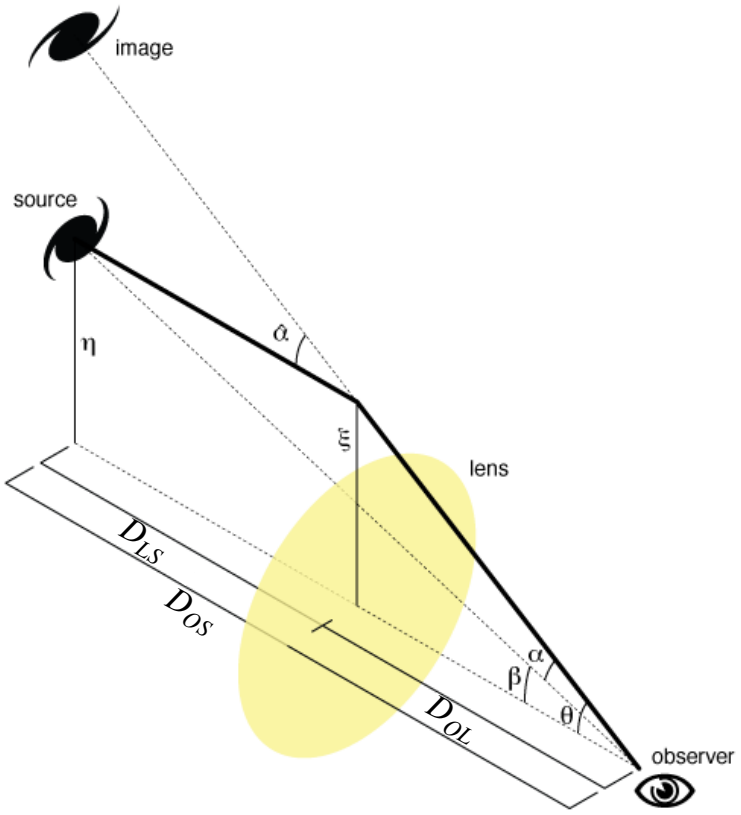


Figure 3.1: Schematic representation of a gravitational lensing system. The gravitational lens (in yellow) has an angular diameter distance D_{LS} from the source and D_{OL} from the observer. Instead the distance between the observer and the source is D_{OS}. The 2D distances, respect to the optical axis, on the lens and source planes are $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ respectively. (Figure from: en.wikipedia.org/wiki/Gravitational_lensing_formalism).

3.3 Lens equation

From the Eq. 3.11 and assuming small deviations for the light rays we can immediately derive, through simple geometrical considerations (see Fig. 3.1), the

following lens equation:

$$\boldsymbol{\theta} D_{OS} = \boldsymbol{\beta} D_{OS} + \hat{\boldsymbol{\alpha}} D_{LS}. \tag{3.12}$$

The lens equation combines the main quantities of a lens system i.e.: the deflection angle vector ($\hat{\boldsymbol{\alpha}}$); the angular diameter distances between observer-source (D_{OS}), observer-lens (D_{OL}) and lens-source (D_{LS}); and the angular positions of the source ($\boldsymbol{\beta}$) and of the lensed images ($\boldsymbol{\theta}$) respect to the optical axis.

Defining the reduced deflection angle vector $\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{D_{LS}}{D_{OS}} \hat{\boldsymbol{\alpha}}(\boldsymbol{\theta})$, the lens equation assumes the simpler form:

$$\boldsymbol{\theta} = \boldsymbol{\beta} + \boldsymbol{\alpha}. \tag{3.13}$$

The lens equation can also be expressed in a dimensionless form introducing a convenient length scale, ξ_0, on the lens plane and the corresponding distance, $\eta_0 = \xi_0 \frac{D_{OS}}{D_{OL}}$, on the source plane. Since $\boldsymbol{\eta} = \boldsymbol{\beta} D_{OS}$ is the source position respect to the optical axis and $\boldsymbol{\xi} = \boldsymbol{\theta} D_{OL}$ is the impact parameter of the light ray at the lens position (see Fig. 3.1), we can define the two dimensionless quantities $\mathbf{x} = \frac{\boldsymbol{\xi}}{\xi_0}$ and $\mathbf{y} = \frac{\boldsymbol{\eta}}{\eta_0}$. Similarly, the dimensionless deflection angle vector is given by: $\boldsymbol{\alpha}(\mathbf{x}) = \frac{D_{OL} D_{LS}}{\xi_0 D_{OS}} \hat{\boldsymbol{\alpha}}(\xi_0 \mathbf{x})$. Thus, replacing these quantities in Eq. 3.13 we finally obtain the dimensionless lens equation:

$$\mathbf{y} = \mathbf{x} - \boldsymbol{\alpha}(\mathbf{x}). \tag{3.14}$$

3.4 Effective lensing potential, convergence and shear

Another useful quantity to describe the effects of a gravitational lens on the light rays emitted by a background source is the effective lensing potential, $\hat{\psi}(\boldsymbol{\theta})$. This quantity is defined as a scaled lens-plane projection of the 3D gravitational potential ϕ (Meneghetti 2019):

$$\hat{\psi}(\boldsymbol{\theta}) = \frac{2 D_{LS}}{c^2 D_{OL} D_{OS}} \int_{-\infty}^{+\infty} \phi\left(D_{OL}\boldsymbol{\theta}, z\right) \mathrm{d}z. \tag{3.15}$$

While the dimensionless form of the effective lensing potential is given by:

$$\psi(\mathbf{x}) = \frac{D_L^2}{\xi_0^2}\hat{\psi}(\boldsymbol{\theta}).$$

From Eqs. 3.15 and 3.11, we can derive the following two relations for the effective lensing potential:

$$\boldsymbol{\nabla}_{\boldsymbol{\theta}}\hat{\psi}(\boldsymbol{\theta}) = \frac{D_{LS}}{D_{OL}D_{LS}}\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}), \text{ or dimensionless } \boldsymbol{\nabla}_{\mathbf{x}}\psi(\mathbf{x}) = \boldsymbol{\alpha}(\mathbf{x}). \tag{3.16}$$

The gravitational lenses produce deformations on the shape of the images of a background source. These deformations are due to the different deflections suffered by the source light rays intersecting the lens plane with different impact parameters. In theory, to derive the final shape of an image we need to solve the lens equation for a huge number of light rays. However, if the source angular size is much smaller than the angular size over which the physical properties of the lens change, we can define two quantities that describe the first-order distortions of the lensed images. These two quantities are known as convergence and shear.

- **Convergence:** The convergence is given by the expression:

$$\kappa(\boldsymbol{\theta}) = \frac{\Sigma(D_{OL}\boldsymbol{\theta})}{\Sigma_{cr}}, \text{ or dimensionless } \kappa(\mathbf{x}) = \frac{\Sigma(\mathbf{x})}{\Sigma_{cr}}. \tag{3.17}$$

  where $\Sigma(D_{OL}\boldsymbol{\theta})$ ($\Sigma(\mathbf{x})$) is the surface mass density distribution (dimensionless surface mass density distribution) of the gravitational lens and Σ_{cr} is the critical surface mass density of the lensing system:

$$\Sigma_{cr} = \frac{c^2 D_S}{4\pi G D_L D_{LS}}. \tag{3.18}$$

  The convergence can also be expressed in term of the effective potential through the simple equation:

$$\kappa(\boldsymbol{\theta}) = \frac{1}{2}\triangle_{\boldsymbol{\theta}}\hat{\psi}(\boldsymbol{\theta}), \text{ or dimensionless } \kappa(\mathbf{x}) = \frac{1}{2}\triangle_{\mathbf{x}}\psi(\mathbf{x}). \tag{3.19}$$

- **Shear:** The shear, $\boldsymbol{\gamma} = (\gamma_1, \gamma_2)$, is a 2D lens plane pseudo-vector whose components are given by:

$$\begin{aligned}\gamma_1(\boldsymbol{\theta}) &= \frac{1}{2}\left(\frac{\partial^2\hat{\psi}(\boldsymbol{\theta})}{\partial\theta_1\partial\theta_1} - \frac{\partial^2\hat{\psi}(\boldsymbol{\theta})}{\partial\theta_2\partial\theta_2}\right), \text{ or } \gamma_1(\mathbf{x}) = \frac{1}{2}\left(\frac{\partial^2\psi(\mathbf{x})}{\partial x_1\partial x_1} - \frac{\partial^2\psi(\mathbf{x})}{\partial x_2\partial x_2}\right). \\ \gamma_2(\boldsymbol{\theta}) &= \frac{\partial^2\hat{\psi}(\boldsymbol{\theta})}{\partial\theta_1\partial\theta_2} = \frac{\partial^2\hat{\psi}(\boldsymbol{\theta})}{\partial\theta_2\partial\theta_1}, \text{ or } \gamma_2(\boldsymbol{x}) = \frac{\partial^2\psi(\mathbf{x})}{\partial x1\partial x_2} = \frac{\partial^2\psi(\mathbf{x})}{\partial x_2\partial x_1}.\end{aligned} \tag{3.20}$$

The convergence describes an isotropic deformation on an image of the source, i.e. the image preserves the original shape of the source but with rescaled dimensions. Conversely, the shear stretches the source intrinsic shape along one privileged direction.
In particular, a circular source with radius r, is mapped by the lens equation into an ellipse whose semi-major and semi-minor axes are respectively:

$$d_a = \frac{r}{1-\kappa-\gamma}, \qquad d_b = \frac{r}{1-\kappa+\gamma}, \quad \text{with} \quad \gamma = \sqrt{\gamma_1^2+\gamma_2^2}.$$

The combination of the image distortions with the surface brightness conservation (Liuville's theorem) produces an effect called magnification. If the angular size of an image increases (or decreases) respect to the angular size of the source without the lens, the received flux from the source is magnified (de-magnified). The magnification, μ, is defined as the ratio between the flux received from a lensed image and the those coming from the unlensed source. The magnification can be expressed in term of convergence and shear modulus through the following expression:

$$\mu = \frac{1}{(1-\kappa^2)-\gamma^2}. \tag{3.21}$$

We can also quantify the magnification in tangential and radial directions with the equations:

$$\mu_t = \frac{1}{1-\kappa-\gamma}, \qquad \mu_r = \frac{1}{1-\kappa+\gamma}$$

For $1-\kappa-\gamma=0$ or $1-\kappa+\gamma=0$ the magnification of the images goes to infinity along the radial or tangential directions respectively. These two conditions define two curves on the lens plan known as radial and tangential critical lines. Images close to radial critical lines are distorted perpendicularly respect to these lines while those near to the tangential critical lines are distorted tangentially (see Fig. 3.2). The points on the source plane that are mapped by lens equation on to critical lines define source plane curves called caustics.

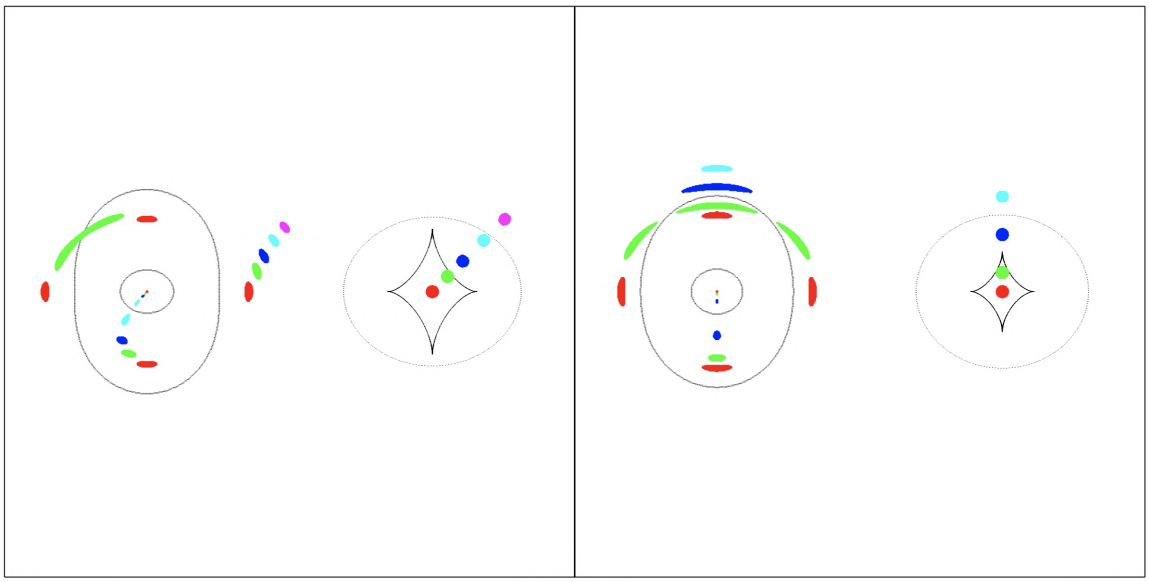


Figure 3.2: Compact source moving away from the center of an elliptical lens. *Left panel:* source crossing a fold caustic; *right panel:* source crossing a cusp caustic. Within each panel, the diagram on the left shows critical lines and image positions and the diagram on the right shows caustics and source positions. (Figure from: Narayan & Bartelmann 1996).

3.5 Circular and elliptical gravitational lenses

If the gravitational lens has a circularly symmetric surface mass distribution all the previous equations can be expressed as a function of the lens-plane projected distance, R, from the lens center.
In particular, the convergence becomes:

$$\kappa(R) = \frac{\Sigma(R)}{\Sigma_{cr}}, \tag{3.22}$$

where $\Sigma(R)$ is the surface mass density profile of the lens.
For the deflection angle, the effective lensing potential and the shear we have the following expressions (Oguri 2019):

$$\begin{aligned} \alpha(R) &= \frac{2}{R}\int_0^R R'\kappa(R')dR', \\ \hat{\psi}(R) &= \int_0^R \alpha(R')dR' = 2\int_0^R \kappa(R')R' \ln\left(\frac{R}{R'}\right)dR', \\ \gamma(R) &= \frac{\alpha(R)}{R} - \kappa(R) \end{aligned} \tag{3.23}$$

Due to circular symmetry both the deflection angle vector and the shear are scalar quantities. Replacing in Eqs. 3.22 and 3.23 the expression for the projected surface mass distribution of a SIS (Eq. 2.47) or of a circular dPIE (Eq. 2.54) we can immediately derive all the relevant lensing quantities for these two kinds of models.

All the previous formulas can be generalized to the case of elliptical lenses through the simple substitution:

$$R \to \tilde{R} \equiv \sqrt{\frac{\tilde{r}_1^2}{(1-\tilde{e})} + (1-\tilde{e})\tilde{r}_2^2}, \tag{3.24}$$

where the ellipticity is defined as $\tilde{e} = 1 - \frac{b}{a}$, with a and b the semi-minor and semi-major axis of the ellipse. Instead, the two coordinated $\tilde{r}_1$ and $\tilde{r}_2$ are given by:

$$\tilde{r}_1 = r_1 \cos\theta_e + r_2 \sin\theta_e, \qquad \tilde{r}_2 = -r_1 \sin\theta_e + r_2 \cos\theta_e,$$

where r_1 and r_2 are the two components of the lens-plane vector $\mathbf{r}$ (a scalar in the circular case), while θ_e is the angle respect to the r_1 direction.

3.6 Deflection angle for a mass distribution

While a circular or an elliptical mass distribution may be sufficient to describe a single lens galaxy, galaxy clusters are usually the sum of hundreds of halos with different shapes and sizes (see the next section). In this section we generalize the Eq. 3.11 for the deflection angle vector to the case of an asymmetrical, even complex, mass distribution.

We start considering a gravitational lens system made by N point masses distributed on the lens plane. If a source light ray intersects the lens plane in a position $\boldsymbol{\xi}$, the total deflection angle vector is simply the sum of the deflection angles computed for each point mass:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \sum_i^N \hat{\boldsymbol{\alpha}}_i(\boldsymbol{\xi} - \boldsymbol{\xi}_i) = \frac{4G}{c^2} \sum_i M_i \frac{\boldsymbol{\xi} - \boldsymbol{\xi_i}}{|\boldsymbol{\xi} - \boldsymbol{\xi}_i|^2}.$$

The $\boldsymbol{\xi}_i$ are the positions of the point masses, M_i, on the lens plane.

This formula can be generalized to the more realistic case of a lens with a continuous 3D mass density distribution. If the longitudinal size of mass distribution

is much smaller than the distances D_{OL} and D_{LS}, we can consider the lens mass as distributed on the lens plane with surface mass density $\Sigma(\boldsymbol{\xi})$ (thin lens approximation). Under this approximation, the total deflection angle vector is the sum of all the contributions coming from the infinitesimal surface mass elements, $\Sigma(\boldsymbol{\xi})d^2\xi$, i.e. (Meneghetti 2019):

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4G}{c^2}\int \frac{(\boldsymbol{\xi}-\boldsymbol{\xi'})\Sigma(\boldsymbol{\xi'})}{|\boldsymbol{\xi}-\boldsymbol{\xi'}|^2}d^2\xi'.$$

3.7 Cluster lens models through the parametric software `LensTool`

In this section, we briefly describe the public software `LensTool` (Kneib et al. 1996; Jullo et al. 2007; Jullo & Kneib 2009) used to create the cluster lens models in the following chapters of the thesis. Like several other codes, `LensTool` was developed to constrain, in a parametric way, the total mass distribution of a galaxy cluster from the observed multiple image positions (representing the constrains to the lens model).
Unlike non-parametric software, in the parametric codes the number and the type of halos, used to model the cluster mass distribution, has to be defined a priory. The parametric approach has two main advantages: firstly, the assumption of meaningful models for the mass distribution leads to results that are directly comparable to real observed physical quantities (e.g. ellipticities, velocity dispersions, scale radii, etc.); secondly, the number of free parameters in the model is strongly reduced. This is particularly important to avoid over-fitting problems and the necessity of regularization terms.

`LensTool` determines the best-fit cluster lens model using a Bayesian Markov chain Monte Carlo (MCMC) approach. In particular it looks for the set of model free parameters that minimize the following lens-plane likelihood (Jullo et al. 2007):

$$\mathcal{L} = \prod_{i=1}^{N} \frac{1}{\prod_{j=1}^{n_i}\Delta x_{i,j}\sqrt{2\pi}} e^{-\chi_i^2/2}, \tag{3.25}$$

where N is the number of sources and n_i is the number of multiple images associated to the source i (usually called a family). The χ^2 in the equation is given by:

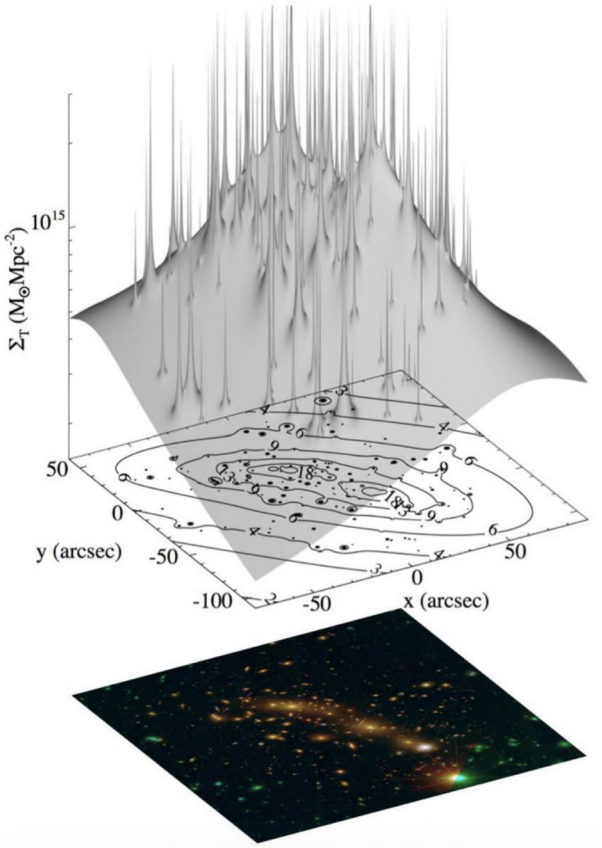


Figure 3.3: Total surface mass density $\Sigma_{\rm T}$ in the inner regions of MACS J0416.1−0403 reconstructed from the best-fitting strong lensing model. The different contributions of the two extended dark-matter halo and many candidate cluster member components are visible. The contour levels on the lens plane are in units of $10^{14}\,{\rm M}_\odot{\rm Mpc}^{-2}$. (Figure from Grillo et al. 2015).

$$\chi_i^2 = \sum_{j=1}^{n_i} \frac{\left[\mathbf{x}_{obs}^j - \mathbf{x}^j(\Psi)\right]^2}{\Delta x_{i,j}^2}. \tag{3.26}$$

The $\mathbf{x}_{obs}^j$ are the observed positions of the multiple images on the lens plane; the $\mathbf{x}^j(\mathbf{\Psi})$ are their predicted positions given the set of model parameters, $\mathbf{\Psi}$; and the $\Delta x_{i,j}$ are the uncertainties on observed positions.

The total mass distribution of a galaxy cluster (or equivalently the total gravitational potential) is usually parametrized as the sum of three main components (see Fig. 3.3; e.g. Natarajan & Kneib 1997; Natarajan et al. 2017; Caminha et al. 2017b,a, 2016; Bonamigo et al. 2018; Grillo et al. 2015):

$$\phi_{tot} = \sum_{i=1}^{N_h} \phi_i^{halo} + \sum_{j=1}^{N_{sf}} \phi_k^{shear+foreg} + \sum_{k=1}^{N_g} \phi_i^{gal} \tag{3.27}$$

The first sum runs over the smooth cluster-scale halos of the cluster. These halos are mostly made by Dark Matter, while a minor fraction of their mass is in the form of intra-cluster gas and intra-cluster light (stars that are not within the cluster galaxies). The lens models of spherical symmetric clusters usually contain a single cluster-scale halo, while to model asymmetric clusters (such as the merger clusters) we need several large halos.
The second sum in Eq. 3.27 takes into account the massive structures in outer regions of the cluster and the mass distributions along the LoS. Indeed, large masses in the outskirt of a cluster can produce additional shear terms in the lens model, while the massive foreground structure along the LoS must be parametrized as additional halos.
Finally, the last sum describes the clumpy sub-halo component of the cluster, i.e. the DM and baryonic content of its N_g (usually $N_g > 100$) cluster member galaxies. These sub-halos are usually parametrized as circular dPIE profiles (see Eq. 2.45) of vanishing core radii, while their positions are fixed on the centroids of the light emission of the galaxies. Thus, every cluster member introduces two additional free parameters (σ_0 and r_{cut}) in the lens model. However, the number of constrains to the model (proportional to the number of observed multiple images) is usually not sufficient to constrain these $2N_g$ extra free parameters. To overcame this problem the following scaling relations for the cluster member velocity dispersions and truncation radii are assumed (Jullo et al. 2007; Brainerd et al. 1996):

$$\sigma_{LT,i}^{gal} = \sigma_{LT}^{ref} \left(\frac{L_i}{L_0} \right)^{\alpha}, \tag{3.28}$$

$$r_{cut,i}^{gal} = r_{cut}^{ref} \left(\frac{L_i}{L_0} \right)^{\beta_{cut}}. \tag{3.29}$$

L_0 is a reference luminosity, while L_i is the luminosity of the i-th cluster galaxy with velocity dispersion $\sigma_{LT,i}^{gal}$ and truncation radius $r_{cut,i}^{gal}$.
Using these relations the number of free parameters of the clumpy component reduces to four, i.e.: the reference velocity dispersion, σ_{LT}^{ref}, corresponding to the luminosity L_0; the reference truncation radius, r_{cut}^{ref}; and the two slopes α and β_{cut}.

Note that in Eq. 3.28 we introduce the fiducial velocity dispersions σ_{LT}. In `LensTool`, the dPIE profiles are always parametrized through σ_{LT}, instead of the

most commonly used σ_0. However, we have the following relation between the fiducial and the central velocity dispersions:

$$\sigma_0 = \sqrt{\frac{3}{2}}\sigma_{LT} \tag{3.30}$$

3.7.1 σ_0-r_{cut} degeneracy

We conclude this chapter describing a well known degeneracy existing between the central velocity dispersion σ_0 (or equally σ_{LT}) and the truncation radius r_{cut} of a circular dPIE lens of vanishing core radius. The study of possible ways to reduce this degeneracy is one of the main goals of the following chapters.

The total projected mass of a circular dPIE, within an aperture of radius R, is given by Eq. 2.56. Considering a vanishing core radius ($r_{core} = 0$), and an aperture radius equal to the galaxy Einstein radius (R_E), this equation reduces to:

$$M_d(R_E) = \frac{\pi\sigma_0}{G}\left(R_E + r_{cut} - \sqrt{r_{cut}^2 + R_E^2}\right). \tag{3.31}$$

The multiple images around a galaxy well constrain its total mass within the Einstein radius. However, the Eq. 3.31 shows that this mass depends on both σ_0 and r_{cut}. In particular, an increase (decrease) of the central velocity dispersion can leave the $M_d(R_E)$ value unchanged if it is followed by a proper decrease (increase) of the truncation radius. This creates the inverse degeneracy between the values of σ_0 and r_{cut} clearly visible in the Fig. 3.4. This figure is obtained modelling the second brightest cluster member of MACS J1206.2−0847 as described in Chapter 7.

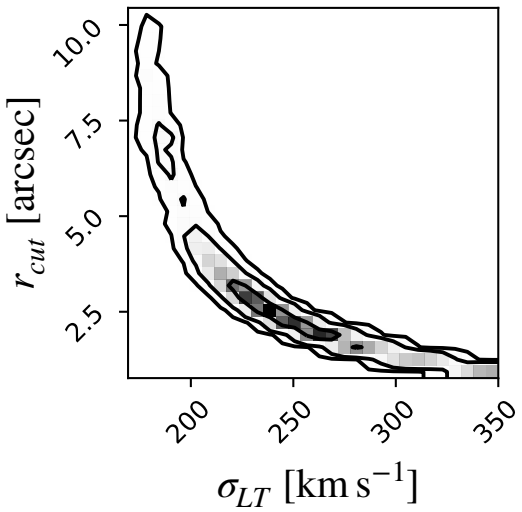


Figure 3.4: Degeneracy between the two lens model parameters σ_{LT} and r_{cut} of the second brightest cluster member of MACS J1206.2−0847. The lens model of this galaxy will be studied in details in Chapter 7.

Chapter 4

Data and instruments

The results presented in the following chapters are based on extensive and mostly coordinated observations of the core of several massive galaxy clusters, which have produce a large volume of spectroscopic and photometric data of unprecedented quality.

In particular, the imaging data were obtained with the Hubble Space Telescope (HST), while the spectroscopic data come from several Very Large Telescope (VLT) programs, utilizing the VIMOS multi-object panoramic spectrograph and the integral field spectrograph MUSE (Multi Unit Spectroscopic Explorer). In this chapter, we briefly describe the instrumentation used and the corresponding imaging and spectroscopic observational campaigns.

4.1 Imaging and photometric data: HST

The Hubble Space Telescope[1] was launched in 1990 in a stable orbit at 600 Km of height, inclined at 28.5 degree respect to the equator and with a revolution time around the earth of 96-97 minutes. Respect to the earth-base telescopes it delivers diffraction limit imaging with an angular resolution of 50 mas. HST is widely recognized as one of the most productive and ground-breaking scientific machine ever built.

The imaging data used in this thesis were obtained with the Advanced Camera for Surveys (ACS) and the Wide Field Camera 3 (WFC3) cameras. Both cameras

[1] http://www.stsci.edu/hst

include a large set of broad and narrow band-pass filters organized in a photometric system from the UV to the near-IR wavelength (see Fig. 4.1).

WFC3[2], which replaced WFPC2 in 2009 during the servicing mission (SM) 4, has two working channels, one in the ultraviolet (UVIS) covering approximately the wavelength range 2000–4000 Å and the other in the near infrared (NIR), over the range 0.9–1.7μm. In this thesis, only observations with the NIR channel are used. Its detector is a 1014×1014 array made of mercuric cadmium telluride (HgCdTe) that guarantees a high ($\sim$50%) broad-band throughput, with a FoV of $123'' \times 137''$ and a pixel scale of $0.13''$.

The **ACS**[3] camera replaced the HST Faint Objects Camera during the SM 3B, in 2002. Its detector consists of two 2048×4096 CCDs, with a pixel size of 15 μm, or $0.049''$ on the sky. It is sensitive to the light from 3700Å up to 11000Å (B, V, I, z bands) and it has a FoV of $202'' \times 202''$. Once installed, its new throughput, image quality and FoV increased the HST potential for discoveries of a factor ten. The ACS had originally two independent working channels, i.e. the Wide Field Channel (WFC) and a High Resolution Channel (HRC). After an electric fault in 2007, only the WFC was recovered during SM4 in 2009. ACS/WFC obsevations are at the core of many HST surveys, including the imaging data used in the following chapters.

4.1.1 Spectroscopic data: VLT/MUSE

The VLT, at the European Southern Observatory of the Cerro Paranal in Chile, consists in an array of four Ritchey-Chrétien Unit Telescopes (UT), with main mirrors of 8.2 m of diameter, and four auxiliary movable telescopes (AT), of 1.8 m of diameter. One of the main features of the UT telescopes is the adaptive optics. Indeed, the mirrors of the telescopes can be deformed to compensate the aberrations on the observed images induced by the atmosphere. The maximum FoV of a UT telescope is of 27 arcmin , slightly smaller than the angular size of the full moon. The eight VLT telescopes can work independently or in a combined mode that increases their total collecting area to those of a single telescope of 16 m of diameter. The telescopes can also be used as a giant interferometer (the ESO Very Large Telescope Interferometer) that allows a resolving power equivalent to those of a telescope of 100 m of diameter. In this configuration, the VLT can resolve two objects 2 m apart at the distance of the moon (0.001 arcsec for a wavelength of 1μm).

[2]http://www.stsci.edu/hst/wfc3

[3]http://www.stsci.edu/hst/acs/

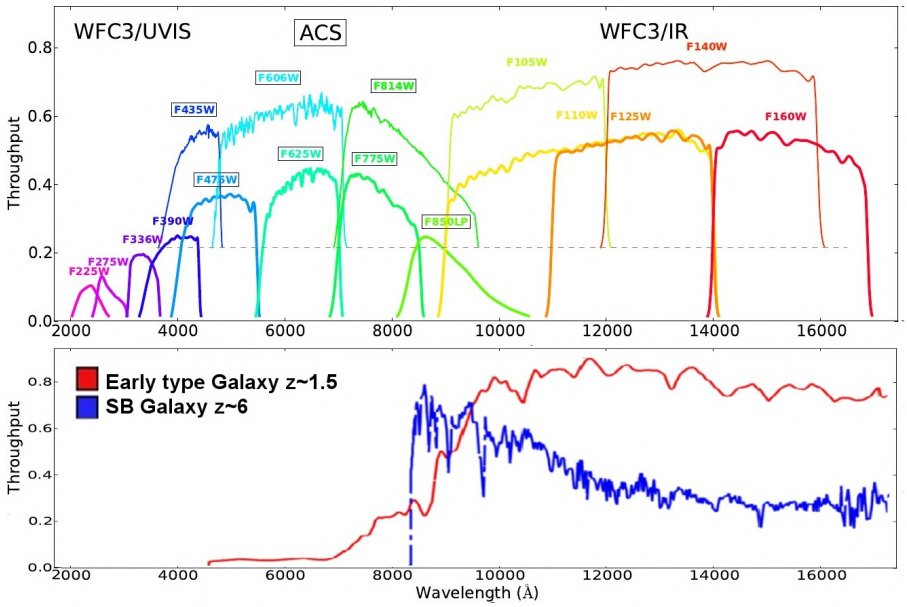


Figure 4.1: *Upper panel:* transmission curves of the HST/ACS and HST/WFC3 filters used in the CLASH survey. *Lower panel:* Comparison of a star-forming spectral energy distribution (SED) redshifted to $z = 6$ (in blue) and an early-type galaxy at $z = 1.5$ (from the COSMOS library, Ilbert et al. 2009). In both cases the strong spectral break falls in the same NIR range, and both these kind of sources appear as optical dropouts. At λ redder than the break the SEDs differ, thus the NIR photometry can be used to discriminate between high-z and low-z dropouts. (Figure from: Monna et al. 2014.)

The spectroscopic data used in the following chapters come from the MUSE[4] integral field spectrograph (Bacon et al. 2012), located at the Nasmyth B focus of the telescope VLT/UT4 (known as Yepun). It utilizes a set of image slicers and 24 spectrographs to build a data cube covering $1' \times 1'$ on the sky, with $0.2''$pixels, and therefore a total of $(60/0.2)^2 = 90000$ spectra.
MUSE combines a wide FoV with the improved spatial resolution of the adaptive optics. It is a second generation instrument (installed in 2014) that operates in the visible wavelength range from the 4650Å up to the 9300Å. It has a spectral sampling of 1.25Å/pix and a spectral resolution of $R \sim 1750$ ($R \sim 3750$) at 4650Å (9300Å). In Fig. 4.2 we show the variation of R as a function of the observed wavelength λ. Since $R = \lambda/(\Delta\lambda)_{\rm FWHM}$, we can assume an almost constant wavelength resolution of $(\Delta\lambda)_{\rm FWHM} = 2.6$Å FWHM in the range 4650-8500Å.

The finite spectral resolution produces a broadening of the lines in the observed spectra as if they were convoluted with a Gaussian function (see Chapter 5). For

[4]https://www.eso.org/sci/facilities/develop/instruments/muse.html

reasons that will become clear in the next section, it is useful to express the spectral resolution in terms of the standard deviation of the Gaussian kernel function, σ_{ist} (i.e. a velocity dispersion). Since $R = c/(\Delta v)_{\rm FWHM}$, and considering that the standard deviation of a Gaussian is 2.355 smaller than its FWHM, we can express σ_{ist} as:

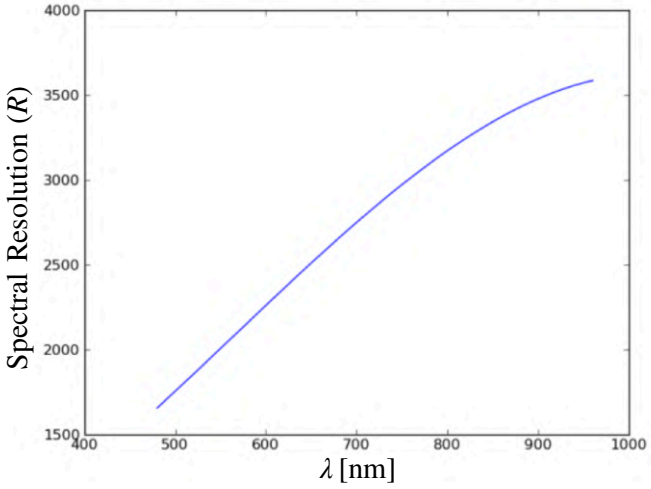
$$\sigma_{ist} = \frac{1}{2.355}\frac{c}{R} \tag{4.1}$$

At an observed wavelength of 4650Å (9300Å) we have $\sigma_{ist} \sim 80\,{\rm km\,s^{-1}}$ ($\sigma_{ist} \sim 35\,{\rm km\,s^{-1}}$).

MUSE has two operating modes (see Fig. 4.2). The first is called Wide Field Mode (WFM) while the second is the Narrow Field Mode (NFM).
The WFM observes with the maximum obtainable FoV of about $1 \times 1\,{\rm arcmin^2}$ (corresponding to 0.32 Mpc at z=0.4) sampled in pixels of $0.2 \times 0.2\,{\rm arcsec}$. Considering the typical size of a galaxy cluster at redshift z=0.4, MUSE gives the opportunity to observe, in a single shot, hundreds of member galaxies in the cluster core. The whole MUSE FoV is divided in sub-fields that are sent to 24 spectrographs, known as Integral Field Units (IFUs), that collect a spectrum for each image pixel. Then, the MUSE DATA REDUCTION SOFTWARE organizes the nearly 90000 observed spectra in a datacube, i.e. a three-dimensional virtual object that contains in two of its dimensions the observed FoV and the pixel spectra as a third dimension. The WFM is the operating mode adopted to observe the data used in this thesis.
In the NFM the FoV is reduced to an area of $7.5 \times 7.5\,{\rm arcsec^2}$ sampled in pixels of $0.025 \times 0.025\,{\rm arcsec^2}$.

4.2 HST Galaxy Clusters Programs

Most of the data used in the following chapters were collected as part of three major HST surveys of galaxy clusters: CLASH, CLASH-VLT and HFF. These surveys have provided a new spectro-photometric dataset with unprecedented quality of 27 massive clusters with redshift ranging from $z = 0.187$ (Abell 383) to $z = 0.890$ (CLJ1226+3332).
In this section, we summarize the main characteristics and scientific objectives of the three surveys.



MUSE	
Spectral range	4650.0 - 9300.0 Å
Spectral resolution at 4650Å	1750
Spectral resolution at 9300Å	3750

Wide Field Mode (WFM)	
Field of view	$1 \times 1\,\mathrm{arcmin}^2$
Spatial sampling	$0.2 \times 0.2\,\mathrm{arcsec}^2$

Narrow Field Mode (NFM)	
Field of view	$7.5 \times 7.5\,\mathrm{arcsec}^2$
Spatial sampling	$0.025 \times 0.025\,\mathrm{arcsec}^2$

Figure 4.2: *Left:* MUSE spectral resolution as a function of the wavelength, λ. *Right:* Main features of the MUSE integral field spectrograph in the WFM and NFM operating modes. (Figure from: www.eso.org).

4.2.1 CLASH survey

The Cluster Lensing And Supernova survey with Hubble (CLASH[5], Postman et al. 2012) with Principal Investigator (P.I.) Marc Postman, was one of the three selected HST Multy-Cycle Treasury Programs in 2011. During this survey, 25 massive galaxy clusters were observed in the wavelength range 2000-17000Å, using the 16 filters of the HST/WFC3 and HST/ACS cameras, for a total of 524 HST orbits (see Fig. 4.3). Moreover, Chandra observations through the Advanced CCD Imaging Spectrometer[6] (ACIS) provided X-Rays images of the clusters.
The 25 target clusters were chosen according to the following selection criteria:

- 20 clusters were chosen because they are perfect prototypes of relaxed clusters on the basis of their symmetric and smooth X-ray emission. The other five clusters were included as famous powerful gravitational lenses to maximize the chance of finding lensed galaxies.
- All the clusters were selected to be behind regions of low Galactic extinction (median $E(B-V) \sim 0.026$)

The CLASH program was completed in 2003 and its data are still producing

[5]http://www.stsci.edu/ postman/CLASH/Home.html

[6]http://cxc.harvard.edu/cal/Acis/

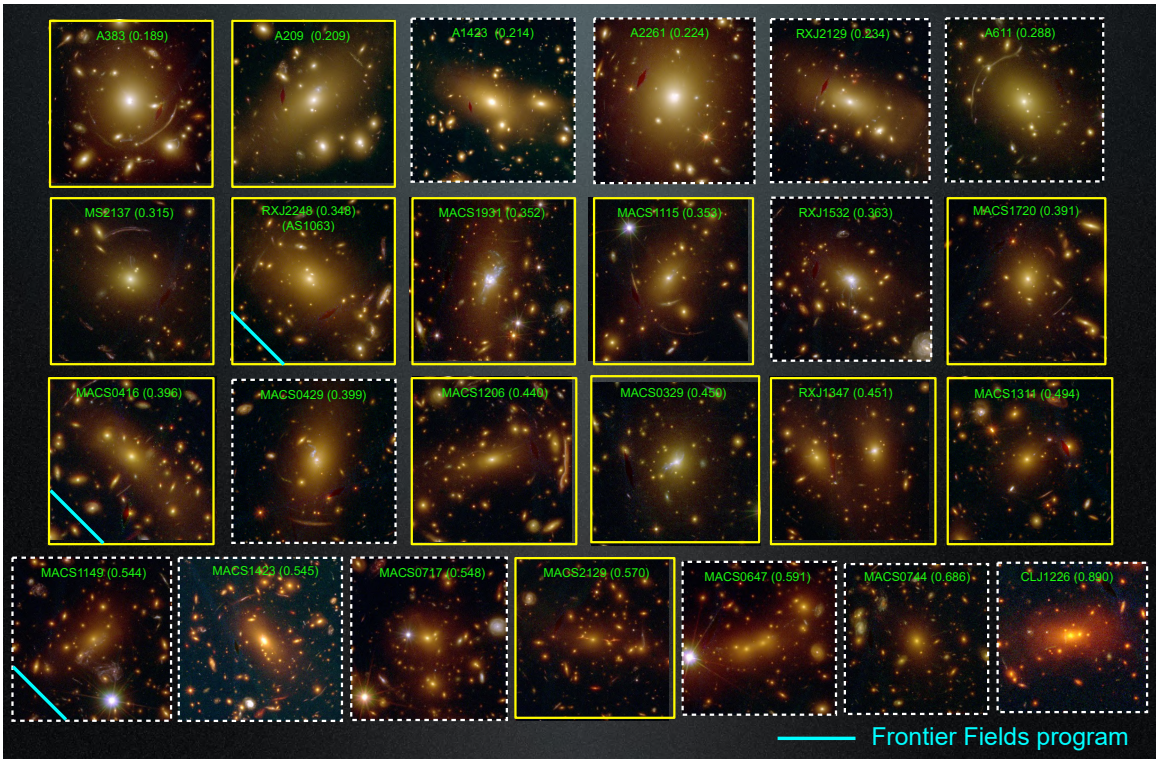


Figure 4.3: RGB cutouts of the core regions of the 25 CLASH clusters. IDs and redshifts of the clusters are written in green. Yellow boxes encircle the 13 clusters included in the CLASH-VLT program, while the blue bars mark the HFF clusters.

a great variety of new results. They range from the study of the degree of concentration of the mass profiles as a function of cluster mass by Merten et al. (2015), up to the study of the star-forming galaxies at $z \sim$ 9-10 (Coe et al. 2013; Bouwens et al. 2014), or to the detection of supernovae at high redshift (up to $z \sim 2$). An exhaustive list of publications can be found here: http://www.stsci.edu/ postman/CLASH/Publications.html.

4.2.2 CLASH-VLT program

The CLASH survey was conceived as a 16-band HST imaging program to take advantage of good photometric redshift information, essential to characterize the cluster galaxy populations, to identify high-redshift magnified galaxies, and obtain large numbers of multiply lensed images. The latter were used as constraints for new set of strong lensing models of the cluster mass distribution. Clearly, to fully address the CLASH scientific objectives, spectroscopic follow-up observations are needed. To this aim, an ESO-Large Program to obtain an extensive wide-field

spectroscopic coverage of the southern CLASH clusters was proposed and approved in 2014 (CLASH-VLT[7], P.I. Piero Rosati).

The CLASH-VLT survey carried out a spectroscopic campaign of the 13 CLASH clusters (see Fig. 4.3) accessible from the VLT, providing literally a third dimension to the CLASH images (Rosati et al. 2014). All the spectroscopic data were obtained by the Visible wide field Imager and Multi-Object Spectrograph (VIMOS), mounted on the third UT telescope of the VLT. VIMOS has a spectral range between 3600Å and 10,000Å and a spectral resolution between 13Å and 28Å FWHM. The survey provided a total of 225 hours of observations divided in 25 hours of pre-imaging acquisition and 200 hours of Multi-Object Spectroscopy (MOS). Additional key observations of the program were also the Subaru Suprime-Cam imaging data obtained for most of the clusters, and the ESO Wide Field Imager (WFI) images for the southernmost target RXJ 2248.74431.
The main goals of the CLASH-VLT survey can be summarized as follows:

- Obtain spectroscopic confirmation of at least 500 cluster member galaxies in each cluster. With this large number of members' velocities the cluster mass density profiles can be determined from a dynamical Jeans analysis with the same accuracy as the strong and weak lensing profiles.
- Measure the redshifts of over 200 lensed galaxies, including several highly magnified galaxies up to $z \sim 7$. The identification of new multiple images lensed by the clusters provides critical constrains for the cluster strong lensing models.
- Obtain an unprecedented collection of spectro-photometric data for studies of galaxy populations in a variety of environments.

4.2.3 Frontier Fields Program

Finally, the Hubble Frontier Fields Program[8] (HFF) is the third important survey designed for the study of the galaxy clusters (Lotz et al. 2017, 2014; Koekemoer et al. 2014). The HFF program provided ultra-deep HST photometric observations of six massive clusters (see Fig. 4.4), in the seven WFC3 and ACS filters: F435W (ACS); F606W (ACS); F814W (ACS); F105W (WFC3); F125W (WFC3); F140W (WFC3);

[7] https://sites.google.com/site/vltclashpublic/home

[8] https://frontierfields.org

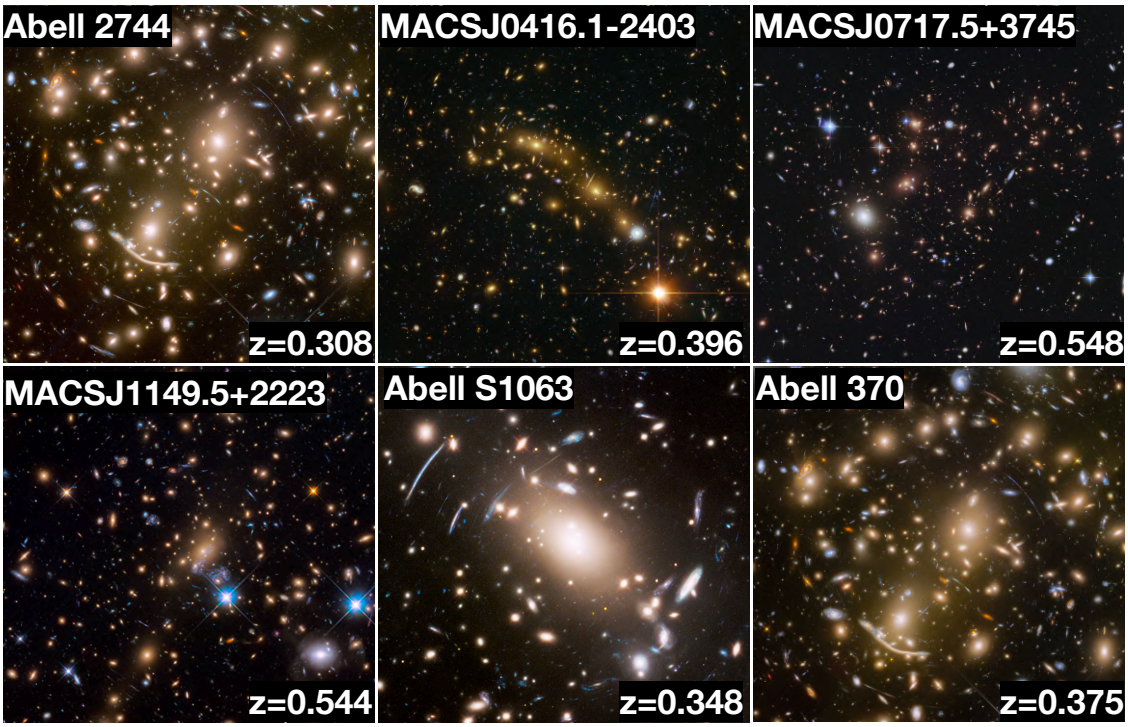


Figure 4.4: RGB cutouts of the core region of the six HFF clusters (Lotz et al. 2017). Clusters IDs and redshifts are written in white.

and F160W (WFC3). The six HFF clusters were selected for their strong lensing features, the absence of bright stars in the observed fields and for the possibility of complementary observations by other telescopes (in particular Spitzer and the James Webb Space Telescope). The whole program consists in 840 HST orbits (about 140 per cluster) including the parallel fields associated to each cluster. The HFF/HST images reach a 5σ point-source depths of $\sim$29th ABmag, i.e. about 1.5 magnitudes deeper than CLASH observations.

The main goals of the HFF program are the discovery and the study of high redshift galaxies lensed by the gravitational potential of the clusters, and the development of more accurate cluster lensing models.

Chapter 5

Cluster member kinematics from observed spectra

5.1 The galaxy spectral energy distribution

The light emission of a galaxy can be considered as the superposition of different types of stellar spectra from billions of moving stars. Since the shape of a stellar spectrum depends on the mass, the metallicity and the age of the star, the same quantities affect also the galactic spectrum. Moreover, we expect an evolution of the galaxy spectrum in time due to the evolution of the stars along the Hertzsprung-Russel diagram (HR diagram). For example, since the massive bluer stars leave their main sequence in few millions of years, a modification of the shape of the galaxy spectrum toward a redder color is expected.

To model a galaxy spectrum, a first assumption on the choice of the initial mass function (IMF), $\hat{\phi}(m)$, has to be made. The IMF defines the initial mass distribution of the stars in the galaxy and it is usually normalized such that:

$$\int_{m_L}^{m_U} \hat{\phi}(m)\, m\, dm = 1\,\mathrm{M}_\odot.$$

The integration limits range from $m_L \sim 0.1\,\mathrm{M}_\odot$, corresponding to the minimum mass to turn on the fusion processes, to $m_U \sim 100\,\mathrm{M}_\odot$, i.e. the mass of the heaviest observed stars.

A common choice for $\hat{\phi}(m)$ is the Salpeter IMF (Salpeter 1955) defined as:

$$\hat{\phi}_S(m) = m^{-2.35}. \tag{5.1}$$

This equation well describes the stars with $m \geq 1\,M_\odot$, while less massive stars are better characterized by a shallower Chabrier (Chabrier 2003) or Kroupa (Kroupa 2001) IMF.
A second fundamental ingredient necessary to model galaxy spectra is the star-formation rate (SFR). The SFR is defined as the fraction of gas mass that is converted into stars per unit time:

$$\psi(t) = -\frac{\mathrm{d}m_{gas}}{\mathrm{d}t}. \tag{5.2}$$

Since the amount of gas available to form new stars decreases with time, the SFR can be considered as a decreasing function. A common functional form for the SFR is the delayed exponential expressed by the following equation:

$$\psi(t) = \frac{t}{\tau^2} e^{-\frac{t-T}{\tau}} H(t-T), \tag{5.3}$$

where H is the Heaviside step function. For $t \geq T$ the Eq. 5.3 is different from zero, thus T corresponds to the reference time at which the star formation begins.
The second parameter in the Eq. 5.3, τ, defines a typical time-scale for the SFR. For $t \geq \tau$ the formation of new stars is strongly suppressed and the galaxy becomes sensibly redder.

A third physical quantity influencing the shape of galaxy spectra is the metallicity Z. Indeed, an higher metallicity produces redder stars with a larger M/L ratio, and a galaxy rich of metals is expected to hold a redder spectrum respect to a metal poor galaxy.
In general, the metallicity of a stars can be considered equal to the metallicity of the Inter-Stellar Medium (ISM) at the time of its formation, $Z(t)$. Since the stellar evolution processes create metals that are injected into the ISM by stellar winds, planetary nebulae and supernovae explosions, $Z(t)$ is an increasing function of the time.

Given an IMF and the function $Z(t)$, we can define a quantity $S_{\lambda,Z}(t)$ corresponding to the emitted energy, at the wavelength λ, by a population of stars with initial metallicity Z and age t. The $S_{\lambda,Z}(t)$ is usually normalized to one initial solar mass, and contains the physics of the stellar evolution inside the HR diagram. From the $S_{\lambda,Z}(t)$, we can compute the total flux emitted by a galaxy at a wavelength λ and time t by a convolution with the SFR:

$$F_\lambda(t) = \int_0^t \mathrm{d}t' \psi(t-t') S_{\lambda, Z(t-t')}(t'). \tag{5.4}$$

A primary feature in elliptical galaxy spectra is a prominent 4000Å break caused by the absorption of higher energy radiations from the metals (in various states of ionization) in the stellar atmospheres. Early-type galaxy spectra are also characterized by deep CaII H and K absorption lines (and high-order lines of the Balmer series), at the rest-frame wavelengths 3969Å and 3934Å respectively, typical of an old and methal-rich stellar population (Sánchez Almeida et al. 2012). Instead, the lack of strong emission lines is symptom of a low star formation.

5.2 Stellar kinematics from observed galaxy spectra

In the last section we showed how a galaxy spectrum can be obtained convolving the SFR with the spectral energy distribution of its stellar populations. However, the shape of the observed galaxy spectrum is also affected by the stellar kinematics. We begin by considering a star that is moving with the LoS component of its velocity given by v_{los}. A spectral feature that in the star rest frame corresponds to a wavelength λ_0 is shifted, due to the Doppler shift, to an observed wavelength $\lambda_{obs} = \lambda_0 + \Delta\lambda$, with:

$$\Delta\lambda = \left(\frac{v_{los}}{c}\right) \lambda_0. \tag{5.5}$$

Defining the spectral velocity $u = c \ln \lambda$ we can also write:

$$u_{obs} = c \ln \lambda_{obs} = c \ln \left[\lambda_0 \left(1 + \frac{v_{los}}{c}\right)\right] \sim c \ln \lambda_0 + v_{los}, \tag{5.6}$$

where we assume $v_{los} \ll c$. Thus, a spectral line observed at the spectral velocity u_{obs} is emitted, in the star rest-frame, at $u - v_{los} \sim c \ln \sigma_0$. The addition of the Doppler-shifted spectra of billions unresolved stars, results in a galaxy spectrum with shifted and broadened emission and absorption lines.

A useful quantity to describe the stellar kinematics of a galaxy is the Line-Of-Sight Velocity Distribution (LOSVD), $\mathcal{L}(v_{los})$. In particular, the fraction of stars with LoS velocities between v_{los} and $v_{los} + \mathrm{d}v_{los}$ is given by $\mathcal{L}(v_{los})\mathrm{d}v_{los}$.

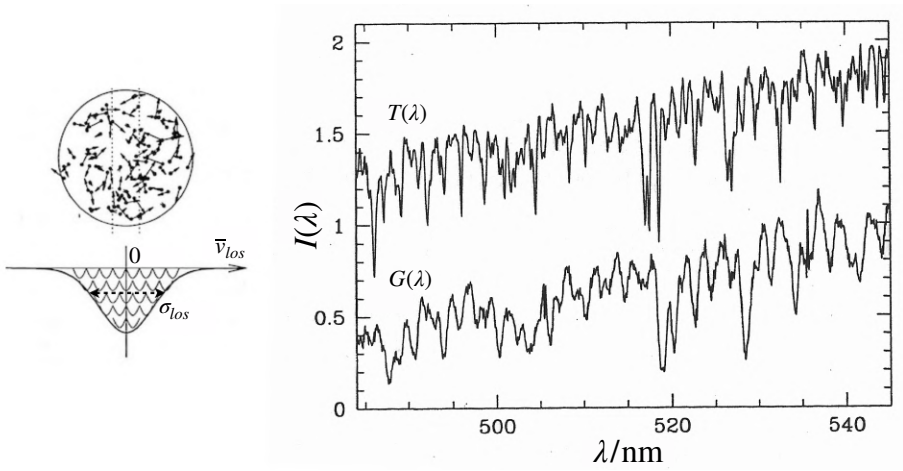


Figure 5.1: Left: Broadening of a spectral absorption line caused by the velocity dispersions of the stars in the galaxy. Right: Comparison between the spectra of a K0 giant star ($T(\lambda)$) and the lenticular galaxy NGC 2549 ($G(\lambda)$).

From the LOSVD we can easily compute the average LoS peculiar velocity of the galaxy, $\overline{v}_{los}$, and the average LoS velocity dispersion of its stars, σ^2_{los}:

$$\overline{v}_{los} = \int \mathrm{d}v_{los} v_{los} \mathcal{L}(v_{los}), \qquad \sigma^2_{los} = \int \mathrm{d}v_{los} \left(v_{los} - \overline{v}_{los}\right)^2 \mathcal{L}(v_{los}). \tag{5.7}$$

While the rigid shift of the galaxy spectral features depend on $\overline{v}_{los}$ (obviously we are ignoring every contribution from the cosmological redshift), the broadening of spectral lines is proportional to σ^2_{los} (see Fig. 5.1).

Considering, for simplicity, a galaxy made by a single spectral type of stars with spectral energy distribution $T(u)$ (known as stellar template), we can obtain the galaxy spectrum, $G(u)$, through the following convolution:

$$G(u) \propto \int \mathrm{d}v_{los}\, \mathcal{L}(v_{los})\, T\left(v_{los} - \overline{v}_{los}\right) = T * \mathcal{L}. \tag{5.8}$$

Conversely, we can measure the stellar LOSVD de-convolving the observed galaxy spectrum, $G(u)$, through the stellar template. Despite the de-convolution process is an intrinsically ill problem that amplifies the noise in observations, several possible solutions exist. In the next section we discuss the solution proposed by Cappellari & Emsellem (2004) and implemented in their Penalized Pixel Fitting (pPXF) code. The problem is even more complicated by the fact that real galaxies contain several spectral types of stars with different spectral templates. To overcome this problem pPXF gives the possibility to use combinations of stellar templates to de-convolve

the LOSVD. The selection of a reasonable library of stellar templates, matching the complete stellar population of the galaxy, is fundamental for a correct measurement of the LOSVD. Indeed, a wrong template selection can cause a systematic error in the LOSVD parameters optimization known as *templates mismatch*.

5.3 LOSVD fitting by pPXF

The pPXF code was developed by Cappellari & Emsellem 2004 to measure the stellar LOSVD of observed galaxy spectra. In this section, we briefly describe the methodology implemented into the software and its main input and out parameters. Most of the results obtained in the following chapters rely on the cluster member velocity dispersions measured by pPXF. For this reason, in the next section, we test the code performances using extensive spectra simulations.

pPXF assumes a Gauss-Hermit parameterization for the $\mathcal{L}(v_{los})$ given by:

$$\mathcal{L}(v) = \frac{e^{-\frac{y^2}{2}}}{\sigma\sqrt{2\pi}}\left[1 + \sum_{m=3}^{M} h_m H_m(y)\right], \quad \text{where} \quad y = \frac{v - V}{\sigma}. \tag{5.9}$$

In this equation, σ is the measured stellar velocity dispersion, V is the measured peculiar velocity of the galaxy, and $H_m(y)$ are Hermit polynomials weighted by the coefficients h_m. Obviously, if the galaxy spectrum, $G(u)$, is extracted inside an aperture, both the σ and the V are light-weighted average quantities within the aperture (see Eq. 2.32).

Given a set of parameters for the LOSVD in Eq. 5.9, pPXF creates a model for the galaxy spectrum through the following equation:

$$G_{model} = \sum_{n=1}^{N} w_n \left\{ [T_n(x) * \mathcal{L}_n(cx)] \sum_{k=1}^{K} a_k P_k(x) \right\} + \sum_{l=0}^{L} b_l P_l(x) + \sum_{j=1}^{J} c_j S_j(x). \tag{5.10}$$

Conversely to Eq. 5.8, now we are considering a library of stellar templates $\{T_n\}$. Every stellar template can be convoluted with a different LOSVD, $\mathcal{L}_n(cx)$, and the weights, w_n, quantify the contributions of different T_n to the final galaxy spectrum. The Eq. 5.10 contains also the following additional terms: the $P_k(x)$ are multiplicative Legendre polynomials used to compensate a non accurate spectral calibration and the reddening by dust; the $P_l(x)$ are additive polynomials useful to minimize the template mismatch, giving the possibility to change the intensity of

individual absorption lines, and to reduce the impact of a bad sky subtraction and of the scattered light; finally the $S_j(x)$ are sky spectra (see Cappellari 2017 for more details).

The best-fit parameters of the LOSVD (V, σ, h_m), the weights (w_n), and the constants (a_k, b_l, c_j) are determined minimizing a penalized χ^2 of the form (Cappellari & Emsellem 2004):

$$\chi^2_p = \chi^2 \left(1 + \hat{\lambda}^2 \mathcal{D}^2\right). \tag{5.11}$$

The standard $\chi^2 = \sum_{n=1}^{N} \frac{G_{mod(n_x)} - G(x_n)}{\Delta G(x_n)}$ quantifies the agreement between the model and the observed galaxy spectrum over the set of N spectral pixels. Instead, the penalty function, $\mathcal{D}$, is given by:

$$\mathcal{D}^2 = \frac{\int_{-\infty}^{+\infty} [\mathcal{L}(v) - \mathcal{G}(v)]^2 \, dv}{\int_{-\infty}^{+\infty} \mathcal{G}(v)^2 dv}, \tag{5.12}$$

where $\mathcal{G}$ is the best-fit LOSVD (see Eq. 5.9) but considering all the $h_m = 0$ (Gaussian shape). The penalty function acts biasing the LOSVD fit toward a Gaussian shape (suppressing all the h_m moments) when the measured velocity dispersion is smaller than the instrumental resolution and the signal-to-noise (hereafter S/N) is too low to constrain higher velocity moments. The impact of the penalty term is tuned by the bias coefficient $\hat{\lambda}$. The value of this coefficient depends, in most of the cases, by the mean signal-to-noise (hereafter $\langle S/N \rangle$) of the observed galaxy spectra.

Since pPXF fits the LOSVD parameters in the spectral pixel space, we can easily exclude the bad pixels of the spectra from the fitting procedure. These bad pixels are for example noisy regions of the observed galaxy spectra, regions of bad sky subtractions, or emission lines coming from contaminant sources.
The pPXF fit of the high redshift galaxy spectra has to be performed in the rest-frame wavelengths. Thus, we need to divide the wavelength axis of each observed galaxy by a factor $1+z_{gal}$ (where z_{gal} is the galaxy measured redshift). This procedure transform the first moment of the LOSVD, V, in a fine correction on z_{gal}. In particular the corrected galaxy redshift, z_{new}, is given by $1 + z_{new} = (1 + z_{gal}) \times (1 + V/c)$.

In the following chapters, we use the python 2018-version of pPXF (Cappellari 2017). In this latest release, the convolution between the LOSVD and the stellar templates is performed assuming an analytic Fourier transform for the LOSVD. If the $\langle S/N \rangle$ is sufficiently high, this methodology allows an accurate measurement of the stellar kinematics even at small velocity dispersions.

5.4 pPXF simulations

In this final section of the Chapter 5, we test the reliability of the pPXF kinematic measurements on a set of mock cluster member spectra designed to perfectly reproduce the MUSE/VLT spectroscopic observations of CLASH and HFF galaxy clusters at intermediate redshift ($0.2 < z < 0.5$). Indeed, the results presented in the following chapters critically depend on the robustness of cluster member velocity dispersion measurements.

The simulated spectra span a wide range of S/N and in each spectrum the S/N is generally a function of wavelengths to reproduce the MUSE spectrograph response. Here and in the following sections, we compute the mean signal-to-noise, $\langle S/N \rangle$, of the spectra over the analyzed wavelength range, from the residuals between the spectral data and the best-fit pPXF spectral model, after averaging over N_{bin} wavelength bins, which are $\sim$100Å wide. Thus,

$$\langle S/N \rangle = \sum_{i=1}^{N_{bin}} \frac{|data|_i}{\sqrt{\mathrm{Var}(residuals)_i}} \tag{5.13}$$

The simulations are used to determine:

- the minimum value of the average signal-to-noise, $\langle S/N \rangle_{min}$, that ensure robust velocity dispersion measurements.
- the lower limit of the measured velocity dispersions, σ_{min}, below which the MUSE spectral resolution can produce biases on LOSVD output parameters.
- the value of the penalty coefficient, $\hat{\lambda}$ in Eq. 5.11, as a function of the $\langle S/N \rangle$ of the spectra
- correction functions for the presence of systematic biases in the measured velocity dispersions. They include the bias induced by the penalty term in Eq. 5.11.
- the relation between the statistical errors measured by pPXF and the real errors committed in the velocity dispersion measurements.

To perform the pPXF measurements, we use a subset of 105 stellar templates of different spectral types, drawn from the National Optical Astronomy Observatory library (Valdes et al. 2004). To match the typical underlying stellar populations of early-type galaxies in the cluster cores, most of the templates are of G, K, M spectral

classes. In addition, we include 10 A-stars to extend spectral fits to a non-negligible fraction of E+A galaxies in our cluster sample and a few O and B stars. The stellar templates cover a wavelength range from 3465 Å to 9469 Å, with a sampling of 0.4 Å/pix and have an intrinsic resolution of 1.35 Å full width at half-maximum. The same collection of stellar templates will be use also in the following chapters.

5.4.1 The simulated sample of cluster galaxies

The simulated cluster member sample consists in 10000 spectra created according to the following pipeline:

1) Synthetic cluster member spectra from linear combination of stellar templates

The staring point of our simulations are the stellar templates. In particular we want to determine the best stellar template combinations that faithfully reproduce the stellar composition of observed cluster member galaxies. This can be done with a first run of pPXF over a sample of real cluster member spectra. In our case this sample contains all the galaxies of the CLASH cluster MACS J1206.2−0847 within the MUSE pointings and with $m_{F160W} < 24$ and $\langle S/N \rangle > 20$. The galaxy spectra were extracted inside apertures of $R = 0.8''$ from the MUSE datacube. There are not particular reasons to use MACS J1206.2−0847 instead of another cluster. Every sample of galaxy spectra with a stellar population similar to those of the galaxies that we want to study may in principle works.
The first pPXF run is used to determine the weights, w_n, in the Eq. 5.10 that better reproduce the stellar composition of the observed galaxies (a zero degree multiplicative Legendre polynomial is assumed in this run). The stellar template linear combinations obtained using these measured w_n correspond to our *bare* simulated galaxy spectra. In this context, *bare* means rest-frame galaxy templates not yet convoluted with a LOSVD and without noise.
Since the bare galaxy spectra are linear combinations of stellar templates they maintain a wavelength range from 3465 Å to 9469 Å, a dispersion of 0.4 Å/pix and a spectral resolution of 1.35 Å Full Width at Half Maximum (FWHM).
Our simulations span the same spectral type diversity and stellar composition of the observed cluster members. Sometimes both the observed and simulated galaxies contain younger stellar populations with significant Balmer lines (e.g. E+A galaxies).

2) Redshift and LOSVD

In this second step, we convolve every bare galaxy spectrum with a stellar LOSVD. The LOSVD has the form of Eq. 5.9 and we consider the Hermit polynomials up to the 4^{th} moment (h_4).
The final sample of simulated galaxies is divided into six subsets that differ for the ranges in which the input velocity dispersions (σ_{in}) and $\langle S/N \rangle$ are uniformly distributed. The chosen ranges in each subset are in Table 5.1. The first (V_{in}), third (h_3^{in}) and fourth (h_4^{in}) moments of the LOSVD are uniformly distributed within the intervals $(-50.0, 50.0)$ km s^{-1}, $(-0.1, 0.1)$ and $(-0.1, 0.1)$ respectively, in all the subsets.

N	$\boldsymbol{\sigma_{in}}$ **[km s^{-1}]**	$\boldsymbol{\langle S/N \rangle}$
1000	0.10 - 250.0	1.0 - 100.0
1000	10.0 - 150.0	1.0 - 30.0
1000	10.0 - 150.0	1.0 - 15.0
2000	45.0 - 250.0	1.0 - 100.0
2000	45.0 - 250.0	1.0 - 20.0
3000	50.0 - 250.0	1.0 - 15.0

Table 5.1: Number of simulated spectra of cluster members in six bins of velocity dispersion, σ_{in}, and $\langle S/N \rangle$. In each subset, LOSVD moments (V_{in}, h_3^{in}, h_4^{in}) are randomly distributed (see text).

The simulated spectra are then redshifted such to reproduce the observed redshift distribution of MACS J1206.2−0847 cluster members.

3) Simulation of MUSE observations

In this final step of the simulation pipeline, the mock galaxy spectra are reshaped to mimic MUSE observations. Thus, we reduce their wavelength range to the MUSE spectral range, 4650-9300 Å, and we degrade their spectral resolution to the lower MUSE instrumental resolution of 2.6 Å FWHM approximately constant over the whole considered wavelength range (see Chapter 4.1.1). Then, the mock spectra are re-sampled to the MUSE 1.25 Å/pix scale.
Finally, a Gaussian noise, drawn from real MUSE variance spectra, is added to

the mock spectra in such a way to reproduce a $\langle S/N \rangle$ between ~ 1 and ~ 100 (see Table 5.1). Thus, the high variance due to sky subtraction around sky lines is included in the simulations.

5.4.2 Simulation results

Using pPXF we measure the LOSVD parameters of the mock spectra (see Fig. 5.6). Then, these measurements are compared to the simulation inputs to test the robustness of the pPXF fit.
The LOSVD in Eq. 5.9 is considered up to the forth order, thus the fitted LOSVD moments are the peculiar velocity V, the velocity dispersion σ and the two coefficients of the Hermit polynomials h_3 and h_4. To perform the pPXF fit, we consider also a fourth degree additive Legendre polynomial, $P_l(x)$, (see Eq. 5.10) and a zero degree multiplicative polynomial, $P_k(x)$).
Finally, we assume the following expression for the bias parameter $\hat{\lambda}$ (see Eq. 5.11):

$$\hat{\lambda} = 0.15 + 0.0107 \langle S/N \rangle - 0.00004 \langle S/N \rangle^2 , \tag{5.14}$$

where $\langle S/N \rangle$ is the mean measured signal-to-noise of the spectra (see Eq. 5.13). This relation was empirically found by Cappellari et al. (2011) for SAURON integral field spectrograph observations, with spectral resolution of 4.2Å FWHM and a wavelength range of 4800-5300Å.
The same pPXF parameter configuration will be adopted in the following chapters to measure the velocity dispersions of the observed cluster member galaxies.

The mock galaxy spectra are fitted in the rest-frame wavelength range 3600-4900Å. Indeed, for redshifts close to $z = 0.439$ (i.e. the MACS J1206.2−0847 mean redshift), this represents the optimal wavelength range to exclude regions of low $\langle S/N \rangle$ due to the MUSE sensitivity curve, particularly on the red side of the spectrum strongly affected by sky lines residuals.

As first result of the simulations, we show in Fig. 5.2 the variation of the third and the fourth momenta of the LOSVD (h_3, h_4) as function of the input velocity dispersion. In this figure we consider only the galaxies with $\langle S/N \rangle > 60$. The black solid curve in the plots join the medians of the galaxy distributions inside eight bins of input velocity dispersion, while the dark-gray areas are bounded by the 16-th and 84-th percentiles computed inside the same eight bins. The vertical blue lines

correspond to a value of $\sigma_{in} = 80\,\mathrm{km\,s^{-1}}$, i.e. the average MUSE spectral resolution (2.6Å FWHM) in the selected wavelength range (see Section 4.1.1).
For velocity dispersions smaller than the spectral resolution, the penalty function correctly acts biasing the h_3 and h_4 values towards zero to keep the noise on the measured V and σ under control (see Section 5.3 and Cappellari et al. 2011). Repeating this analysis at different $\langle S/N \rangle$ we proved that the expression in Eq. 5.14 can be adopted also for our MUSE observations.

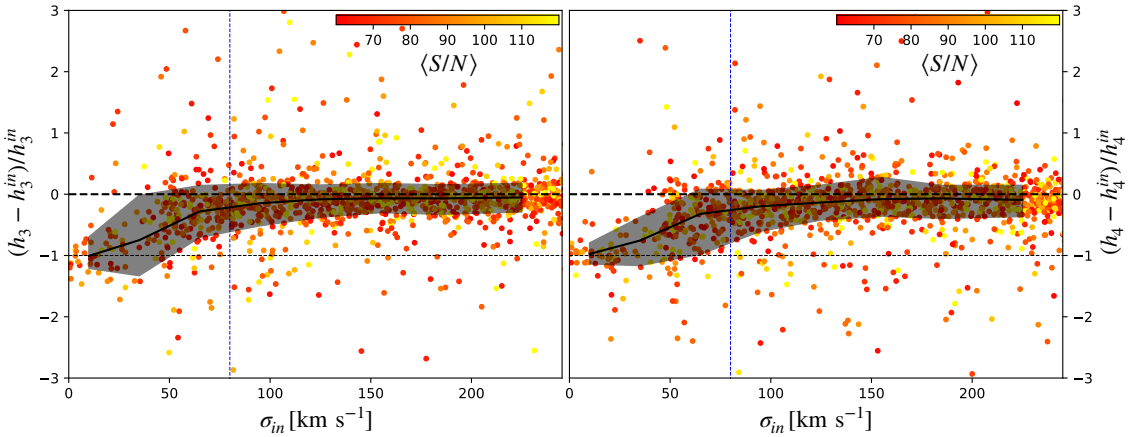


Figure 5.2: *Left:* Normalized difference between the pPXF measured h_3 and the input h_3^{in} in simulated spectra, as a function of the input velocity dispersions, σ_{in}. Only the galaxies with $\langle S/N \rangle > 60$ are shown in the plot (colored dots) and they are color-coded according to their $\langle S/N \rangle$. The black solid curve joins the medians of the mock galaxy distributions inside eight bins of input velocity dispersion. Instead, the gray shaded area is bounded by the 16^{th} and 84^{th} percentiles computed inside the same eight bins. The vertical blue line corresponds to a $\sigma = 80$ km s^{-1}, i.e. the mean MUSE spectral resolution within the wavelength interval 3600-4900Å. *Right:* Same as the left panel but for h_4.

The two panels in Fig. 5.3 show the variation of the differences $V - V_{in}$ and $\sigma - \sigma_{in}$, as a function of the measured velocity dispersion. As in Fig. 5.2, only the galaxies with $\langle S/N \rangle > 60$ are considered. While the first momentum of the LOSVD, V, is well fitted by pPXF at any given σ, we notice a systematic overestimation of the measured velocity dispersion for σ approaching (or below) the MUSE instrumental resolution. This mismatch between the σ and σ_{in} values is largely due to an inverse degeneracy, between the σ and the h_4 parameters, into the χ^2_p expression (see Cappellari & Emsellem 2004). Thus, the suppression of the h_4 values operated by the penalty function $\mathcal{D}$ (see Eq. 5.12), reflects in an overestimation of the measured velocity dispersions at $\sigma < 80\,\mathrm{km\,s^{-1}}$.
Using a polynomial function to fit the median values of the $\sigma - \sigma_{in}$ distributions in

seven σ bins (black solid line in Fig. 5.3), we find the following correction function for the measured velocity dispersions:

$$\sigma_{true} = \sigma - (4.00 \cdot 10^{-5}\sigma^2 - 2.59 \cdot 10^{-2}\sigma + 4.06)\,\mathrm{km\,s^{-1}}. \tag{5.15}$$

The increase in the $\sigma - \sigma_{in}$ scatter at low $\langle S/N \rangle$ makes it difficult to detect such a small bias at lower $\langle S/N \rangle$.
In any case, to avoid the possibility of systematic errors in the pPXF fits, in the following chapters we measure only the velocity dispersion of the galaxies with $\sigma > 80\,\mathrm{km\,s^{-1}}$(i.e. the mean MUSE spectral resolution in the selected wavelength range).

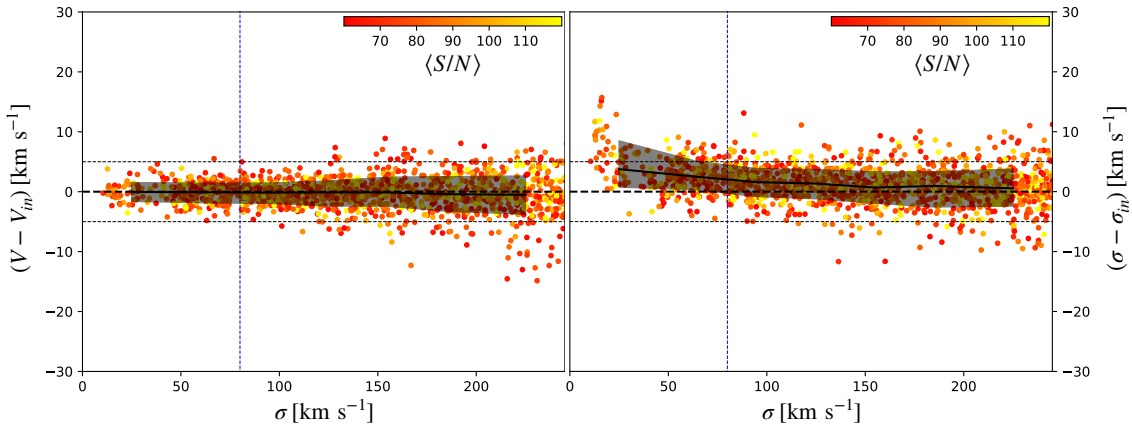


Figure 5.3: *Left:* Difference between the pPXF measured V and the input values V_{in}, as a function of the measured velocity dispersion, σ. Only the mock galaxy spectra with $\langle S/N \rangle > 60$ are considered (dots in the plot). The dots are color-coded according to the $\langle S/N \rangle$ values. The black dashed horizontal lines correspond to velocity differences of -5, 0 and 5 $\mathrm{km\,s^{-1}}$ from the bottom to the top. The black solid curve joins the medians of the galaxy distributions inside seven bins of input velocity dispersion. Instead, the gray shaded area is bounded by the 16^{th} and 84^{th} percentiles computed inside the same seven bins. The vertical dashed blue line correspond to a $\sigma = 80$ $\mathrm{km\,s^{-1}}$, i.e. the mean MUSE spectral resolution within the wavelength interval 3600-4900Å. *Right:* Same as the left panel but for the measured velocity dispersions. (Figure from Bergamini et al. 2019)

To quantify the impact of the $\langle S/N \rangle$ on the pPXF measured V and σ values, we realize the plots in Fig. 5.4. In this figure the differences $V - V_{in}$ and $\sigma - \sigma_{in}$ are plotted against the $\langle S/N \rangle$ of the mock spectra. Similarly to Figs. 5.2 and 5.3, the black solid curves and the dark-grey areas correspond to the medians and the 16^{th} and 84^{th} percentiles of the galaxy distributions but now inside seven bins of $\langle S/N \rangle$. Regardless of the input V_{in} and σ_{in} values (the dot colors in the figure), for

$\langle S/N \rangle > 20$ we observe a scatter, in the $V - V_{in}$ and $\sigma - \sigma_{in}$ differences, almost always smaller than 15 $\mathrm{km\,s^{-1}}$ (horizontal black dashed lines in the plots). Instead, a steep increase of the scatter is visible at $\langle S/N \rangle < 20$ since the measured velocity values become more sensitive to the presence of sky lines in the spectra. In particular, at $\langle S/N \rangle < 10$ the measured velocities became totally unreliable. For this reason, in the following chapters we take $\langle S/N \rangle = 10$ as lower limit for robust velocity dispersion measurements.

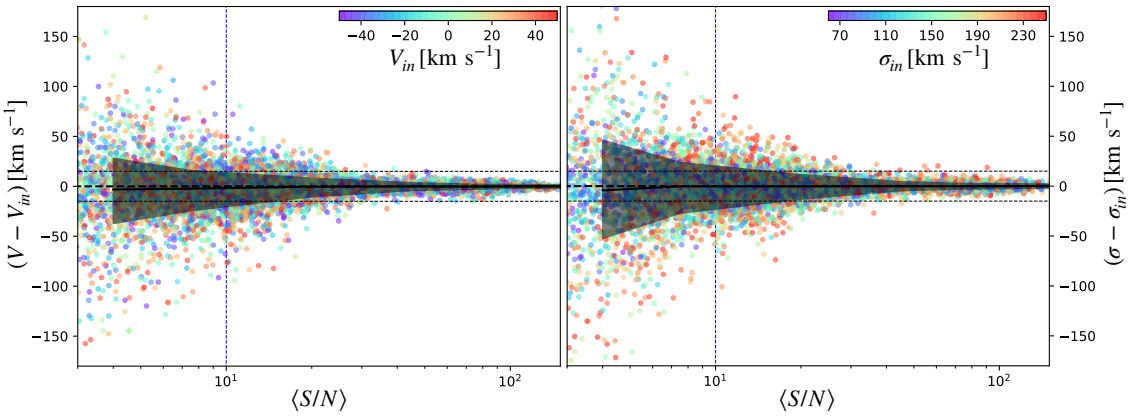


Figure 5.4: *Left:* Difference between the pPXF measured V and the input values V_{in}, as a function of the average signal-to-noise, $\langle S/N \rangle$, of the mock galaxy spectra (the dots in the plot). The dots are color-coded according to the V_{in} values. The black dashed horizontal lines correspond to velocity differences of -15, 0 and 15 $\mathrm{km\,s^{-1}}$ from the bottom to the top. The black solid curve joins the medians of the galaxy distributions inside seven bins of $\langle S/N \rangle$. Instead, the gray shaded area is bounded by the 16^{th} and 84^{th} percentiles computed inside the same seven bins. The vertical dashed blue line correspond to a $\langle S/N \rangle = 10$, i.e. the lower limit that we choose for reliable velocity measurements. *Right:* Same as left panel but for the velocity dispersion. In this case the dots are color-coded according to the σ_{in} values and only the simulated galaxies with $\sigma_{in} > 60$ are plotted.

Finally, we use the simulated galaxy spectra to derive realistic statistical errors on the pPXF measured velocity dispersions. The black solid line in Fig. 5.5 shows the variation of the statistical errors estimated from the simulations as a function of the $\langle S/N \rangle$. To determine these errors we compute the 68^{th} percentiles of the $|\sigma - \sigma_{in}|/\sigma_{in}$ distributions inside five bins of $\langle S/N \rangle$ (vertical black dashed lines in the plot). Instead, the dashed blue line in the plot shows the variation of the pPXF fit errors ($\delta\sigma$) at different $\langle S/N \rangle$. This line is obtained computing the median values of the $\delta\sigma/\sigma$ distributions inside the same five bins.
Performing a polynomial fit of the ratio of the two curves we find a correction function to the errors retrieved by pPXF (red solid line in the plot). In particular,

the relative errors on the velocity dispersions estimated using the simulations, are systematically underestimated by the pPXF fit errors by the $\sim$20% for $\langle S/N \rangle \gg 15$, up to the $\sim$25% for $\langle S/N \rangle \sim 10$.

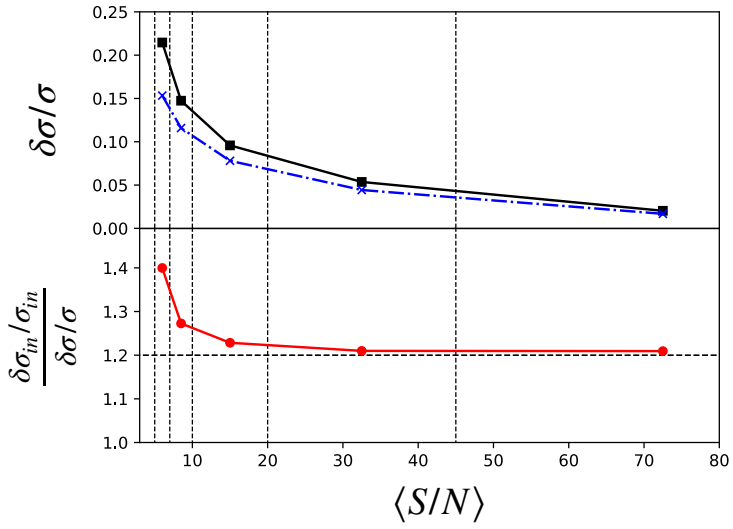
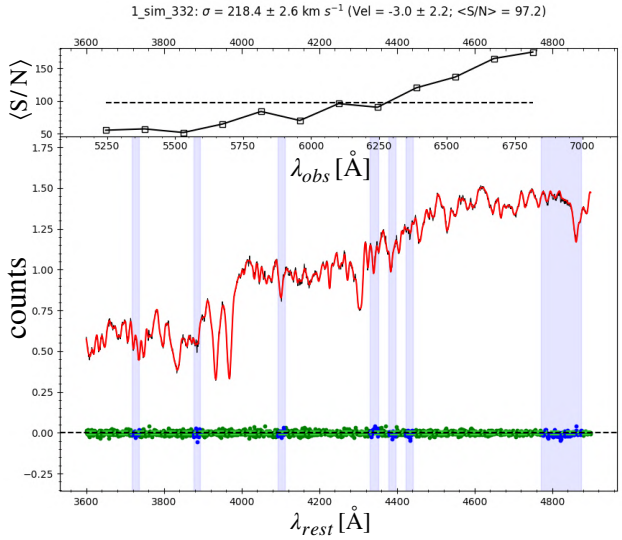
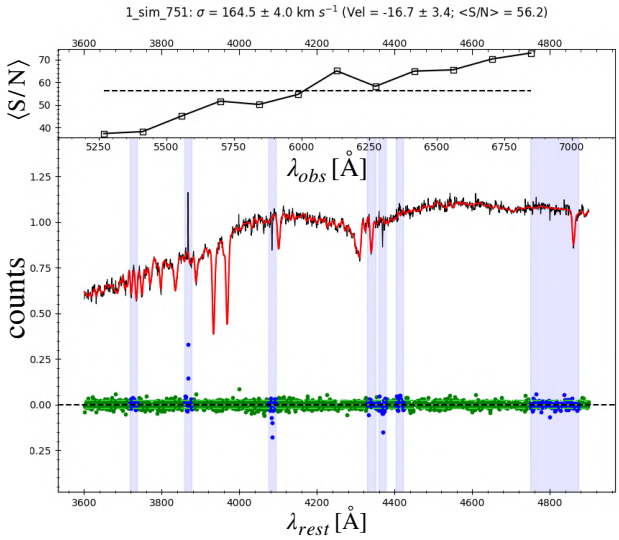
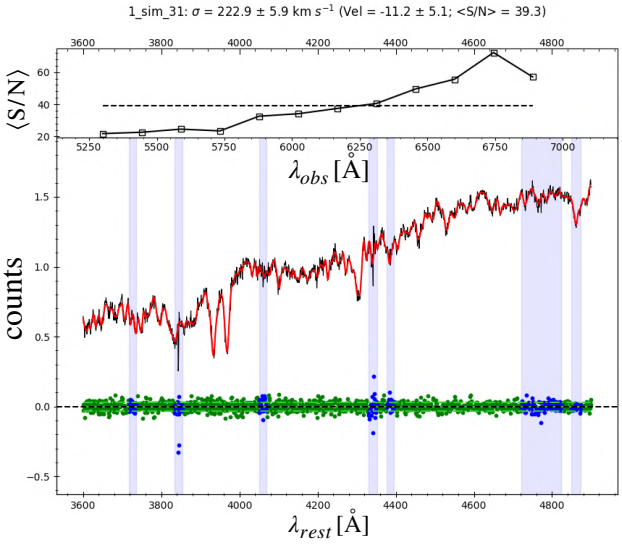


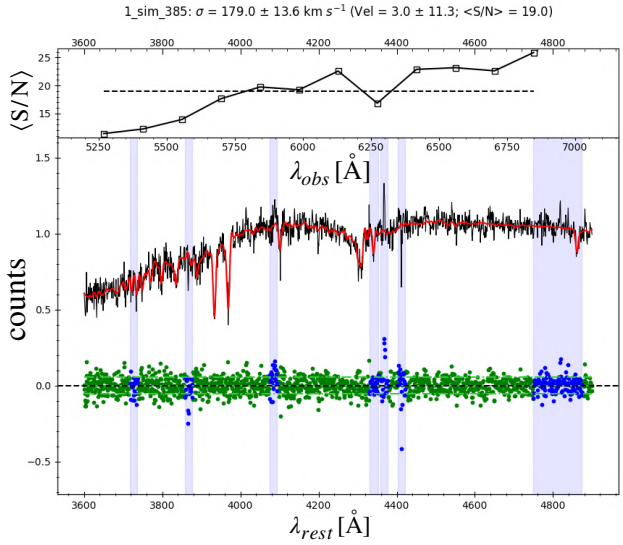
Figure 5.5: The black solid (blue dashed) line shows the relative statistical error estimated from simulations (provided by pPXF), as a function of mean signal-to-noise $\langle S/N \rangle$. The red curve corresponds to the ratio between the two; the statistical relative errors appear to be underestimated, respect to the pPXF values, by $\sim$ 20% in our spectra. (Figure from Bergamini et al. 2019)

In the following chapters, we apply all the corrective functions found in this section to obtain robust measurements of the cluster member velocity dispersions and their associated statistical errors. Moreover, we assume the two values $\langle S/N \rangle = 10$ and $\sigma = 80\,\mathrm{km\,s^{-1}}$ as lower limits to have reliable pPXF fits.
The current version of our simulations does not address directly the impact of template mismatch issue suffered by stellar velocity dispersion measurements. However, the possibility of pPXF software to fit a wide range of stellar templates (105 stellar templates selected to properly describe the stellar composition of cluster member galaxies are used in our study, see Sec. 5.4) together with the galaxy stellar kinematics is expected to strongly reduce the template mismatch problem (Cappellari & Emsellem 2004). To precisely quantify the impact of template mismatch in kinematic measurements, specific simulations will be developed in the future.


1_sim_332: σ = 218.4 ± 2.6 km s^{-1} (Vel = -3.0 ± 2.2; <S/N> = 97.2)
⟨S/N⟩
counts
λ_{obs} [Å]
λ_{rest} [Å]


1_sim_751: σ = 164.5 ± 4.0 km s^{-1} (Vel = -16.7 ± 3.4; <S/N> = 56.2)
⟨S/N⟩
counts
λ_{obs} [Å]
λ_{rest} [Å]


1_sim_31: σ = 222.9 ± 5.9 km s^{-1} (Vel = -11.2 ± 5.1; <S/N> = 39.3)
⟨S/N⟩
counts
λ_{obs} [Å]
λ_{rest} [Å]


1_sim_385: σ = 179.0 ± 13.6 km s^{-1} (Vel = 3.0 ± 11.3; <S/N> = 19.0)
⟨S/N⟩
counts
λ_{obs} [Å]
λ_{rest} [Å]

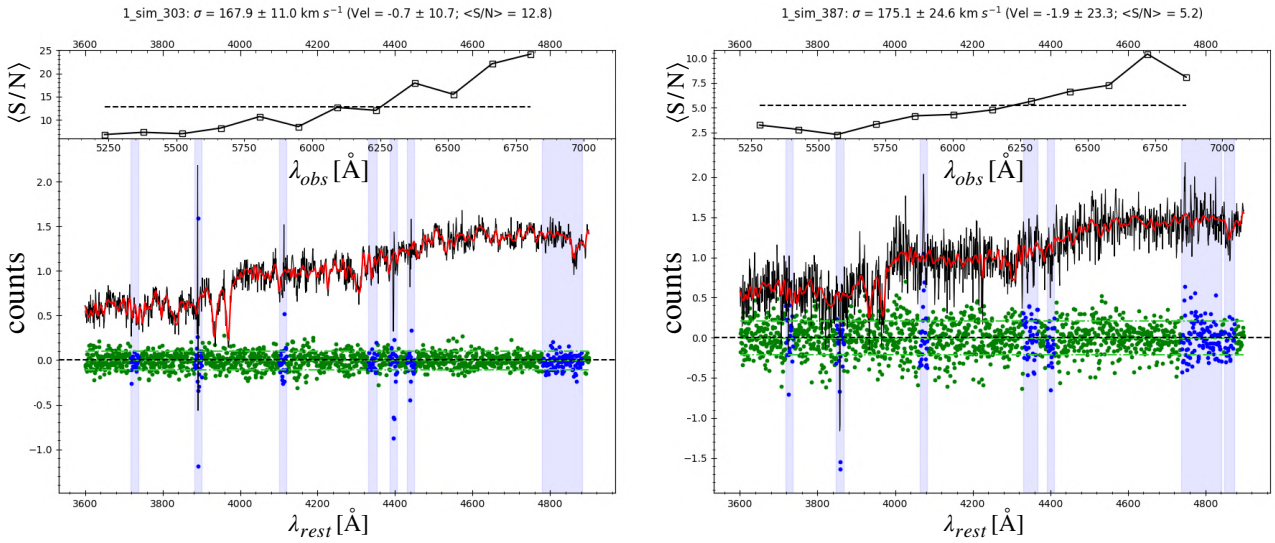


Figure 5.6: pPXF fit of six mock galaxy spectra (solid black lines in the plots) with decreasing values of $\langle S/N \rangle$ from the top left to the bottom right corners. The red solid lines are the pPXF best-fit models, while the green points correspond to the data-model residuals. The blue shaded regions along the wavelength axes were excluded in the fitting procedure due to the presence of sky lines or emission lines from the galaxies. The corresponding residuals are in blue. In the titles we report the ID of the galaxies, the measured velocity dispersion σ, the first moment V (Vel in the plots) and the $\langle S/N \rangle$. In the top panel of each plot we show the variation of the $\langle S/N \rangle$ as a function of the wavelength. In this case the $\langle S/N \rangle$ is computed within twelve bins whose central values are the squares in the plots.

Chapter 6

Enhanced cluster lensing models with measured galaxy kinematics

In this chapter we present an improved determination of the total mass distribution of three CLASH/Hubble Frontier Fields massive clusters, MACS J1206.2−0847 ($z = 0.44$), MACS J0416.1−0403 ($z = 0.40$), Abell S1063 ($z = 0.35$). We specifically reconstruct the sub-halo mass component with robust stellar kinematics information of cluster galaxies, in combination with precise strong lensing models based on large samples of spectroscopically identified multiple images. We use VLT/MUSE integral-field spectroscopy in the cluster cores to measure the stellar velocity dispersion, σ, of 40-60 member galaxies per cluster, covering 4-5 magnitudes to $m_{F160W} \simeq 21.5$. In Section 5.4 we verified the robustness and quantify the accuracy of the velocity dispersion measurements with extensive spectral simulations, thus determining the limiting acceptable signal-to-noise ($\langle S/N \rangle > 10$) and minimum velocity dispersion ($\sigma > 80$ km s^{-1}) for the depth of the spectroscopic data presented in this chapter. With these data, we determine the normalization and slope of the galaxy L-σ Faber-Jackson relation in each cluster and use these parameters as a prior for the scaling relations of the sub-halo population in the mass distribution modeling (see Section 3.7). When compared to our previous lens models, the inclusion of member galaxies' kinematics provides a similar precision in reproducing the positions of the multiple images. However, the inherent degeneracy between the central effective velocity dispersion, σ_0, and truncation radius, r_{cut}, of sub-halos is strongly reduced, thus significantly alleviating possible systematics in the measurements of sub-halo masses (see Section 3.7.1). The three independent determinations of the σ_0-r_{cut} scaling relation in each cluster are found to be fully consistent, enabling a statistical determination of sub-halo sizes as a function of σ_0, or halo masses. Finally, we derive the galaxy central velocity dispersion functions of the three

clusters projected within 16% of their virial radius, finding that they are well in agreement with each other. We argue that such a methodology, when applied to high-quality kinematics and strong lensing data, allows the sub-halo mass functions to be determined and compared with those obtained from cosmological simulations (see Section 8.3.4).

This chapter is organized as follows. In Section 6.1, we describe our imaging and spectroscopic data sets. In Section 6.2, we detail how internal velocity dispersions of member galaxies are measured. Strong lensing models for the three clusters under study are described in Section 6.3, while the specific methodology to incorporate galaxy kinematics information into our lens models is discussed in Section 6.4. Results are discussed in Section 6.5, where we also present the velocity dispersion functions for the three clusters. In Section 6.6, we summarize the main conclusions of the chapter.

Throughout this and the following chapters, we adopt a flat ΛCDM cosmology with $\Omega_m = 0.3$ and $H_0 = 70\,km\,s^{-1}Mpc^{-1}$. With these parameters, $1''$ corresponds to a physical scale of 5.68, 5.34 and 4.92 kpc at $z = 0.439, 0.396, 0.348$, respectively the redshift of the three clusters: MACS J1206.2−0847, MACS J0416.1−0403 and Abell S1063 (see below). All magnitudes refer to the AB system.

6.1 Photometric and spectroscopic data

This section summarizes the photometric and spectroscopic data sets for the three galaxy clusters used in this chapter, namely MACS J1206.2−0847, MACS J0416.1−2403, and Abell S1063 (a.k.a. RXJ 2248.7−4431), hereafter MACSJ1206, MACSJ0416 and AS1063, at redshifts 0.439, 0.396 and 0.348, respectively (see Table 6.1). These clusters were observed with HST in 16 broad band filters, from UV to near-IR, as part of the CLASH program (see Section 4.2.1). HST imaging of MACSJ0416 and AS1063 was significantly augmented with the HFF program (see Section 4.2.3), by adding deep exposures in seven filters (F435, F606W, F814W, F105W, F125W, F140W, F160W). The three clusters were also part of an extensive spectroscopic campaign with the CLASH-VLT Large Programme (P.I. P. Rosati), using the VIMOS high-multiplexing spectrograph, which provided over 4000 redshifts in each of the three clusters, over an area of $\sim 25 \times 25\,arcmin^2$ (see Section 4.2.2). These data sets yielded approximately 600 spectroscopic members for MACSJ1206 (Biviano et al. 2013, Girardi et al. 2015), 900 members for MACSJ0416 (Balestra et al. 2016) and over 1200 members for AS1063 (Mercurio et al. 2019, in prep.). Spectroscopic information in the cores of the three clusters has been

significantly enhanced with the MUSE integral field spectrograph at the VLT, which is at the basis of the kinematic measurements presented in this chapter (see Fig. 6.1 and Section 4.1.1). These data, which are described in more detail below, have enabled new high-precision strong lensing models based on large samples of multiply lensed sources (see Caminha et al. 2017a, Caminha et al. 2017b, Caminha et al. 2016 and Sec. 6.3). As we saw in Section 4.1.1, MUSE has a field of view of 1 arcmin^2, a spatial sampling of 0.2″, a spectral resolution of $\sim$2.6 Å FWHM over the spectral range 4650 - 9300 Å, with a spectral sampling of 1.25 Å/pix (see Fig. 4.2).

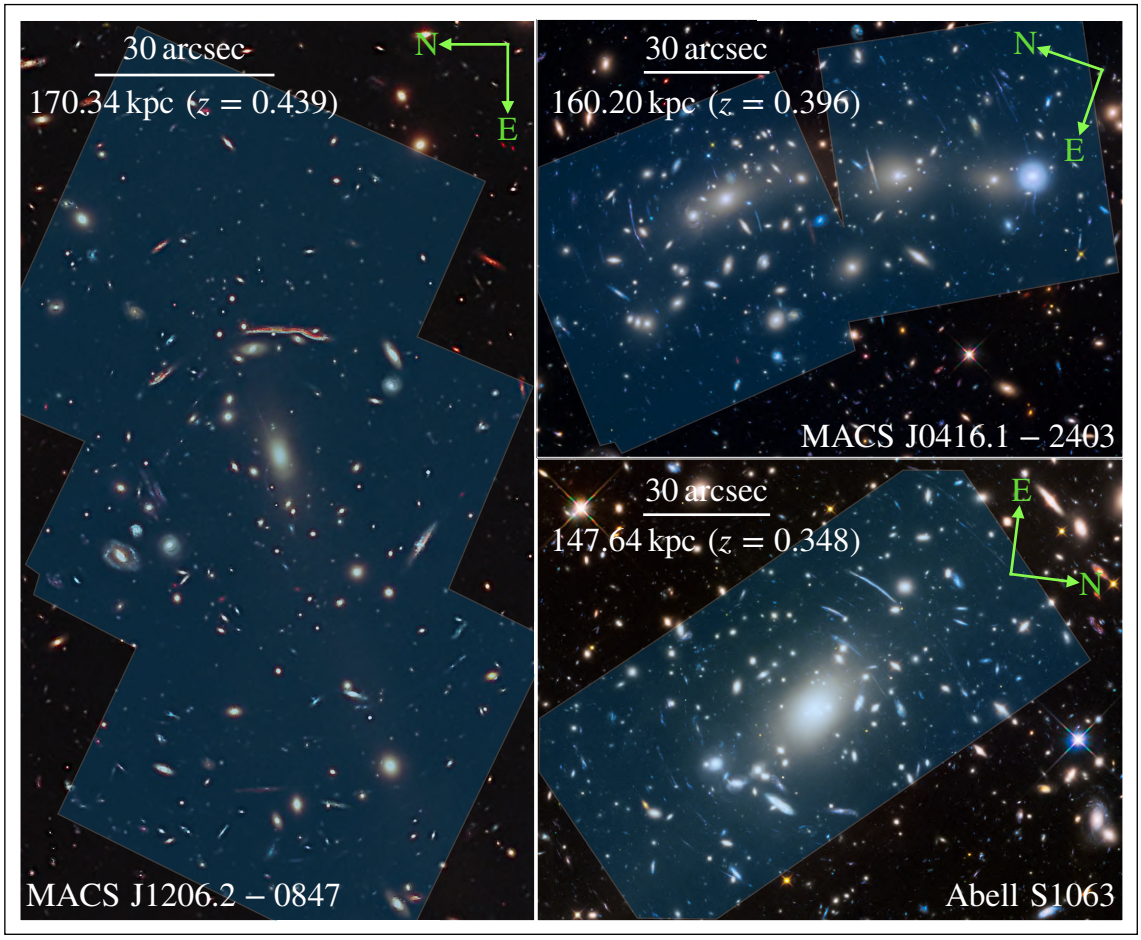


Figure 6.1: RGB cutouts of the three clusters MACSJ1206 (on the left), MACSJ0416 (top right corner) and AS1063 (bottom right corner). The MUSE pointings over the clusters are shown as light-blue regions.

MACSJ1206: MUSE data were obtained between 2015 and 2016[1], the redshift

[1]ID 095.A-0181(A) and 097.A-0269(A) (P.I. J. Richard)

measurements of the member galaxies and a large sample of multiple images were presented in Caminha et al. (2017b) (hereafter in this chapter C17b). Three MUSE pointings cover a total area of 2.63 arcmin^2, mapping the SE–NW elongation of the cluster (see Figure 6.1). The exposure time is 8.5 hours in the central ~ 0.5 arcmin^2 and 4 hours in the remaining area.

MACSJ0416: MUSE archival observations used in this work were presented in Caminha et al. (2017a) (hereafter in this chapter C17a), along with the redshift catalog, and consist of two pointings[2], one of 2 hours in the NE region, with a seeing of 0.5″, and a SW pointing of 11 hours, with a seeing of 1″ (see Figure 6.1).

AS1063: MUSE data consist of two pointings, which were presented in Karman et al. (2015) and Karman et al. (2017). The SW pointing[3] has an exposure of 3.1 hours and seeing $\sim 1.1''$, the NE pointing[4] has an exposure of 4.8 hours and seeing of 0.9″ (see Figure 6.1).

6.2 Cluster members: spectral extraction and internal kinematics

In this section, we describe the methodology adopted to extract the spectra of the cluster members from the MUSE data-cubes and to measure their internal stellar velocity dispersions.

Catalogues of cluster members for MACSJ1206, MACSJ0416, AS1063 were presented in C17b, C17a and Caminha et al. (2016) (hereafter in this chapter C16), respectively, with the main objective of identifying the sub-halos to be included in the lens models (see below). Cluster members were defined as galaxies in the redshift intervals 0.425 - 0.453, 0.382 - 0.410 and 0.335 - 0.362 for the three clusters respectively, lying within a rest-frame velocity of approximately ± 3000 km s^{-1}around the median redshift of the cluster. With 114 - 145 spectroscopic members per cluster in the HST FoV, the extensive CLASH multi-band information was used to obtain highly ($\sim$95%) complete and pure samples of photometric members down to $m_{\rm F160W} = 24$, following the method described in Grillo et al. (2015).

[2]ID 094.A-0115B (PI J. Richard) and 094.A0525(A) (PI F.E. Bauer)

[3]ID 60.A-9345 (P.I.: K. Caputi & C. Grillo)

[4]ID 095.A-0653 (P.I. K. Caputi)

The spectra of cluster members are extracted from the MUSE data-cubes within apertures of $R_{ap} = 0.8''$ radius, which is found to be a good compromise between the signal-to-noise of the spectra and the contamination from other sources in the field (either interlopers along the line-of-sight or close members). All the extractions are visually inspected to assess possible contamination from nearby sources and the apertures are reduced in specific cases down to $0.6''$. The spectra in which the contamination from bright nearby members is too strong are discarded, as the velocity dispersion of the fainter galaxy is likely biased.

Cluster	z	N_m^{meas} (N_m^{tot})	N_{im} (N_{fam})	$M_{200c}[10^{15} M_\odot]$	R_{200c}[Mpc]	$N_m(< 0.16\,R_{200c})$
MACS J1206.2−0847	0.439	58 (258)	82 (27)	(1.59 ± 0.36)	(2.06 ± 0.16)	179
MACS J0416.1−0403	0.396	49 (193)	102 (37)	(1.04 ± 0.22)	(1.82 ± 0.13)	124
Abell S1063	0.348	37 (222)	55 (20)	(2.03 ± 0.67)	(2.32 ± 0.26)	199

Table 6.1: Most relevant parameters of the three clusters of studied in this chapter. Redshift (z), number of cluster members with measured velocity dispersion (N_m^{meas}), total number of cluster members within the HST field with $m_{F160W} < 24$ included in the lens models (N_m^{tot}), number of spectroscopically confirmed multiple images (N_{im}), number of image families (N_{fam}), M_{200c} and R_{200c} values from Umetsu et al. (2014), and number of cluster members within a radius of $0.16\,R_{200c}$ ($N_m(< 0.16\,R_{200c})$). (Table from Bergamini et al. 2019)

We measure the stellar LOSVD of cluster members using the public software `pPXF` described in Section 5.3 (Cappellari & Emsellem 2004), with the latest improvements included in the 02/2018 python version (see Chapter 5 and Cappellari 2017). With `pPXF`, we determine the best-fit LOSVD parameters by performing a cross-correlation of the observed spectrum with the set of spectral templates described in Section 5.4. The best-fit is obtained by minimizing the penalized χ^2 in Eq. 5.11 between the template and the observed spectrum. Moreover, we measure the mean signal-to-noise, $\langle S/N \rangle$, of each spectrum over the selected wavelength range as in Eq. 5.13. For the `pPXF` bias parameter we adopt the relation between λ and $\langle S/N \rangle$ in Eq. 5.14 tested with the extensive spectral simulations developed in Section 5.4 (see also Cappellari et al. 2011).

The simulations in Section 5.4 reproduce our galaxy spectra with varying input LOSVD parameters, redshift and $\langle S/N \rangle$, to quantify the accuracy and precision of `pPXF` in measuring the velocity dispersions σ of member galaxies, thus optimizing `pPXF` input parameters. Simulated spectra are constructed from model spectra spanning the spectral type diversity of the galaxy populations of the three cluster studied in this chapter and by adding noise drawn from the variance map of the reduced MUSE data-cubes.

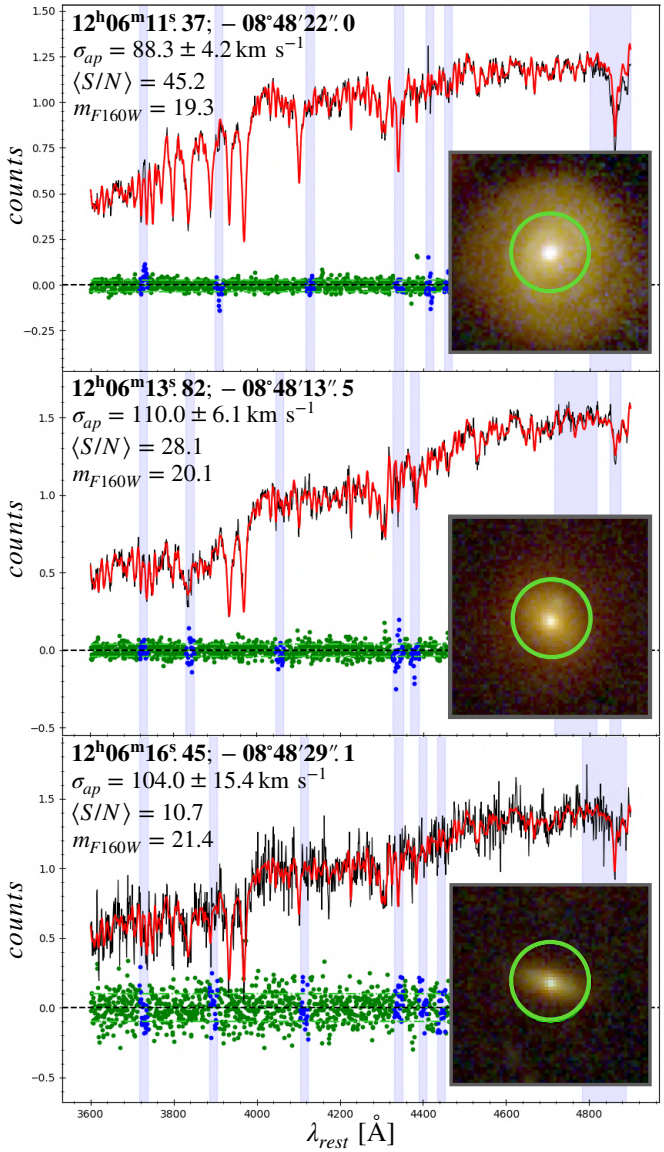


Figure 6.2: Results from the line-of-sight velocity dispersion fitting of the spectra of three cluster members in MACSJ1206, as obtained with pPXF. Galaxy spectra are shown in black; red curves are the pPXF best-fit models, while the green points correspond to the data−model residuals. The blue shaded regions along the wavelength axis were excluded in the fitting procedure due to the presence of noisy sky subtractions around emission lines in the spectra. Corresponding residuals in these regions are marked in blue. The first from the top is a high signal-to-noise E+A galaxy. The second is a passive galaxy spectrum with a $\langle S/N \rangle = 28.1$, corresponding approximately to the mean S/N of our galaxy sample. The bottom spectrum has a $\langle S/N \rangle = 10.7$, close to our lower limit for reliable velocity dispersion measurements. Coordinates, measured velocity dispersion (σ_{ap}), mean signal-to-noise and F160W magnitudes are indicated in each panel. Cutouts are HST RGB images, $4''$ across, showing in green the apertures of $0.8''$ radius used for the spectral extraction. (Figure from Bergamini et al. 2019)

With the simulations (see Section 5.4), we check the reliability of the statistical error provided by `pPXF` on our data, as well as the presence of systematic errors as a function of $\langle S/N \rangle$ and input velocity dispersion ($\sigma_{\rm in}$). The latter can become important especially when measuring velocity dispersions of low mass galaxies, as we approach the MUSE instrument resolution. In particular, we find that statistical errors are generally underestimated by $\sim 20\%$ for $\langle S/N \rangle > 15$, up to $\sim 25\%$ for $\langle S/N \rangle \sim 10$, whereas a positive bias of a few $\mathrm{km\,s^{-1}}$ becomes evident for $\sigma \lesssim 100\ \mathrm{km\,s^{-1}}$ at high $\langle S/N \rangle$. In addition, we find that measurements become increasingly uncertain at $\langle S/N \rangle < 10$. We therefore include in our galaxy kinematic sample only galaxies with $\langle S/N \rangle > 10$ and $\sigma > 80\ \mathrm{km\,s^{-1}}$. In all cases, we use the empirical formulae in Eq. 5.15 and in Fig. 5.5 to correct the measured velocity dispersions and their uncertainties in different $\langle S/N \rangle$ and σ regimes.
Simulations were also used to choose the optimal wavelength range for `pPXF` fits. The resulting selected rest-frame wavelength ranges are 3600 - 4900 Å, 3600 - 5200 Å and 3600 - 5300 Å for MACSJ1206, MACSJ0416 and AS1063, respectively.
The LOSVD input velocity value (zero moment) is taken from our redshift catalogs. The measured parameters are the velocity shift (V, typically within $50\ \mathrm{km\,s^{-1}}$), the velocity dispersion σ, and higher moments (h3, h4) and their one-standard deviation errors. All velocity dispersions measured with pPXF are labeled in the following with the subscript *ap*, to emphasize that these are line-of-sight quantities within an aperture, as opposed to σ_0, and σ_{LT}, which refer to parameters inferred by the lens models (see next session). Fig. 6.2 shows examples of `pPXF` spectral fitting for three cluster galaxies in MACSJ1206, one in the high $\langle S/N \rangle$ regime, one for the median $\langle S/N \rangle$ of our sample (~ 30), and one at the limiting $\langle S/N \rangle \sim 10$.

In Table 6.1, we quote the number of uncontaminated spectra extracted in each cluster ($\mathrm{N_m^{meas}}$), for which we can reliably measure velocity dispersions, together with the total number of spectro-photometric members and relevant cluster parameters. In Fig. 6.3, the data points correspond to measured velocity dispersions of member galaxies as a function of their F160W magnitude, defining the Faber-Jackson (see Eq. 2.65) relation in the three clusters.

6.3 Strong lensing models

Accurate strong lensing models were developed for MACSJ1206, MACSJ0416 and AS1063 in C17b, C17a and C16, respectively. These models were further refined in Bonamigo et al. (2018) (hereafter in this chapter B18), who included the mass distribution of the hot gas component in each cluster, as derived from the Chandra

X-ray data, dominating the smooth baryonic cluster component. As customary in cluster strong lensing modeling, none of these models included any kinematic information on cluster galaxies. We describe here our methodology which combines the B18 lens models with internal stellar kinematics derived from a large number of velocity dispersions measured with MUSE. For each cluster, we use the same catalogs of multiple images and cluster galaxies as in B18 and the Caminha et al. models, and likewise we employ the public software `LensTool` (see Section 3.7 and Kneib et al. 1996, Jullo et al. 2007, Jullo & Kneib 2009).
A parametric lens model for the total mass distribution of each cluster is optimized searching for the set of parameters $\mathbf{\Psi}$, which maximize the likelihood defined in Eq. 3.25.
Following B18, we perform a first optimization of our lensing models assuming a positional error of $0.5''$ for images identified in HST and $1''$ for those only found in the MUSE data. We multiply these errors by a constant factor ensuring that the best-fit χ^2 is close to the number of degree of freedom of the models. These updated errors are then used to sample the posterior distributions of the free parameters.
As in the B18 models, the total mass distribution (or equivalently the gravitational potential ϕ) of each cluster is described as the sum of three contributions (see Eq. 3.27 and Section 3.7): 1) an elliptical large-scale smooth halo, which is further decomposed in a DM component and a smooth gas mass component, both parametrized as elliptical dPIE profiles; the latter is obtained from deep Chandra observations, as multiple dPIE fits, as described in B18; 2) a clumpy component representing the cluster member galaxies (DM+baryons), modeled as spherical dPIE halos; 3) a shear+foreground-structure term to take into account the presence of massive structures in the outer cluster regions and line-of-sight mass distributions. Specific details on these multiple components are given below in the description of each lens model. The general functional form for the spherical dPIE, including the relations between 3D and projected mass densities, as well as the expressions to derive aperture projected line-of-sight velocity dispersions, are given in Section 2.5.

Following our previous models and the general `LensTool` methodology, the dPIE parameters for the sub-halo population follow a scaling relation for the central velocity dispersion and the truncation radius given by Eqs. 3.28 and 3.29. A similar scaling relation for the core radius, r_{core}, is used in `LensTool`, however it is not relevant here since a vanishing core radius is adopted. In these equations, $\sigma^{gal}_{LT,i}$ is the `LensTool` fiducial velocity dispersion of each member, which is related to the central velocity dispersion of the dPIE profile by $\sigma^{gal}_{0,i} = \sqrt{3/2}\,\sigma^{gal}_{LT,i}$ (see Section 3.7).

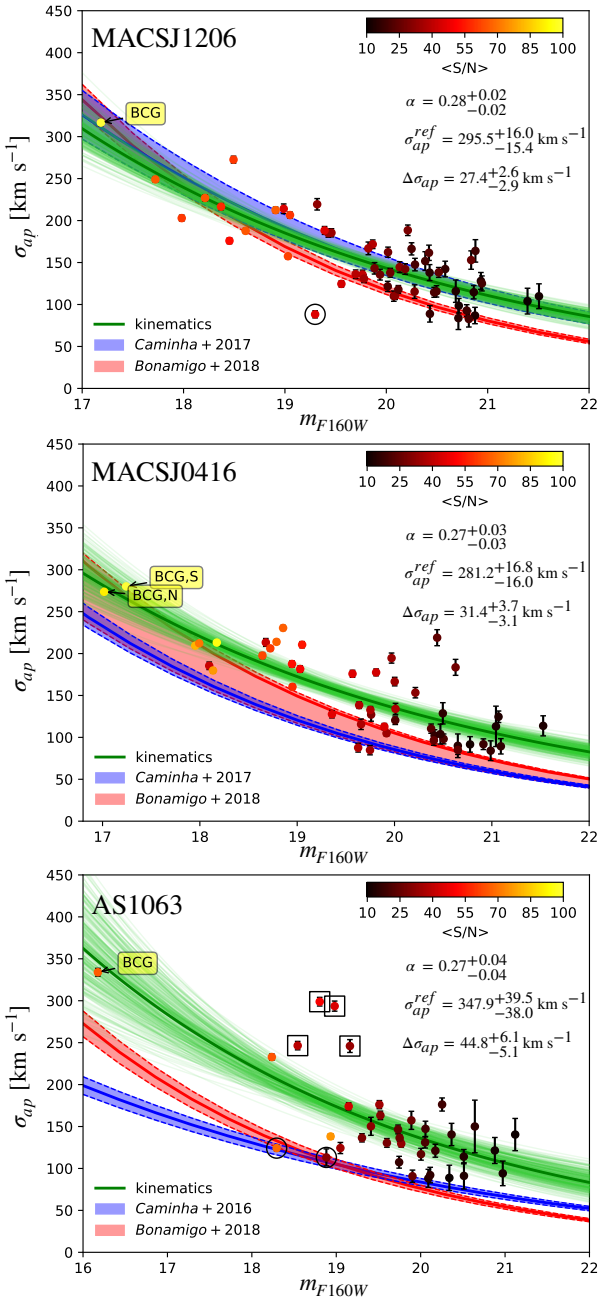


Figure 6.3: Data points are the measured stellar velocity dispersions of the cluster members, color-coded according to their spectral $\langle S/N \rangle$. Green lines are 300 scaling relations randomly drawn from the posterior distributions of the σ_{ap}-mag scaling relation parameters, α, σ_{ap}^{ref}, $\Delta\sigma_{ap}$, obtained from fitting the data points (see Fig. 6.4). Optimized parameters are also quoted. Red and blue areas are obtained by projecting the 3D σ-L scaling relations from our previous lens models with no kinematics prior (see text). Some velocity dispersion measurements, which deviate significantly from the scaling relations, are marked with squares and circles and discussed in the text. (Figure from Bergamini et al. 2019)

It is common practice to fix the slopes α and β_{cut}, so that the model optimization is performed over only two free parameters, i.e. the normalizations of the velocity dispersion and the truncation radius, corresponding to the reference luminosity L_0. We measure the luminosities L_i and L_0 using the HST F160W Kron magnitudes, which are a good proxy of the stellar mass of the cluster members (see Grillo et al. 2015) and include members down to $m_{F160W} = 24$. This leads to a minimum of 193 sub-halos to be included in the lens model (see Table 6.1), fixed at the galaxy positions.
The total mass of a circular dPIE profile is given by (see Section 2.5, Elíasdóttir et al. 2007, Limousin et al. 2005) the relation: $M_{tot} = \pi\sigma_0^2 r_{cut}/G$.
Assuming a fixed scaling between the cluster members luminosity L_i (in the same band considered by Eqs. 3.28 and 3.29) and its total mass $M_{tot,i}$, i.e. $M_{tot,i}/L_i \propto L_i^{\gamma}$, one can obtain the following relation between the slopes of Eqs. 3.28 and 3.29:

$$\beta_{cut} = \gamma - 2\alpha + 1. \tag{6.1}$$

The main goal of this chapter is to use prior information for the scaling relations derived directly from measured velocity dispersions of cluster galaxies. The latter are light-weighted projected values of the 3D-velocity dispersions within the extracted spectroscopic apertures, hereafter σ_{ap}. In order to compare σ_{LT} - L scaling relations with measured quantities, we need to compute the aperture-averaged line-of-sight velocity dispersion from the adopted dPIE mass models. The projection coefficients needed to transform σ_{LT} into σ_{ap} are obtained with a numerical integration depending on r_{cut}, r_{core} and R_{ap} (see Eq. 2.63). In Section 2.5 we show all the equations to perform this projection procedure.

In Fig. 6.3, we show the best-fit scaling relations obtained from the previous models for each cluster (red and blue curves), which did not include any prior from internal kinematics of cluster members. These curves are computed by projecting the `LensTool` scaling relations, using the posterior distribution of the parameters r_{cut} and σ_{LT}^{ref} at different magnitudes, yielding the model σ_{ap} as a function of m_{F160W}. In this process, we use an aperture of $0.8''$, so that the projected scaling relations can be directly compared with our kinematic measurements.

As customary in previous lens models to date, all the slopes for scaling relations α, β_{cut}, γ (see Eq. 6.1) are fixed, here instead we take advantage of the measured stellar velocity dispersions to directly fit the normalization σ_{ap}^{ref} and slope α of the σ - mag scaling relation. We then derive β_{cut} by adopting $\gamma = 0.2$, which is consistent with the canonical fundamental plane (see Eq. 2.68, Faber et al. 1987, Bender et al. 1992), and that we verify to be appropriate using photometric and morphological

data for AS1063 (see below and Mercurio et al. in prep.).

In order to fit the σ - mag relation, the following Bayesian approach is used. Given a set of N cluster members of magnitude m_i^{gal} with measured velocity dispersions, $\sigma_{ap,i}^{gal} \pm \delta\sigma_{ap,i}^{gal}$, the σ - mag scaling relation, corresponding to the σ - L^{α} relation in Eq. 3.28, can be written as:

$$\hat{\sigma}_{ap,i}^{gal} = \sigma_{ap}^{ref} 10^{0.4\,(m_{F160W}^{ref} - m_i^{gal})\,\alpha}, \tag{6.2}$$

where $\hat{\sigma}_{ap,i}^{gal}$ are the model predicted velocity dispersions for a cluster member with F160W magnitude m_i^{gal}.
To estimate the model parameters of the scaling relation and their uncertainties, we sample the posterior distribution of σ_{ap}^{ref} and α, including the intrinsic scatter $\Delta\sigma_{ap}$ of the measured velocities around the backbone of the scaling relation. Using Bayes' theorem, the posterior probability function can be written as:

$$p\left(\sigma_{ap}^{ref}, \alpha, \Delta\sigma_{ap} \mid m^{gal}, \sigma_{ap}^{gal}, \delta\sigma_{ap}^{gal}\right) \propto \\ p\left(\sigma_{ap}^{gal} \mid m^{gal}, \delta\sigma_{ap}^{gal}, \sigma_{ap}^{ref}, \alpha, \Delta\sigma_{ap}\right) p\left(\sigma_{ap}^{ref}, \alpha, \Delta\sigma_{ap}\right). \tag{6.3}$$

In particular, the posterior is the product of a likelihood function (Eq. 6.4) and a prior (Eq. 6.5):

$$\ln\left\{p\left(\sigma_{ap}^{gal} \mid m^{gal}, \delta\sigma_{ap}^{gal}, \sigma_{ap}^{ref}, \alpha, \Delta\sigma_{ap}\right)\right\} = \\ = -\frac{1}{2}\sum_{i=1}^{N}\left[\frac{\left(\sigma_{ap,i}^{gal} - \hat{\sigma}_{ap,i}^{gal}\right)^2}{\left(\delta\sigma_{ap,i}^{gal}\right)^2 + \Delta\sigma_{ap}^2} + \ln\left\{2\pi\left[\left(\delta\sigma_{ap,i}^{gal}\right)^2 + \Delta\sigma_{ap}^2\right]\right\}\right], \tag{6.4}$$

$$\ln\left[p\left(\sigma_{ap}^{ref}, \alpha, \Delta\sigma_{ap}\right)\right] = \begin{cases} -\ln(\Delta\sigma_{ap}), & \text{if } \sigma_{min}^{ref} < \sigma_{ap}^{ref} < \sigma_{max}^{ref} \\ & \text{and } \alpha_{min} < \alpha < \alpha_{max} \\ & \text{and } (\Delta\sigma_{ap})_{min} < \Delta\sigma_{ap} < (\Delta\sigma_{ap})_{max} \\ -\infty, & \text{otherwise} \end{cases}. \tag{6.5}$$

The boundaries σ_{min}^{ref}, σ_{max}^{ref}, α_{min}, α_{max}, $(\Delta\sigma_{ap})_{min}$ and $(\Delta\sigma_{ap})_{max}$ were chosen to limit the parameter space around the measured velocity dispersions.

To sample the log-posterior in the 3D parameters space (σ_{ap}^{ref}, α, $\Delta\sigma_{ap}$), we use the Affine-Invariant Markov Chain Monte Carlo (MCMC) Ensemble sampler developed by Goodman and Weare (Goodman & Weare 2010), and in particular its `python` implementation[5] (Foreman-Mackey et al. 2013). The parameter space is explored with 100 walkers with 5000 steps each, which are initialized in a narrow Gaussian sphere around the maximum-likelihood point. To ensure that the final distributions are independent from the initial walker positions, we remove 80 steps in the burn-in phase based on the auto-correlation time computed for each parameter.
The posterior probability distributions of the scaling relation model parameters (Eq. 6.3) obtained with this procedure are shown in Fig. 6.4 for MACSJ1206 and 6.5 for MACSJ0416 and AS1063.
Further details on the Affine-Invariant approach will be given in Chapter 9.

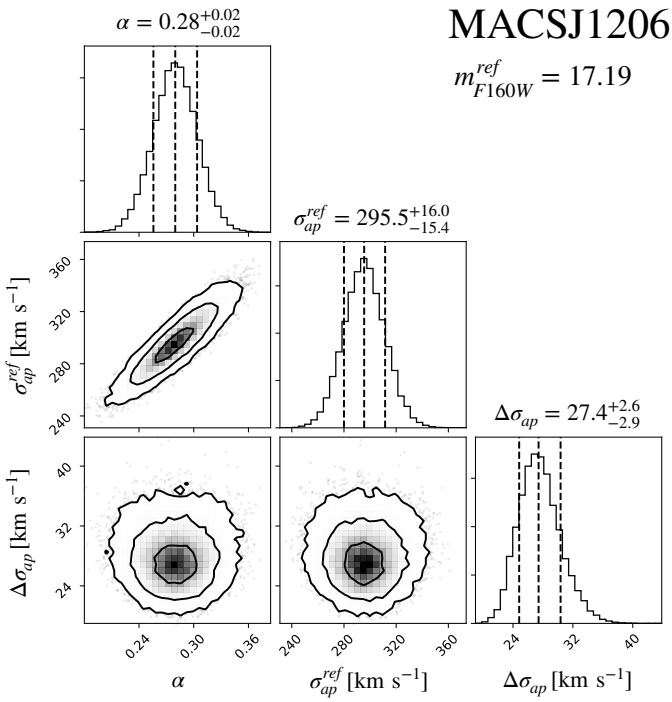


Figure 6.4: Posterior probability distributions for the σ_{ap}-mag scaling relation parameters, obtained from the velocity dispersion measurements of the 58 cluster members in MACSJ1206. The 16^{th}, 50^{th} and 84^{th} percentiles of the marginalized distributions for the slope (α), normalization (σ_{ap}^{ref}) and scatter around the scaling relation ($\Delta\sigma_{ap}$) are quoted and shown as vertical dashed lines. (Figure from Bergamini et al. 2019)

[5] https://emcee.readthedocs.io/en/latest/

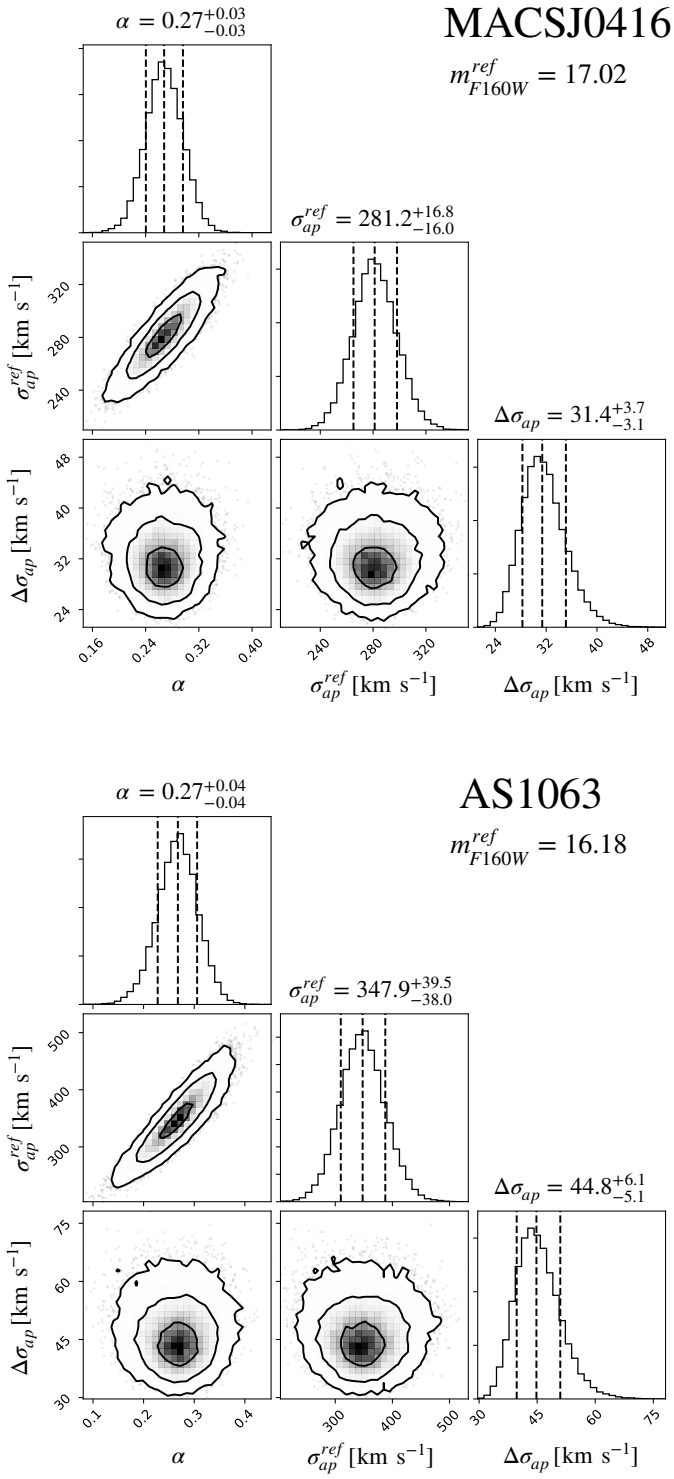


Figure 6.5: Same as Fig. 6.4 but for MACSJ0416 and AS1063. The σ_{ap}-mag scaling relation parameters are optimized using the velocity dispersion measurements of 49 and 37 cluster members in MACSJ0416 and AS1063 respectively. (Figure from Bergamini et al. 2019)

The best-fit scaling relations obtained with this method are shown in Fig. 6.3, including the corresponding uncertainties (green curves), which are derived by sampling 300 times the posterior distributions. These best-fit parameters for each scaling relation (see Table 6.2) are then used as dynamical priors in our lens models, as described below. It is interesting to note that the slope that we obtain for the Faber-Jackson relation, L - $\sigma^{1/\alpha}$, is very similar for the three clusters ($\alpha = 0.27$ - 0.28) and consistent with several spectro-photometric studies of cluster early-type galaxy populations in the literature (e.g. Kormendy & Bender 2013, Focardi & Malavasi 2012).

Fig. 6.3 shows that some galaxies deviate significantly from the best-fit scaling relations (as marked by boxes and circles in the figure). This is however expected, as the Faber-Jackson relation is one of the projections of the fundamental plane relation among half-light radius R_e, mean surface brightness μ_e within R_e, and velocity dispersion: at a given luminosity, more compact galaxies tend to have higher velocity dispersion (see Eq. 2.68). For example, the four cluster galaxies at $m_{F160W} \sim 19$ in AS1063, with a σ significantly higher than the best-fit relation (see boxes in Fig. 6.3), have μ_e in the 16-th highest percentile, however we verified that they still lie on the fundamental plane defined by Jorgensen et al. (1996) (Mercurio et al., in prep.).
On the other hand, the few galaxies lying well below the σ - mag relation (see circled data points in Fig. 6.3 and the spectrum in the upper panel of Fig. 6.2) show sign of emission lines and young stellar populations in their spectra, for which lower velocity dispersions are expected when compared to early-type galaxies with similar luminosities.

Cluster	m^{ref}_{F160W}	σ^{ref}_{ap} [km s^{-1}]	α	$\Delta\sigma_{ap}$ [km s^{-1}]	$\beta_{cut}(\gamma = 0.2)$
MACS J1206.2−0847	17.19	$295.5^{+16.0}_{-15.4}$	$0.28^{+0.02}_{-0.02}$	$27.4^{+2.9}_{-2.6}$	$0.64^{+0.04}_{-0.04}$
MACS J0416.1−0403	17.02	$281.2^{+16.8}_{-16.0}$	$0.27^{+0.03}_{-0.03}$	$31.4^{+3.7}_{-3.1}$	$0.66^{+0.06}_{-0.06}$
Abell S1063	16.18	$347.9^{+39.5}_{-38.0}$	$0.27^{+0.04}_{-0.04}$	$44.8^{+6.1}_{-5.1}$	$0.66^{+0.08}_{-0.08}$

Table 6.2: σ_{ap}-L scaling relation parameters derived from measured velocity dispersions of cluster members for MACSJ1206, MACSJ0416 and AS1063. The normalization parameter, σ^{ref}_{ap}, is computed at the reference magnitudes m^{ref}_{F160W}. Median values are derived from the marginalized parameter distributions, while the errors correspond to the 16^{th} and 84^{th} percentiles. The β_{cut} values are obtained using Eq.6.1. (Table from Bergamini et al. 2019)

6.4 Combining lensing models with kinematics measurements

In Fig. 6.3, a comparison between the sub-halo scaling relations obtained from the lens models and those directly constrained from kinematic measurements (green curves) shows significant discrepancies in the case of MACSJ0416 and AS1063, while they are consistent for MACSJ1206. In the case of AS1063, the normalization of the kinematic scaling relations are found to be approximately 100 and 150 $\mathrm{km\,s^{-1}}$ above the values inferred from the B18 and C16 lens models, respectively. In MACSJ0416, the discrepancy is significant ($\sim$50 $\mathrm{km\,s^{-1}}$) for the C17a model, while it is negligible for the B18 model within the errors, albeit with a slope which deviates from the observed one ($\alpha = 0.27$ against 0.35 assumed in B18). This shows how inherent degeneracies of sub-halo population parameters in strong lensing models can lead to inferred velocity dispersion normalizations which are inconsistent with kinematic measurements of cluster galaxies. Nevertheless, these lens models can reproduce the positions of the multiple images with high precision, with a root-mean square value between the observed and model-predicted images on the image plane of $\Delta_{rms} \simeq 0.45$ - $0.6''$ (see Table 6.4). It is in fact well known that parametric cluster lens models are, in general, affected by some degeneracy between the mass distribution of the macro-halo(s) and that of the sub-halos, even when a large number of constraints are available, as in our case (Meneghetti et al. 2017). Despite that the projected total mass value within a given cluster-centric radius remains robust. In addition, a significant degeneracy exists between the central velocity dispersions and the cut-off radii of the sub-halos, as we showed in Section 3.7.1.

In the following, we describe in detail our new lens models for each cluster. We start from the same parameterization and input constraints as in the B18 models, we then proceed to add critical constraints on the sub-halo scaling relation parameters, with priors from our kinematic measurements. The optimization of model parameters is obtained from MCMC chains of approximately 10^5 samples, excluding the burn-in phase. The model input components and parameters are summarized in Table 6.3, while the output results of the lens model optimizations are in Table 6.4. Note that, as customary in the literature, no scatter in the scaling relations is assumed in our lens models.

MACSJ1206

Following C17b and B18 models, the cluster smooth mass distribution includes three dark matter halos which can reproduce the apparent elongated asymmetry

in the distribution of the cluster galaxies and intra-cluster light. The three halos are described as dPIEs profiles whose values of sky positions, ellipticities, position angles, core radii and velocity dispersions are left free to vary with flat priors, while their truncation radii are fixed to a large value. Three dPIE gas clumps are used to model the X-ray surface brightness distribution as in B18. Thus, the cluster-scale mass components have a total of 18 free parameters. The presence of an external shear term introduces 2 extra free parameters in the model. Finally, the clumpy component includes 258 halos describing the cluster members and the BCG, centered on the peaks of their light emission. All these galaxies are described as circular dPIEs profiles, whose values of central velocity dispersion and truncation radius scale with their F160W magnitude, according to Eqs. 3.28 and 3.29, where the BCG magnitude is used as reference luminosity (L_0). The normalization r_{cut}^{ref} is free to vary between 1″and 50″ (i.e. 5.67 - 283.9 kpc at the cluster redshift) with a flat prior. A Gaussian prior derived from our kinematic measurements is then introduced for the σ - L scaling relation. To determine this prior, we deproject the best-fit (median) normalization of the σ - mag relation ($\sigma_{ap}^{ref} = 295.5^{+16.0}_{-15.4}$ km s^{-1}, see Fig. 6.4) to obtain a `LensTool` fiducial reference velocity dispersion, σ_{LT}^{ref}. This deprojection is achieved with an iterative procedure by computing the projection coefficient, $c_p(r_{core}, r_{cut}, R_{ap})$ (see Eq. 2.63), using the best-fit scaling relation parameters from a first model without kinematic prior, as $\sigma_{LT}^{ref} = \sigma_{ap}^{ref}/c_p$. In the case of MACSJ1206, we thus obtain a Gaussian prior on σ_{LT}^{ref} with a mean of 264 km s^{-1} and a standard deviation of 18 km s^{-1}. The standard deviation also includes the uncertainty in the computation of c_p, due to the allowed range of r_{cut}. Finally, the slope α is fixed to the value obtained from the σ - mag fit ($\alpha = 0.28$), while a $\beta_{cut} = 0.64$ is obtained from Eq. 6.1, with $\gamma = 0.2$. The lens model is thus optimized with a total number of 22 free parameters, using the observed positions of 82 spectroscopically confirmed multiple images associated to 27 families (C17b and B18).

MACSJ0416

Referring to Table 6.3, the cluster-scale mass components of this merging cluster (Balestra et al. 2016) include two massive dark matter halos, whose positions are left free to vary in areas of 30″ × 30″ and 15″ × 15″ around the northern and southern BCG, respectively (BCG,N, BCG,S); plus a third circular halo in the NW region to improve the model accuracy (Δ_{rms}) in that area. Four dPIEs with fixed parameters are used to describe the complex X-ray emitting gas distribution in MACSJ0416, as done in B18. The cluster-scale component has therefore 16 free parameters. No shear is present, however an additional circular dPIE halo with free velocity dispersion and truncation radius is added to account for the presence of a foreground galaxy (Caminha et al. 2017a, Jullo et al. 2007).

The sub-halo mass component, including the BCG, comprises 193 halos in this case. The r_{cut}^{ref} parameter is left free to vary over a wide range (1 - 20″, or 5.3 - 106.8 kpc). As described above for MACSJ1206, the deprojection of the best-fit σ_{ap} - mag relation provides a Gaussian prior for the normalization of $\sigma_{LT}^{ref} = 249 \pm 15$ km s^{-1}, while we use the measured slope $\alpha = 0.27$ (see Fig. 6.5/top), from which we derive $\beta_{cut} = 0.66$ (from Eq. 6.1, with $\gamma = 0.2$). With a total number of 20 free parameters, 102 spectroscopically confirmed multiple images (associated to 37 families) are then used to constrain the lens model (as in C17a and B18).

AS1063

The dark-matter macro-halo includes in this case an elliptical dPIE, close to the BCG position, and a circular dPIE whose position is left free to vary in an area of 150″ × 120″ centered in the North-East region of the cluster (Table 6.3). Three dPIEs are used to describe the gas component (B18), and no shear term is included. There are therefore 9 free parameters for the cluster-scale component.
The clumpy sub-halo component is constituted by 222 cluster members, including the BCG. From our kinematic measurements and best-fit scaling relation, we obtain a deprojected value of $\sigma_{LT}^{ref} = 310^{+35}_{-34}$ km s^{-1}. As a result, we adopt the median value of 310 km s^{-1} as a Gaussian prior for the normalization of the σ_{ap} - L scaling relation, while the standard deviation is reduced to 15 km s^{-1} to force the lens model to converge to a χ^2 minimum solution compatible with our galaxy kinematics. The slope is fixed to the measured value $\alpha = 0.27$ (see Fig. 6.5/bottom), which implies $\beta_{cut} = 0.66$, for $\gamma = 0.2$. The lens model, which has 11 free parameters in total, is thus optimized using the positions of 55 spectroscopically confirmed multiple images, divided in 20 families, as implemented in B18.

6.5 Results and discussion

The inclusion of the kinematic prior in the lens models has the main consequence of significantly reducing the intrinsic degeneracies in the scaling relation parameters. Specifically, the inherent degeneracy discussed in Section 3.7.1 between the normalization parameters of the σ - L and r_{cut} - L scaling relation. Indeed, several combinations of σ_0 and r_{cut} will yield similar aperture masses, which are constrained by the multiple image positions. This degeneracy, in combination with those among the parameters describing cluster-scale mass components or shear terms, can lead the lens model to predict a distribution of sub-halo masses not consistent with internal kinematics of member galaxies. This can be noticed in Fig. 6.3, where we

			dPIE parameters							slopes
			x	y	e	θ [°]	σ_{LT} [km s^{-1}]	r_{core} [″]	r_{cut} [arcsec]	$\gamma^{\alpha}_{\beta_{core}}$
MACS J1206.2−0847	HALOS	1^{st}DM	[-3.0, 3.0]	[-3.0, 3.0]	[0.0, 0.9]	[5.0, 35.0]	[450.0, 1200.0]	[0.0, 20.0]	2000.0	-
		2^{nd}DM	[0.0, 30.0]	[-5.0, 25.0]	[0.0, 0.9]	[0.0, 180.0]	[10.0, 1200.0]	[0.0, 40.0]	2000.0	-
		3^{rd}DM	[-60.0, -5.0]	[-60.0, 60.0]	[0.0, 0.9]	[-90.0, 90.0]	[10.0, 1100.0]	[0.0, 40.0]	2000.0	-
		1^{st}Gas	3.11	-6.34	0.12	-0.71	452.2	63.29	403.05	-
		2^{nd}Gas	-13.50	-7.24	0.50	-113.57	342.3	40.52	43.94	-
		3^{rd}Gas	3.31	2.04	0.58	-169.20	186.9	8.23	68.57	-
	S	Shear	-	-	[0.0, 1.0]	[0.0, 180.0]	-	-	-	-
	G	258 [17.19]	-	-	0.0	0.0	(264.0, 18.0)	0.01	[1.0, 50.0]	$0.2^{0.28}_{0.64}$
MACS J0416.1−0403	HALOS	1^{st} DM	[-15.0, 15.0]	[-15.0, 15.0]	[0.2, 0.9]	[100.0, 180.0]	[350.0, 1000.0]	[0.0, 20.0]	2000.0	-
		2^{nd}DM	[15.0, 30.0]	[-45.0, -30.0]	[0.2, 0.9]	[90.0, 170.0]	[700.0, 1200.0]	[0.0, 25.0]	2000.0	-
		3^{rd}DM	[-55.0, -25.0]	[0.0, 30.0]	0.0	0.0	[50.0, 750.0]	[0.0, 35.0]	2000.0	-
		1^{st}Gas	-18.14	-12.13	0.12	-156.76	433.0	149.21	149.82	-
		2^{nd}Gas	30.79	-48.67	0.42	-71.50	249.0	34.77	165.77	-
		3^{rd}Gas	-2.37	-1.26	0.42	-54.74	101.7	8.28	37.59	-
		4^{th}Gas	-20.13	14.74	0.40	-49.32	281.8	51.67	52.34	-
	S	Foregr.	31.96	-65.55	0.0	0.0	[50.0, 350.0]	0.05	[5.0, 100.0]	
	G	193 [17.02]	-	-	0.0	0.0	(249.0, 15.0)	0.05	[1.0, 20.0]	$0.2^{0.27}_{0.66}$
Abell S1063	HALOS	1^{st}DM	[-2.0, 4.0]	[-2.0, 2.0]	[0.4, 0.8]	[-42.0, -36.0]	[1000.0, 1400.0]	[5.0, 35.0]	2000.0	-
		2^{nd}DM	[-100.0, 50.0]	[-50.0, 70.0]	0.0	0.0	[0.0, 1000.0]	0.05	2000.0	-
		1^{st}Gas	18.90	-73.36	0.80	-162.05	335.9	188.40	189.24	-
		2^{st}Gas	-18.05	13.47	0.13	-27.80	442.6	36.32	339.16	-
		3^{rd}Gas	0.20	-1.24	0.34	-15.49	249.7	14.43	356.50	-
	G	222 [16.18]	-	-	0.0	0.0	(310.0, 15.0)	0.05	[0.0, 20.0]	$0.2^{0.27}_{0.66}$

Table 6.3: Input parameters of the three lens models developed in this chapter. The mass components are grouped into (**HALOS**), shears and foreground galaxies (**S**), and cluster galaxies following the scaling relations (**G**). All cluster-scale halos are modeled as dPIEs. Halos making up the hot gas component are taken from B18. Sky x, y coordinates are the offsets in arcsec from the reference BCG positions (MACSJ1206: $R.A. = 12^h06^m12^s.15$, $DEC = -8°48'03''.4$, MACSJ0416: $R.A. = 04^h16^m09^s.15$, $DEC = -24°04'02''.9$ and AS1063: $R.A. = 22^h48^m43^s.97$, $DEC = 44°31'51''.2$). The ellipticity e is defined as $e = \frac{a^2-b^2}{a^2+b^2}$, where a and b are the semi-major and semi-minor axis; when a shear term is present the ellipticity refers to γ_{shear} (see C17b). The position angle θ is computed from West to North; σ_{LT} is the `LensTool` fiducial velocity dispersion, r_{core} and r_{cut} are the core and truncation radius, respectively. For the galaxy component (G), the number of spherical sub-halos included in the model and the normalization F160W magnitude are quoted in the third column. The last column shows the input values for slopes in Eq. 6.1. If a flat prior is assumed on a quantity, the boundaries of the prior are given in square brackets. Instead, the means and the standard deviations of the Gaussian priors are given in round brackets. Single values (without brackets) are fixed.

			dPIE parameters							
			x [arcsec]	y [arcsec]	e	θ [°]	σ_{LT} [km s^{-1}]	r_{core} [arcsec]	r_{cut} [arcsec]	Δ_{rms} [arcsec]
MACS J1206.2−0847	HALOS	1^{st}DM	$-0.75^{+0.40}_{-0.43}$	$0.35^{+0.20}_{-0.21}$	$0.69^{+0.02}_{-0.02}$	$19.92^{+0.84}_{-0.80}$	$786.9^{+30.6}_{-35.0}$	$6.51^{+0.45}_{-0.48}$	2000.0	**This work:** 0.46
		2^{nd}DM	$9.68^{+0.78}_{-0.73}$	$3.88^{+0.74}_{-0.72}$	$0.56^{+0.09}_{-0.10}$	$114.65^{+2.75}_{-2.38}$	$620.9^{+27.3}_{-26.5}$	$14.37^{+1.50}_{-1.18}$	2000.0	**B18:** 0.45
		3^{rd}DM	$-28.46^{+1.38}_{-1.43}$	$-6.61^{+0.78}_{-0.74}$	$0.34^{+0.06}_{-0.06}$	$-23.83^{+10.59}_{-12.07}$	$491.4^{+37.4}_{-31.9}$	$12.02^{+2.15}_{-1.91}$	2000.0	**C17b:** 0.44
		1^{st}Gas	3.11	−6.34	0.12	−0.71	452.2	63.29	403.05	
		2^{nd}Gas	−13.50	−7.24	0.50	−113.57	342.3	40.53	43.94	
		3^{rd}Gas	3.31	2.04	0.58	−169.20	186.9	8.24	68.57	
	S	Shear	-	-	$0.12^{+0.01}_{-0.01}$*	$101.15^{+1.41}_{-1.36}$	-	-	-	
	G	258 [17.19]	-	-	0.0	0.0	$272.6^{+12.7}_{-12.6}$	0.01	$4.10^{+0.76}_{-0.64}$	
MACS J0416.1−0403	HALOS	1^{st} DM	$-2.14^{+1.06}_{-0.84}$	$1.36^{+0.63}_{-0.73}$	$0.84^{+0.01}_{-0.04}$	$144.75^{+1.12}_{-1.01}$	$581.0^{+19.9}_{-24.0}$	$6.56^{+0.59}_{-0.60}$	2000.0	**This work:** 0.61
		2^{nd}DM	$20.01^{+0.22}_{-0.23}$	$-37.20^{+0.44}_{-0.45}$	$0.76^{+0.01}_{-0.01}$	$125.95^{+0.46}_{-0.39}$	$859.9^{+15.0}_{-15.4}$	$11.98^{+0.60}_{-0.58}$	2000.0	**B18:** 0.59
		3^{rd}DM	$-33.99^{+0.93}_{-1.12}$	$8.38^{+2.83}_{-0.85}$	0.0	0.0	$314.2^{+47.7}_{-50.0}$	$5.58^{+2.62}_{-2.76}$	2000.0	**C17a:** 0.59
		1^{st}Gas	−18.14	−12.13	0.12	−156.76	433.0	149.21	149.82	
		2^{nd}Gas	30.79	−48.67	0.42	−71.50	249.0	34.77	165.77	
		3^{rd}Gas	−2.37	−1.26	0.42	−54.74	101.7	8.28	37.59	
		4^{th}Gas	−20.13	14.74	0.40	−49.32	281.8	51.67	52.34	
	S	Foreground	31.96	−65.55	0.0	0.0	$178.0^{+14.6}_{-15.0}$	0.05	$61.9^{+25.07}_{-21.55}$	
	G	193 [17.02]	-	-	0.0	0.0	$262.0^{+8.5}_{-10.2}$	0.05	$5.68^{+0.81}_{-0.69}$	
Abell S1063	HALOS	1^{st}DM	$1.40^{+0.23}_{-0.23}$	$-0.74^{+0.16}_{-0.17}$	$0.63^{+0.01}_{-0.01}$	$-38.95^{+0.22}_{-0.23}$	$1162.4^{+6.4}_{-6.7}$	$18.06^{+0.53}_{-0.52}$	2000.0	
		2^{nd}DM	$-50.16^{+3.70}_{-4.41}$	$26.80^{+2.36}_{-2.17}$	0.0	0.0	$221.4^{+24.2}_{-22.3}$	0.05	2000.0	**This work:** 0.55
		1^{st}Gas	18.90	−73.36	0.80	−162.05	335.9	188.40	189.24	**B18**: 0.51
		2^{st}Gas	−18.05	13.47	0.13	−27.80	442.6	36.32	339.16	**C16:** 0.51
		3^{rd}Gas	0.20	−1.24	0.34	−15.49	249.7	14.43	356.50	
	G	222 [16.18]	-	-	0.0	0.0	$299.4^{+14.3}_{-14.2}$	0.05	$6.83^{+1.69}_{-1.32}$	

Table 6.4: Output parameters of the three lens models developed in this chapter. The mass components are grouped into (**HALOS**), shears and foreground galaxies (**S**), and cluster galaxies following the scaling relations (**G**). Parameters with no errors are fixed. All cluster-scale halos are modeled as dPIEs. Halos making up the hot gas component are taken from B18. Sky x, y coordinates are the offsets in arcsec from the reference BCG positions (MACSJ1206: $R.A. = 12^h06^m12^s.15$, $DEC = -8°48'03''.4$, MACSJ0416: $R.A. = 04^h16^m09^s.15$, $DEC = -24°04'02''.9$ and AS1063: $R.A. = 22^h48^m43^s.97$, $DEC = 44°31'51''.2$). The ellipticity e is defined as $e = \frac{a^2-b^2}{a^2+b^2}$, where a and b are the semi-major and semi-minor axis; when a shear term is present the ellipticity refers to γ_{shear} (see C17b). The position angle θ is computed from West to North; σ_{LT} is the `LensTool` fiducial velocity dispersion, r_{core} and r_{cut} are the core and truncation radius, respectively. For the galaxy component (G), the number of spherical sub-halos included in the model and the normalization F160W magnitude are quoted in the third column. For each optimized parameter, we quote the median, and the 16^{th} and 84^{th} percentiles as errors. The last column shows the total root-mean-square separation between the model-predicted and observed positions of the multiple images. (Table from Bergamini et al. 2019)

display the case of MACSJ0416 and AS1063.

The implementation of the kinematic prior on the normalization and slopes of the

scaling relation leads to models with a global precision which is similar to that of the previous lens models, as apparent from a comparison of Δ_{rms} values (see Table 6.4). The new Δ_{rms} values are larger than the B18 models by $0.01''$ for MACSJ1206, $0.02''$ for MACSJ0416 and $0.04''$ for AS1063, with the same number of degrees of freedom (88 for MACSJ1206, 110 for MACSJ0416 and 59 for AS1063). Since the informative Gaussian priors on the cluster member scaling relations reduce the freedom of the lens models, a small increase of the Δ_{rms} is not surprising. Indeed, while an uninformative prior on σ^{ref} leave the freedom to scaling relations to compensate the imperfections in the parameterization of cluster-scale mass components, the kinematic Gaussian priors force the scaling relation to kinematic compatible solutions. Moreover, a completely free scaling relation is mostly constrained by those galaxies close to multiple images. Due to the intrinsic scatter of galaxies around the σ - L relation, this might generate biases respect to the measured scaling relation. A solution to this issue is discussed in Chapter 9.
In Fig. 6.6, we show the posterior probability distributions for the projected values of the normalization of the σ - L and r_{cut} - L relations, σ_{ap}^{ref} and r_{cut}^{ref}, respectively, for the new (in green) and previous models. The kinematic priors lead in general to significant smaller uncertainties for these parameters, up to factor of 10 for the truncation radius of AS1063, with respect to the B18 results.

While for MACSJ1206 the new solution is consistent with the kinematic data, for the other two clusters the kinematic prior moves the best-fit solution to a different region of the parameter space. In the case of MACSJ0416, the B18 model had already found a different solution with a higher σ-normalization with respect to the C17a, by extending the MCMC parameter search and including the hot gas mass component. Our new lens model is consistent within 1σ with B18, further reducing parameters' uncertainties especially on r_{cut}. In the case of AS1063, the kinematic prior moves the σ_{ap}^{ref} (for the BCG) to a much larger value, which is 70 $\mathrm{km\,s^{-1}}$ higher, and provides a much tighter constraint for r_{cut}^{ref}, moving the solution away from previous values of ~ 190 kpc (note that a dPIE encircles 90% of the total projected mass within $5r_{cut}$).

It is relevant to note that, despite the higher normalization of the sub-halo scaling relations for MACSJ0416 and AS1063, the cluster total projected mass profile does not vary appreciably when compared to previous models. This is expected, since the multiple image positions provide information about the cluster total mass projected within circles with radii equal to the average distance of the multiple images of each family. In Fig. 6.7, one can appreciate that the differences in the cumulative projected mass between the new and previous models are well within 10% over the radial region with multiple image constraints.

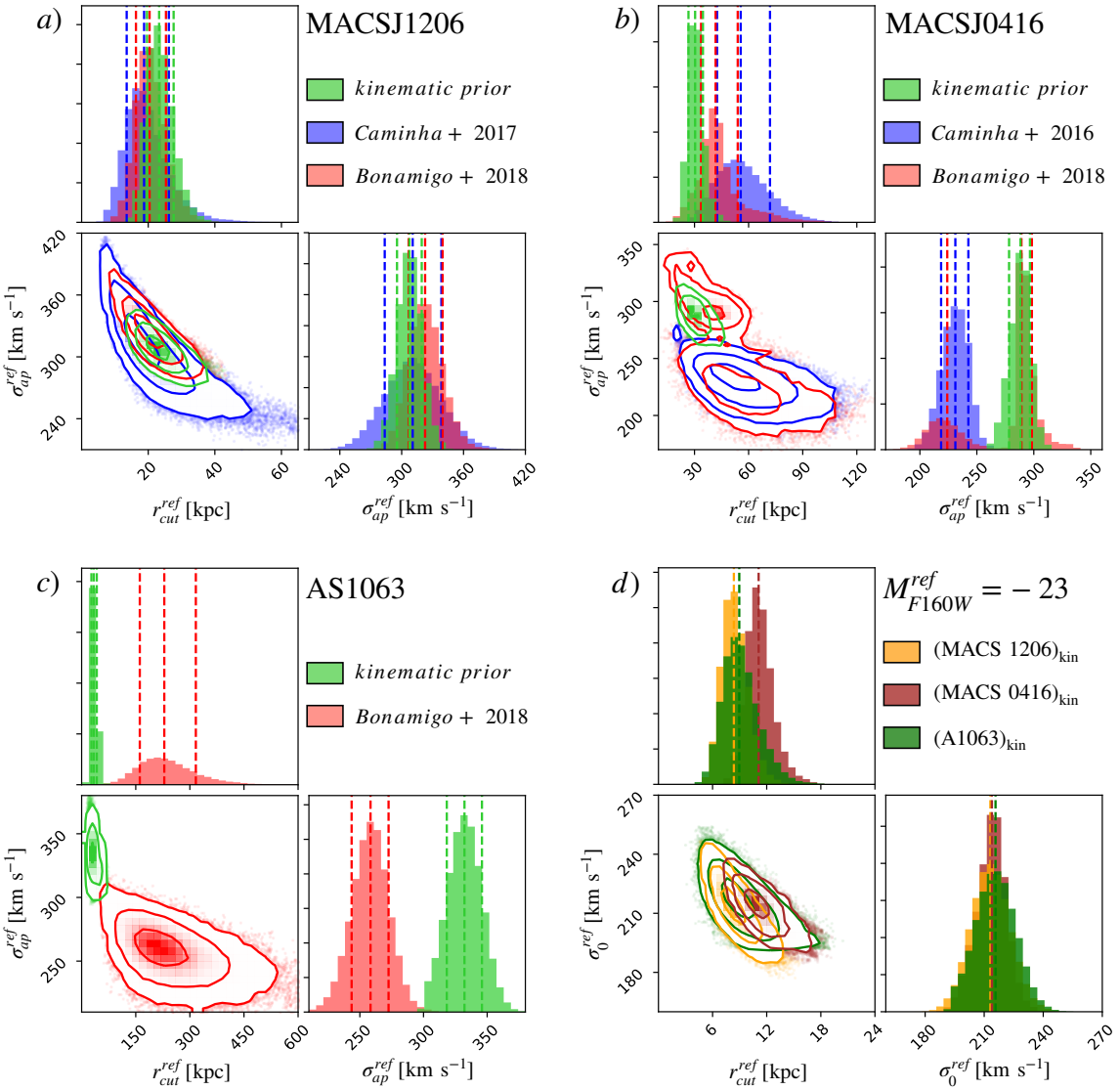


Figure 6.6: Panels a,b,c: posterior probability distributions for the normalization of the σ_{ap}-L scaling relation, i.e. the aperture projected velocity dispersion σ^{ref}_{ap}, and the truncation radius (r^{ref}_{cut}), obtained with different lens models for the three clusters under study. The green contours and distributions refer to our best models which include the kinematic prior based on measured velocity dispersions of member galaxies (Fig. 6.3). For visualization clarity, in panel c we omit the C16 results because they are outside the chosen range of velocity dispersion, as visible in the bottom panel of Fig. 6.3. The reference magnitudes for each cluster are in Table 6.2. In panel (d), we show the posterior distribution of the reference central velocity (σ^{ref}_0) and reference truncation radius (r^{ref}_{cut}) in each cluster, normalized to an absolute mag $M^{ref}_{F160W} = -23$, close to L^* of the early-type galaxy population. The different colors correspond to the models with kinematic priors and contours refer to 1, 2, 3σ confidence levels. Once normalized to the same luminosity, the three scaling relations for σ and r_{cut} are consistent. (Figure from Bergamini et al. 2019)

Therefore, although the fraction of total mass in sub-halos differs for MACSJ0416 and AS1063, the total encircled mass remains a robust measurement. This however shows that the inclusion of the kinematic prior is important to obtain a reliable estimate of the relative contribution of the diffuse and clumpy (sub-halo) mass components from strong lensing models. In general, an underestimate of the sub-halo mass component will lead to a higher contribution of the cluster-scale halos, in particular the central one.

The consistency of the sub-halo scaling relations, based on kinematic measurements in the three clusters, suggests a rescaling of the normalization of the σ - L and r_{cut} - L relations to the same absolute luminosity for a meaningful comparison. To this aim, we choose as reference an absolute magnitude of $M^{ref}_{F160W} = -23$, which is close to the value of L^*, estimated from the F160W luminosity function of MACSJ0416 at z=0.4 (Connor et al. 2017 and Mercurio et al. 2019, in prep.). We then rescale the normalization of the scaling relations obtained from the kinematic lens models, anchored on the BCG magnitude (m^{ref}_{F160W}), to the new reference M^{ref}_{F160W} using the distance modulus μ of each cluster, as follows:

$$\sigma_0^M = \sigma_0^{ref} 10^{0.4\,[m^{ref}_{F160W} - \mu(z_{cl}) - M^{ref}_{F160W}]\,\alpha}, \tag{6.6}$$

where the central reference velocity dispersion is $\sigma_0^{ref} = \sqrt{3/2}\,\sigma_{LT}^{ref}$ in Eq. 3.28, and

$$r_{cut}^M = r_{cut}^{ref} 10^{0.4\,[m^{ref}_{F160W} - \mu(z_{cl}) - M^{ref}_{F160W}]\,\beta_{cut}}. \tag{6.7}$$

The comparison of the posterior distributions of σ_0, and r_{cut}, obtained from our lens models with kinematic prior, once renormalized to the same L^* luminosity (i.e. M^{ref}_{F160W}), is shown in the bottom right panel (d) in Fig. 6.6. The consistency of these distributions among the three clusters is quite remarkable, particularly for r_{cut}, and suggests an empirical relation for the truncation radius, which is generally poorly constrained by lens models without kinematic priors. By combining Eqs. 3.28 and 3.29, one obtains:

$$r_{cut} = 10.1^{(13.1)}_{(7.3)}\ \mathrm{kpc} \left(\frac{\sigma_0}{220\ (\mathrm{km\,s^{-1}})}\right)^{2.43^{(2.45)}_{(2.38)}}, \tag{6.8}$$

where the range for each parameter represents the 16^{th}-84^{th} percentiles of the combined posterior distributions of σ_0^{ref}, and r_{cut}^{ref}, and the slope is obtained by the distribution of β_{cut}/α.

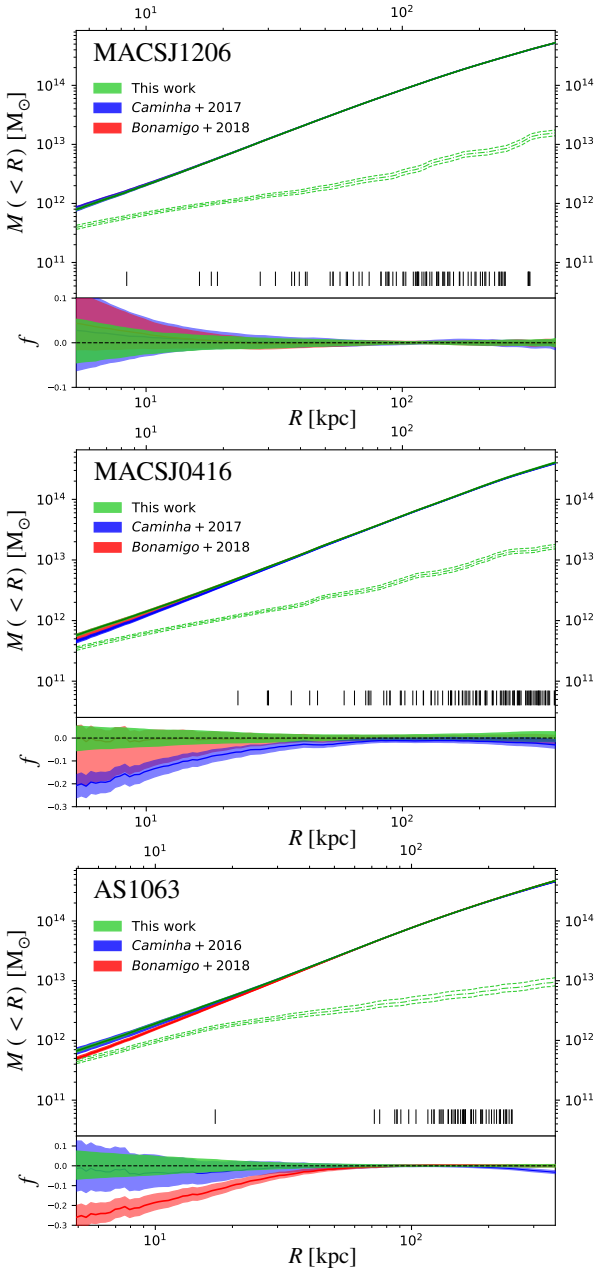


Figure 6.7: *Top:* Projected cumulative mass profiles, as a function of the projected distance from the the BCGs, corresponding to previous and current lens models for each cluster. The colored regions encompass the 16^{th} and the 84^{th} percentiles; colored solid lines are the median values. The dashed green lines correspond to the mass component associated to cluster members (i.e. sub-halos) from the new lens models with kinematic prior (16^{th}, 50^{th}, 84^{th} percentiles). The multiple image projected distances from the cluster centers are marked with vertical black lines. *Bottom:* Relative variation of the cumulative projected total masses with respect to our reference (green) model. (Figure from Bergamini et al. 2019)

A similar relation was derived by Monna et al. (2015) for Abell 383 at $z = 0.18$, incorporating the velocity dispersion measurements of 21 member galaxies. However, in that study the slopes of the $\sigma \propto L^{\alpha}$ and $r_{cut} \propto L^{\beta_{cut}}$ relations are fixed, based on literature studies, assuming $M_{tot}/L = const.$ In our case, the slope α is directly measured from the MUSE spectra of a cluster galaxy sample larger by more than a factor of 2, while β_{cut} is derived assuming galaxies lying on the fundamental plane.

The slope $\beta_{cut}/\alpha = 2.43$ that we find for the r_{cut} - σ_0 relation is likewise the result of our kinematic measurements in three different clusters, whereas the slope of 1.25 quoted in Monna et al. (2015) is derived from the assumed values of α and β_{cut}. However, the normalization of the relation in Eq. 6.8 turns out to be consistent with the Monna et al. (2015) result. Other attempts to constrain halo sizes of cluster galaxies have involved modeling of single strong lensing systems in clusters, with no kinematics information from spectroscopic measurements. For example, Suyu & Halkola (2010) found a $r_{cut} = 6.0^{+2.9}_{-2.0}$ kpc for a galaxy with a lens-based velocity dispersion $\sigma_0 = 127^{+21}_{-12}$ km s^{-1} in a group, in agreement with our results. Eichner et al. (2013) found instead a higher normalization by modeling the surface brightness distribution of the "snake arc" in MACSJ1206. However, we should note that the lens model used in that work was based on a first limited sample of multiple images, when MUSE spectroscopy was not available. Limousin et al. (2007) estimated a scaling relation between halo sizes and σ_0 using galaxy-galaxy weak lensing in 5 clusters at $z \sim 0.2$, which is broadly consistent with our results.

The consistent constraints we find on the central velocity dispersions and truncation radii of the cluster sub-halos from the independent analysis of three clusters suggest that we can combine Eqs. 2.57 and 6.8 to obtain an empirical relation between σ_0 and the total mass of member galaxies. By propagating the uncertainties derived from the posterior distributions of σ_0 and r_{cut}, we obtain the following M_{tot} - σ_0 relation:

$$M_{tot} = 3.5^{(4.6)}_{(2.6)} \times 10^{11}\,\mathrm{M}_{\odot} \left(\frac{\sigma_0}{220\,(\mathrm{km\,s^{-1}})}\right)^{4.43^{(4.45)}_{(4.38)}}, \tag{6.9}$$

where the parameter range refers to the 16^{th} and 84^{th} percentiles.

6.5.1 Sub-halo velocity dispersion function

In light of our results, we present a central velocity dispersion function of cluster members, which can be derived from the sub-halo component of the lens models. This was first presented, in the form of circular velocity function ($v_c = \sqrt{2}\sigma_0$, see

Section 2.5) in Grillo et al. (2015) for MACSJ0416 and later extended by B18 to the same sample of 3 clusters of the present study, including the hot-gas components. The new sub-halo velocity functions presented here are particularly relevant, since the new lens models incorporate a prior on the σ - L scaling relation based on measured velocity dispersions of a large number of cluster members.

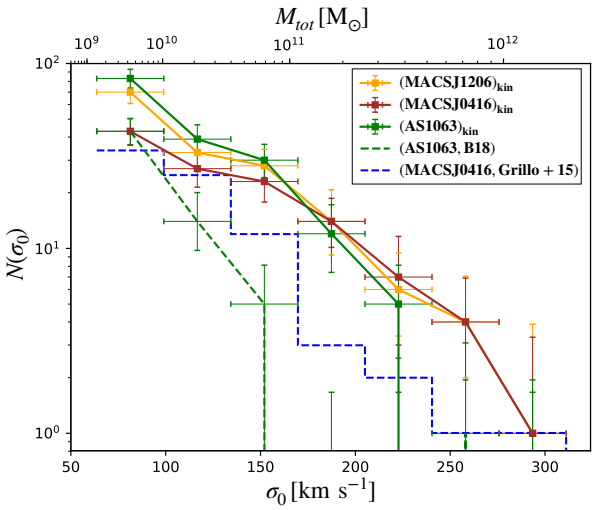


Figure 6.8: Sub-halo central velocity dispersion function derived from our lens models for MACSJ1206, MACSJ0416 and AS1063 with kinematics prior based on measured velocity dispersions of cluster members. Previous models with no kinematic prior for AS1063 and MACSJ0416 are shown for comparison. Data points correspond to median galaxy counts in 35 $\mathrm{km\,s^{-1}}$ bins, drawn from the posterior distributions of sub-halo model parameters. Errors also include poissonian number statistics in each bin. Only galaxies within a cluster-centric radius of $0.16 \times R_{200c}$ are included (see Table 6.1). The top axis gives the total mass of sub-halos related to the central velocity dispersion by Eq. 6.9 (see text). (Figure from Bergamini et al. 2019)

Each data point in Fig. 6.8 is obtained by extracting the central velocity dispersion of each member from 500 random realizations of the model posterior distributions, thereby computing the median number of members in each velocity bin 35 $\mathrm{km\,s^{-1}}$ wide. We adopt the same bin-size as in B18 to allow a direct comparison with previous results. The vertical error bars are obtained by combining the error associated to the 16^{th} and 84^{th} percentiles of these realizations with a Poissonian error generally appropriate for low number counts[6]. Note that in order to compare the velocity functions of the three clusters, we include only member galaxies with cluster-centric distances $R < 0.16 \times R_{200c}$ (see Table 6.1), which is the maximum

[6] $\Delta N^{up} = \sqrt{n+1} + 2/3$ and $\Delta N^{down} = \sqrt{n}$, where n is the number of objects per bin (Gehrels 1986).

aperture within which we have a highly complete sample of member galaxies in all three clusters.

In Fig. 6.8, we compare the new velocity functions with the previous determinations by Grillo et al. (2015) (for MACSJ0416, dashed blue line) and B18 (for AS1063, dashed green line). Not surprisingly, these previous determinations were biased low due to the lower normalization of the scaling relations in previous models, whereas the new lens models with kinematics prior produce velocity functions which are quite consistent with each other.

6.6 Conclusions

In this chapter, we have improved previous strong lensing models of three Hubble Frontier Fields/CLASH clusters, based on a large number of spectroscopically confirmed multiple images, using robust measurements of the internal stellar velocity dispersion of large samples of clusters galaxies, thanks to MUSE integral-field spectroscopy in cluster cores. Such measurements constrain independently the normalization and slope of the Faber-Jackson σ - L relation, so that the sub-halo component of the cluster mass distribution in the lens models is now bound to the kinematic measurements allowing a significant improvement on the reconstruction of the cluster internal sub-structure. The results of our study can be summarized as follows:

1. Using spectra of cluster early-type galaxies with a mean S/N of 25, we measure robust velocity dispersions of 37, 49 and 58 members in AS1063, MACSJ0416, MACSJ1206, whose accuracy is tested with the extensive spectral simulations described in Section 5.4. By obtaining kinematics measurements over 4 - 5 magnitudes in each cluster, down to $m_{F160W} \simeq 21.5$ which corresponds to $\sim$2.5 mag below L^*, we sample well the Faber-Jackson relation. A maximum likelihood modeling of this relation yields well consistent normalizations and slopes for the three clusters.

2. The new lens models incorporating kinematics information of cluster galaxies reproduce the positions of multiple images with similar precision to previous models. While the total projected cluster mass profile remains essentially unchanged, the mass of sub-halos are now robustly constrained, thus reducing degeneracies with other mass components and parameters of the lens model.

3. The inherent degeneracy of lens models between the central velocity dispersion

(σ_0) and truncation radius (r_{cut}) of sub-halos is strongly reduced, thus providing robust estimates of sub-halo masses and sizes. Once normalized to the same absolute luminosity, the three σ_0 - r_{cut} scaling relations, independently derived for each cluster, are found in very good agreement. This is particularly interesting in consideration of the different dynamical states of the three systems. As a result, we obtain a statistical determination of sub-halo truncation radii with masses ranging from that of the BCG down to $\sim 10^{10}\,M_{\odot}$. Our findings extend the initial results derived by Monna et al. (2015) in A383, by using now a high-quality data set which includes precision lens models based on large number bona-fide multiple images and high S/N kinematic measurements of cluster members.

4. With such a robust determination of the scaling relations of the sub-halo populations, we infer fully consistent velocity dispersion functions for the three clusters, unlike in previous lens models. In addition, our new constraints on the sub-halo masses provide an empirical M_{tot} - σ_0 relation, which can be used to translate the velocity functions into sub-halo mass functions.

After this thesis, we plan to extend this methodology to a larger sample of CLASH clusters with strong lensing models based on MUSE observations (Caminha et al. 2019) to better explore the role of systematics in parametric lens models when constraining different mass components of galaxy clusters. Further constraints on the mass profile of cluster galaxies, including the dark matter fraction, will require the use of galaxy-scale strong lensing systems (e.g. Suyu & Halkola 2010, Grillo et al. 2014, Monna et al. 2015), in combination with the internal kinematics of the lenses to extend in dense environments the extensive work carried out in field early-type galaxies. In the next chapter we will study three of such systems found in MACSJ1206 and Abell 2163. A complementary analysis based on the statistics of galaxy-galaxy strong lensing events will also be presented in Chapter 8 and in an upcoming paper by Meneghetti et al. (2019).

Chapter 7

Galaxy scale strong lensing systems in clusters

Cluster member galaxies act as small scale, secondary, gravitational lenses embedded in the deep potential wells of cluster-scale halos. If these galaxies are compact and massive enough, they can produce systems of multiple images on a scale of few arcseconds around their critical lines. These systems are commonly known as Galaxy Scale Strong Lensing Systems (hereafter GSSLSs).
Multiply lensed images around GSSLSs allow us to develop precise strong lensing models of cluster member galaxies, which accurately constrain their total mass within their Einstein radius, as well as to larger radii when combined with the global lens model of the hosting cluster.

In this section, we present a detailed study of the mass distribution of four cluster member galaxies belonging to three GSSLSs.
The first GSSLS consists of two sources at $z_S = 3.753$ in MACSJ1206 which are imaged five times by two close member galaxies (G1 and G2, see panel **a** of Fig. 7.1). This system, which in the following will be referred to as 'G1/G2', was first studied by (Grillo et al. 2015) when the MUSE data were not available.
The second GSSLS is a single cluster galaxy in MACSJ1206 which lenses a single source at $z_S = 4.996$ into four multiple images in an Einstein cross configuration (see panel **a** of Fig. 7.2). We will call this second GSSLS, the 'EC' system (EC stands for Einstein Cross).
The third GSSLS belongs to the galaxy cluster Abell 2163 (hereafter A2163) at $z_L = 0.2008$. In this system, a single massive cluster member produces four images of a background spiral galaxy at $z_S = 1.1628$ (see panel **a** of Fig. 7.3). We will refer to this system as 'SG' (Spiral Galaxy).

In keeping with the results of the previous chapter, we will show how the degeneracies among the lens model parameters can be reduced when the velocity dispersion of the lens galaxies is available, and even more when a crude velocity dispersion profile of the lens galaxy can be extracted from the seeing-limited MUSE data.

7.1 Description of the G1/G2, EC and SG systems

In panel **a** of Fig. 7.1, Fig. 7.2 and Fig. 7.3 we indicate with red boxes the positions of the G1/G2, EC and SG systems in the cluster fields.

G1/G2 system
The G1/G2 GSSLS was identified by Zitrin et al. (2012) in MACSJ1206, using HST CLASH data (see Section 4.2.1). The cluster galaxies G1 (ID 2541) and G2 (ID 3910, the second cluster BCG) lens two close compact sources at $z_S = 3.753$ into five multiple images (see Fig. 7.1). The source spectroscopic redshifts were measured by Caminha et al. (2017b) from the MUSE data. In connection with these sources, Caminha et al. (2017b) also identified a diffuse and extended Lyman-α emission (hereafter Lyα) clearly visible in panel **c** of Fig. 7.1. The figure shows a cutout of the continuum subtracted MUSE data-cube at an observed wavelength of $\lambda_{obs} = 5778$ Å, i.e. the Lyα wavelength at the source redshift, $z_S = 3.753$. A staked spectrum of the multiple imaged Lyα clouds is in Fig. 7.4.
A first lens model for the G1/G2 system was developed by Grillo et al. (2014) using CLASH and VIMOS observations. In this model a photometric redshift of $z_S = 3.7$ (Zitrin et al. 2012) was assumed for the multiple images. Moreover, the two galaxies G1 and G2 were parameterize as SIS profiles, while the cluster mass distribution was described through a single massive halo centered on the cluster BCG. Since the SIS profile is defined through the single free-parameter σ_0 and has a divergent total mass (see Section 2.5), this model cannot constrain on the truncation radius and on the total mass of the lens galaxies.

EC system
The EC system in MACSJ1206 is shown in Fig. 7.2. In this case, a single compact Lyα source at $z = 4.996$ is imaged four times, in the shape of an Einstein cross, by a relatively faint cluster galaxy (Gal.$\dot{9}323$) lying 67″(or 380 kpc) from the BCG. Compared to the G1/G2 GSSLS, the multiple images around the EC system do not have a detectable optical counterpart in the HST/CLASH photometric bands. For this reason, these multiple images were initially disregarded and not included in the cluster lens models developed by Caminha et al. (2017b) and Bonamigo et al. (2018).

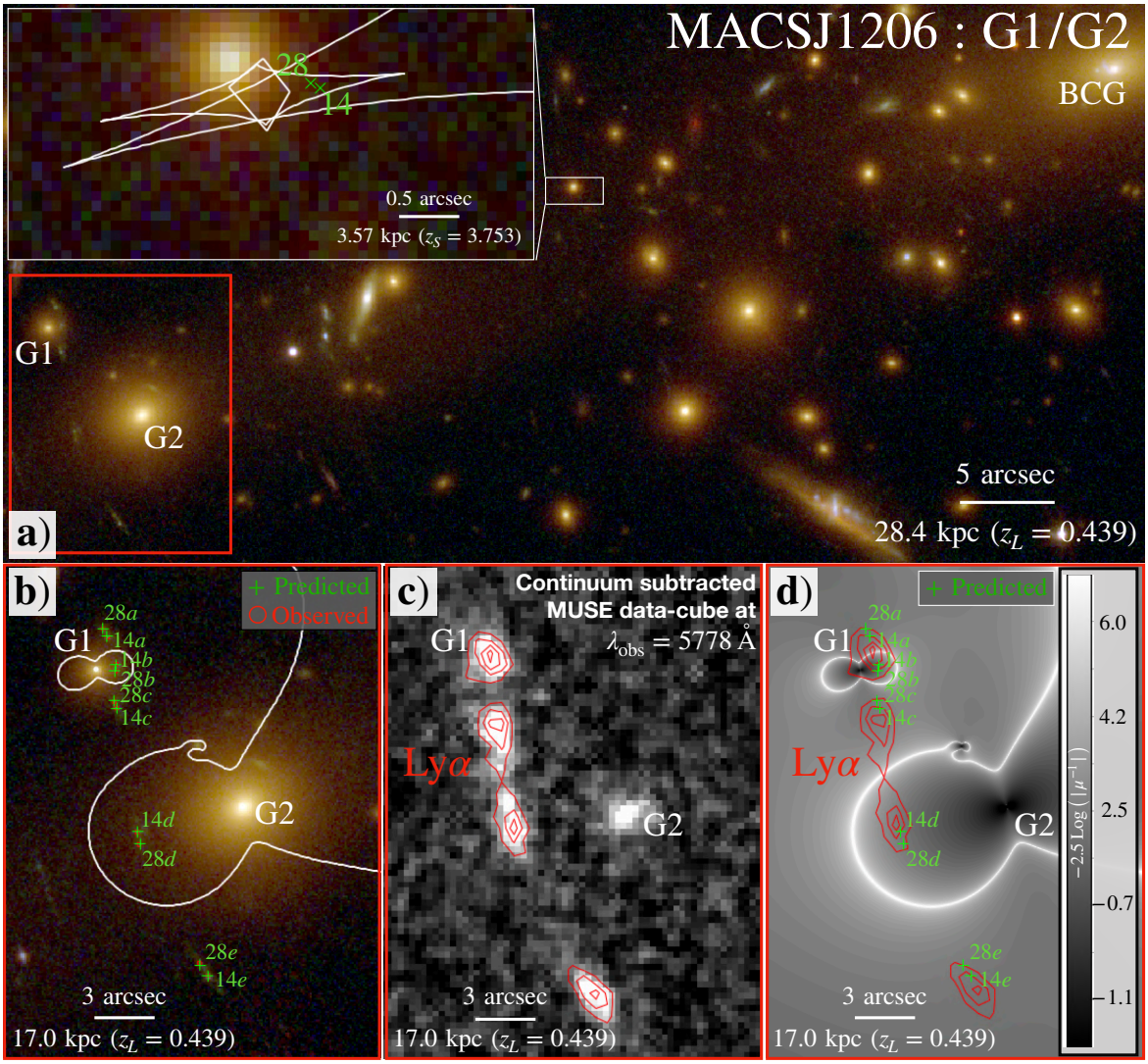


Figure 7.1: a): HST color composite image of the bottom-left region of MACSJ1206. The BCG, G1 and G2 galaxies are labelled in white. The red box encircles the position of the G1/G2 system. On the top-left corner we show a zoom-in on the region where the two imaged sources would be observed in absence of the cluster lens. The caustic lines at $z_S = 3.753$ are plotted in white, while the source positions are shown as green crosses. The distance on the source plane between the two sources is of about 506.4 pc. **b)**: Zoom-in on the G1/G2 GSSLS. The observed and predicted images positions are shown in green and red respectively. The critical lines are plotted in white. **c)**: Continuum subtracted MUSE data-cube at an observed wavelength of $\lambda_{obs} = 5778$ Å, i.e. the Lyα wavelength at the source redshift $z_S = 3.753$. We draw the contours of the Lyα emission in red. **d)**: Magnification map color-coded in $-2.5\mathrm{Log}\left(|\mu^{-1}|\right)$. The map gives the luminosity difference (expressed in absolute magnitude) between images (at $z_S = 0.439$) and sources ($z_S = 3.753$).

In panel **c** of Fig. 7.2, we show the continuum subtracted MUSE data-cube at $\lambda_{obs} = 7289\,\text{Å}$ (i.e. the Lyα line at the source redshift, $z_s = 4.996$). The Lyα Einstein cross has four well resolved peaks, highlighted by the red contours in the figure, which are used as positions of the multiple images. In Fig. 7.4, a spectrum for the lensed source is shown by stacking all spaxels associate to the Lyα emission.

SG system

The third GSSLS studied in this chapter, the SG system in A2163 at $z = 0.2008$, can be easily identified in the HST images obtained as part of the Reionization Lensing Cluster Survey (RELICS) survey (P.I. Coe, Coe et al. 2019). The latter is a recent HST campaign targeting 41 among the most massive *Planck* clusters. RELICS observations are carried out with the same 7 ACS and WFC3 filters used in the HFF campaign (for details see Cerny et al. 2018). A2163, which was observed for a total of 8 HST orbits, is the most massive and X-ray luminous cluster of the RELICS sample, with $M_{500} = 1.5 \times 10^{15}\,\text{M}_{\odot}$ (Planck Collaboration et al. 2016) and $L_X^{[2\text{-}10]\,keV} = 6.0 \times 10^{45}\text{erg}\,\text{s}^{-1}$ (Elbaz et al. 1995; Markevitch et al. 1996). Radio observations carry out by Herbig & Birkinshaw (1994) confirm that A2163 has one of the most extended cluster halos ever detected. An optical study by Maurogordato et al. (2008) suggests that the core of A2163 has undergone a recent merger between two main sub-clusters, while a third large halo is infalling into the cluster core from its outskirts. This merging scenario is also confirmed by weak lensing observations (Soucail 2012; Okabe et al. 2011).

A2163 was observed with MUSE in 2014 as part of the instrument commissioning program (ID 60.A-9100(C)). This observation provided a single MUSE pointing, for a total of 4 hour exposure, centered on the cluster BCG and containing also the SG system. The average seeing during the data acquisition was smaller than 1″. Rescigno et al. (2019), submitted have recently used these data to identify multiple images and develop an accurate lens model (see below).

The SG system consists of a single extended source at $z_S = 1.1628$ lensed four times by a relatively bright cluster galaxy (Gal. 4104), as it can be appreciated in the panel **a** of Fig. 7.3. The lensed source has a prominent [OII], as apparent in the the continuum subtracted MUSE data-cube at $\lambda_{obs} = 8063\,\text{Å}$ (i.e. the wavelength of the [OII] doublet at the source redshift, $z_S = 1.1628$), see panel **c** of Fig. 7.3 and the corresponding spectrum in Fig. 7.4.

By accurately measuring the wavelengths of the [OII] lines in the data-cube, we detect a smooth shift consistent with a rotation pattern for the lensed source. This shows that the lensed galaxy is a nearly edge-on spiral at $z_S = 1.1628$ (see Section 7.3).

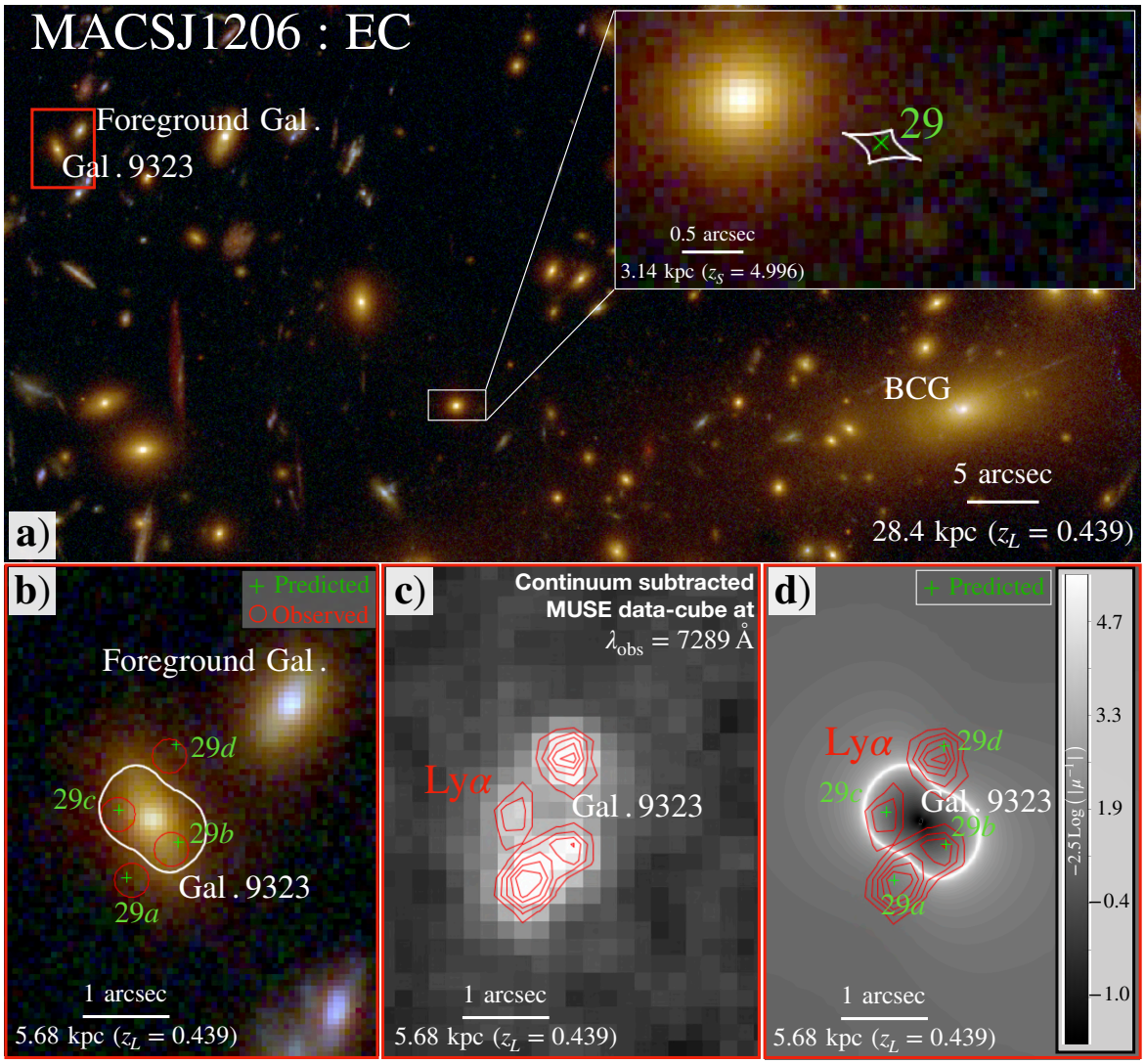


Figure 7.2: Same as Fig. 7.1 but for the EC system. In panel **c** we show the continuum subtracted MUSE data-cube at $\lambda_{obs} = 7289\,$Å. This is the wavelength of the Lyα line at the source redshift, $z_S = 4.996$.

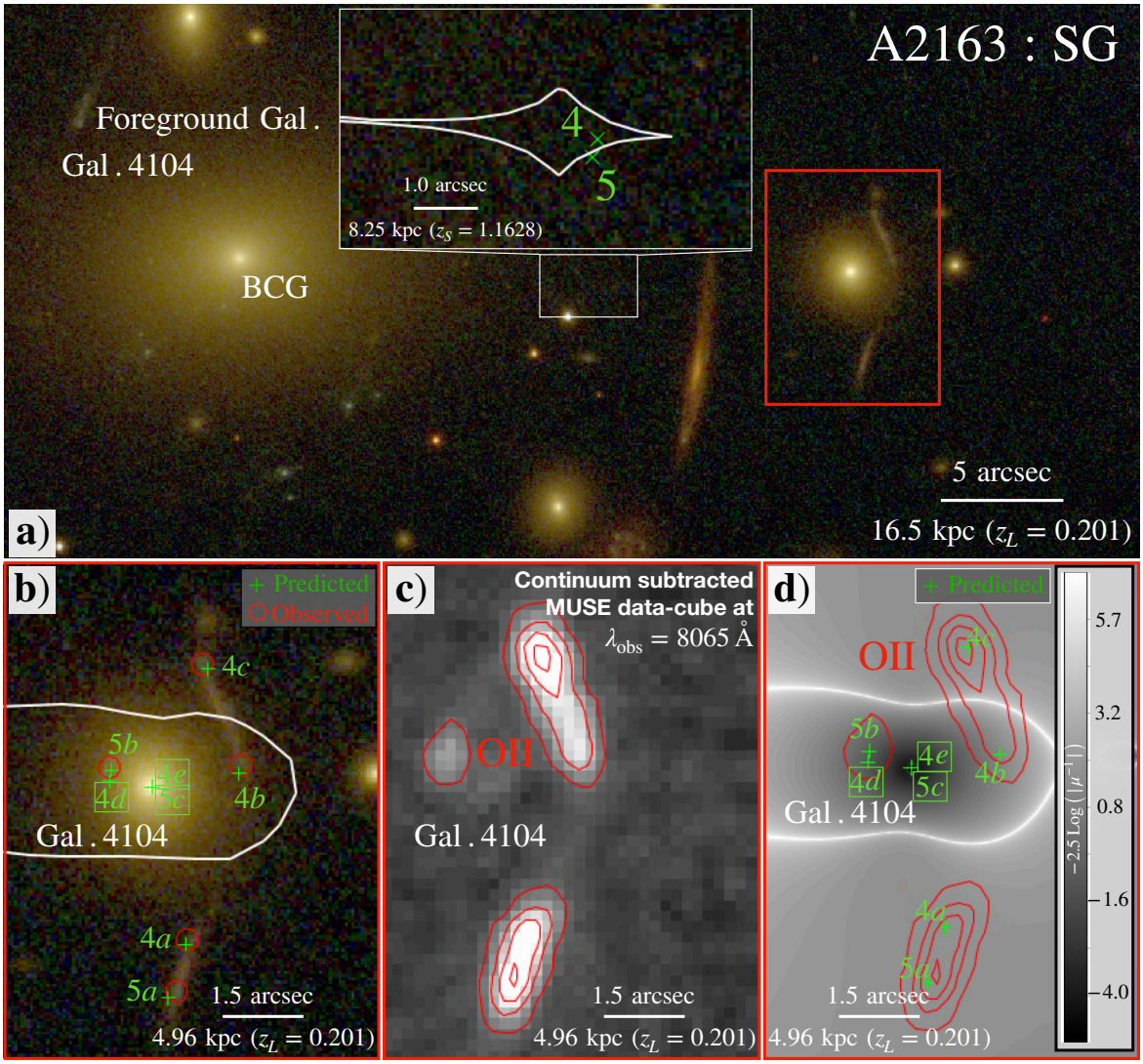
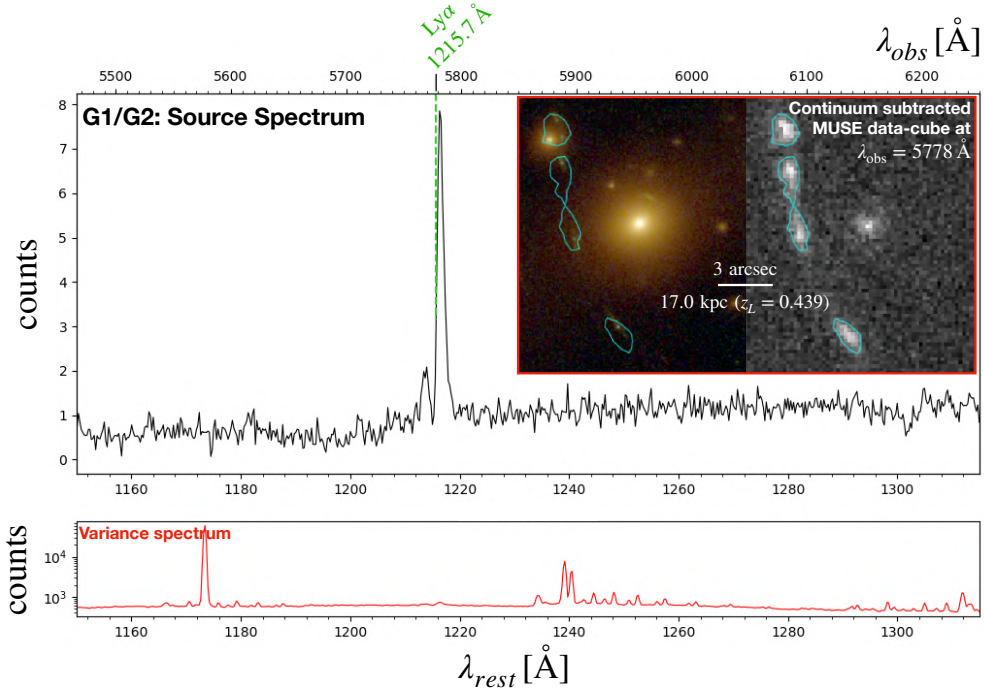
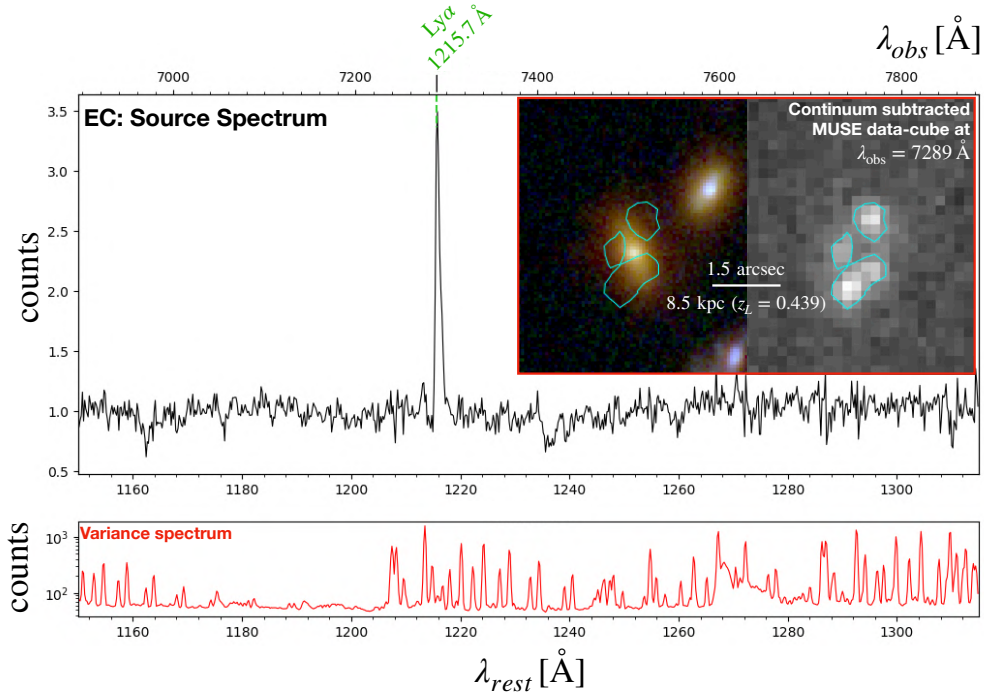


Figure 7.3: Same as Fig. 7.1 but for the SG system. In panel **c** we show the continuum subtracted MUSE data-cube at $\lambda_{obs} = 8065\,$Å. This is the wavelength of the [OII] doublet at the source redshift, $z_S = 1.1628$.


Lyα
1215.7 Å
λ_{obs} [Å]
G1/G2: Source Spectrum
Continuum subtracted
MUSE data-cube at
λ_{obs} = 5778 Å
3 arcsec
17.0 kpc (z_L = 0.439)
counts
Variance spectrum
counts
λ_{rest} [Å]


Lyα
1215.7 Å
λ_{obs} [Å]
EC: Source Spectrum
Continuum subtracted
MUSE data-cube at
λ_{obs} = 7289 Å
1.5 arcsec
8.5 kpc (z_L = 0.439)
counts
Variance spectrum
counts
λ_{rest} [Å]

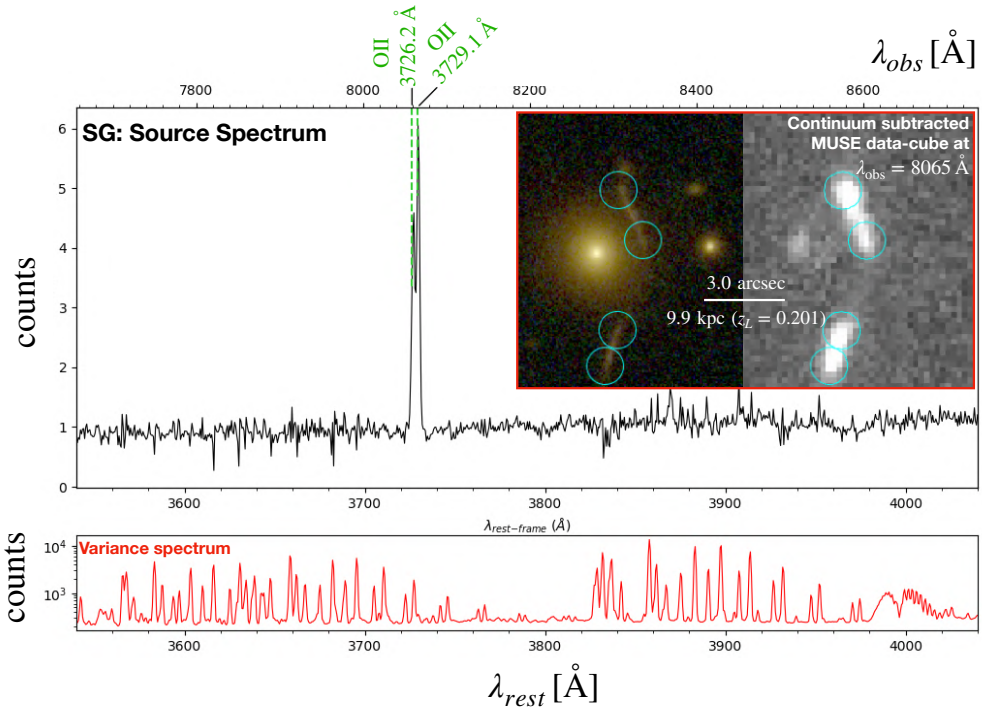


Figure 7.4: Top panel: spectrum and morphology of the diffuse Lyα emission associated to the multiple images in the G1/G2 system. The two cutouts show the HST color composite image and the continuum subtracted MUSE data-cube around the wavelength of the relevant emission line. The source spectrum is extracted by combining all the MUSE spaxels inside the blue emission line contours. The variance spectrum is shown at the bottom in red. Middle panel: the spectrum of the lensed source of the EC system, associated to the Lyα emission of the Einstein quad images. Bottom panel: the spectrum of the extended background source in the SG system in A2163, dominated by the [OII] emission doublet.

7.2 Velocity dispersion measurements and σ profiles

In Chapter 6, we used pPXF to measure the stellar velocity dispersion of cluster members, from MUSE spectra extracted within circular apertures. Applying the same technique, we measure here the stellar velocity dispersion of the three cluster galaxies G2, Gal. 9323 and Gal. 4104 belonging to the G1/G2, EC and SG systems respectively. The G1 galaxy is excluded from our analysis since its measured velocity dispersion ($\sim 65\ \mathrm{km\,s^{-1}}$) is lower than the MUSE spectral resolution and therefore not as robust (see Section 5.4).

For Gal. 9323, we choose a spectral aperture of $0.8''$ radius. As in Chapter 6, the pPXF measurement is performed within the 3600-4900 Å rest-frame wavelength range and considering the stellar LOSVD up to the fourth moment. In Fig. 7.5, we show in black the observed galaxy spectrum and in red the pPXF best fit model. We measure a velocity dispersion of $\sigma_{ap} = (115.9 \pm 14.2)\,\mathrm{km\,s^{-1}}$ and a mean spectral signal-to-noise of $\langle S/N \rangle = 13.9$.

Unlike Gal. 9323, the two galaxies G2 and Gal. 4104 are sufficiently bright and extended to allow a measurement of their velocity dispersion profiles within three spectral apertures of increasing sizes. The radii defining the inner circular apertures and external annular apertures are quoted in each cutout of Fig. 7.5. The external and internal radii are chosen such that the aperture widths are always larger than the PSF (seeing) of MUSE observations. For the pPXF measurements, we use wavelength ranges of 3600-4900 Å and 3975-6300 Å for G2 and Gal. 4104 respectively. In Fig. 7.5, we show the pPXF fit results in each radial bin for Gal. 9323, G2 and Gal. 4104. In each panel, we show the RGB cutouts, corresponding spectral apertures (in green). When extracting the spectra for G2, contaminating sources are masked out (see red circles). The spectral mean S/N and measured velocity dispersions in each radial bin are also quoted in the figure. These measurements are shown as data points in Fig. 7.10, indicating a decreasing velocity dispersion profile for G2 and gal.4104, and will be used to constrain the specific lens models which we describe in the next session.

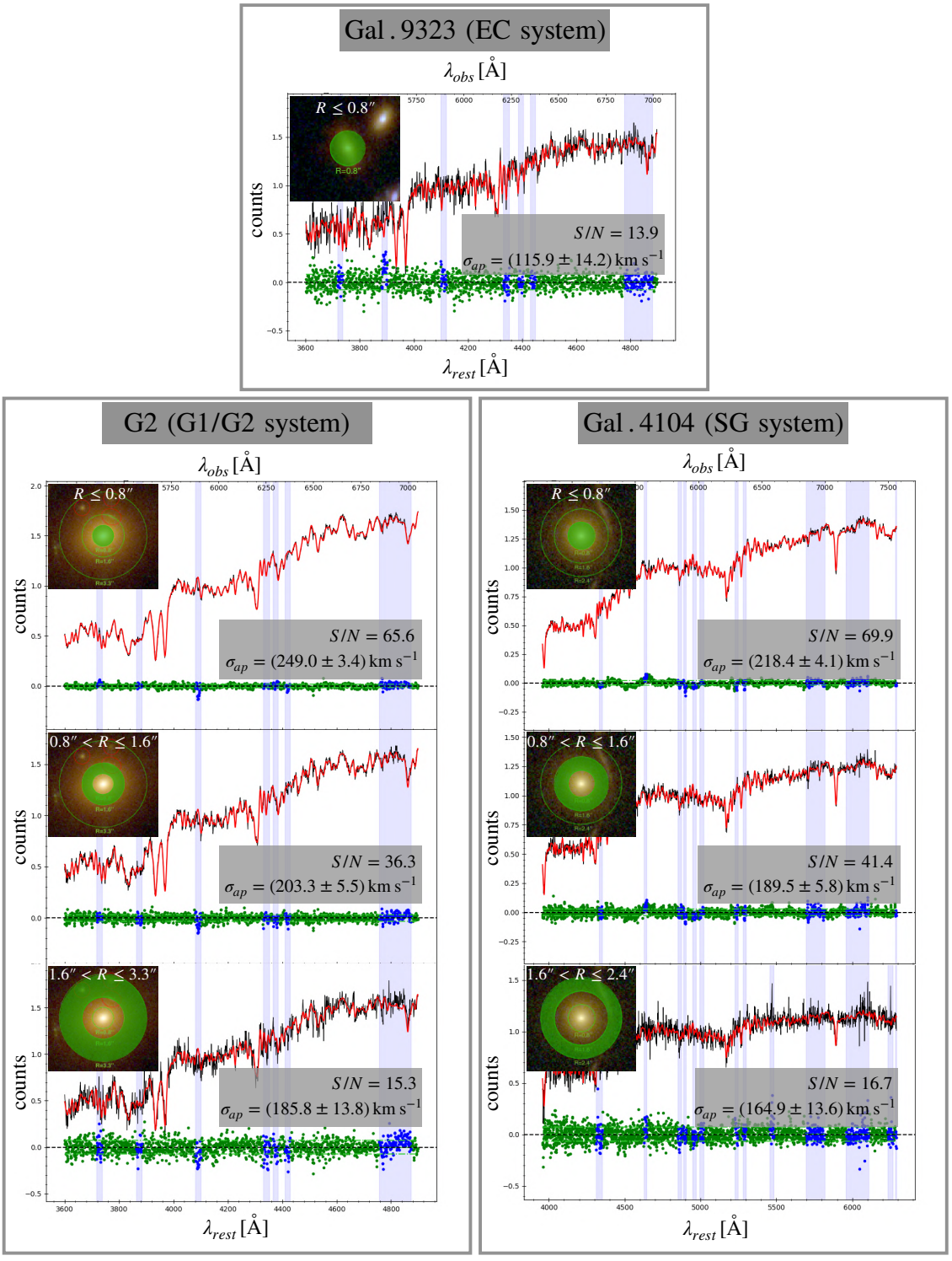


Figure 7.5: pPXF fits to the Gal 9323, G2 and Gal. 4104 cluster galaxy spectra. The Gal. 9323 spectrum is extracted in a circular aperture of radius $R = 0.8''$, while for G2 and Gal. 4104 we consider three ring apertures of radii between: $[(0''\text{-}0.8''), (0.8''\text{-}1.6''), (1.6''\text{-}3.3'')]$ for G1 and $[(0''\text{-}0.8''), (0.8''\text{-}1.6''), (1.6''\text{-}2.4'')]$ for Gal. 4104. For a description of the plots see Fig. 6.2.

7.3 Strong lens models

G1/G2 and EC systems

Extending the work of Grillo et al. (2014), we use the complete MACSJ1206 lens model (with kinematic priors) developed in Chapter 6 to accurately constrain the cluster mass profile in the vicinity of G1 and G2. We assume the same input configuration of Table 6.3, with identical uniform priors on the cluster-scale halo parameters and the kinematic Gaussian prior on the σ-L scaling relation. However, the three galaxies G1 (ID 3910), G2 (ID 2541) and Gal. 9323 are now modeled and optimized outside the scaling relations as separated dPIE sub-halos. In particular, we assume two spherical dPIE mass distributions for the galaxies G1 and G2 and an elliptical dPIE for the Gal. 9323 (EC system). The three dPIEs have a negligible core radius of 0.001″ (5.7 pc at the cluster redshift), while the values chosen for the other parameters are listed in Table 7.1 (upper table).

The observed positions of the ten point-like multiple images belonging to the system G1/G2, and associated to two close sources identified with the indices 14 and 28, are determined from the HST images (see panel **b** of Fig. 7.1). We assign to these positions an error of 0.065″, equal to the HST pixel size. Instead, the four multiple image positions of the EC system (family 29 in the lens model) are computed from the peaks of the Lyα emissions in the continuum subtracted MUSE data-cube (see panel **c** of Fig. 7.2). For these images we assume a conservative error of 0.2″, equal to about 1/4 of the MUSE PSF (panel **b** of Fig. 7.2). The sky-coordinates of the observed multiple images and the position error associated to each image are listed in Table 7.2.
The final "MACSJ1206 + G1/G2 + EC" lens model is optimized using the whole set of multiple images described in Chapter 6, with the addition of the extra images belonging to the G1/G2 and EC systems.

SG system

The first A2163 lens model was developed by Cerny et al. (2018). Despite spectroscopic observations through the Low Dispersion Survey Spectrograph (LDSS3-C18) on the Magellan Clay telescope were attempted for this cluster, Cerny et al. (2018) were unable to measure spectroscopic redshifts for any observed multiple images. A significant improvement of this model was made by Rescigno et al. (2019, submitted), exploiting the MUSE observations described above. This model is optimized using the positions of 16 spectroscopically confirmed multiple images

with $1.16 < z < 2.72$ and F814W magnitudes between 24 and 29. Rescigno et al. model the mass distribution with a single cluster-scale halo, parametrized as a PIEMD (see Eq. 2.44). A second circular dPIE profile, with $r_{core} = 0.01''$, is used to model the Gal. 4104 cluster member (i.e. the lens of the SG system). The sub-halo clumpy component of the cluster contains 110 cluster galaxies (28 spectroscopically confirmed) parametrized as circular dPIE profiles of negligible core radii. Their velocity dispersions and truncation radii vary according to the two scaling relations in Eqs. 3.28 and 3.29, with $\alpha = 0.35$ and $\beta_{cut} = 0.5$.
We adopt the same mass parameterization to model the SG system in this chapter. In Table 7.1 (upper table), we show the input parameter ranges chosen for the SG galaxy.

	INPUT dPIE parameters for the G1/G2, EC and SG systems						
	x [arcsec]	y [arcsec]	e	θ [°]	σ_{LT} [km s^{-1}]	r_{core} [arcsec]	r_{cut} [arcsec]
G1	−57.07	−13.85	0.0	0.0	[50.0; 150.0]	0.001	[0.1; 15.0]
G2	−52.07	−18.52	0.0	0.0	[150.0; 350.0]	0.001	[0.1; 15.0]
9323 (EC)	−64.59	18.44	[0.0; 0.8]	[0.0; 180.0]	[50.0; 350.0]	0.001	[0.1; 10.0]
4104 (SG)	24.41	−0.51	0.0	0.0	[80.0; 350.0]	0.01	[0.5; 15.0]

	OUTPUT dPIE parameters for the G1/G2, EC and SG systems						
	x [arcsec]	y [arcsec]	e	θ [°]	σ_{LT} [km s^{-1}]	r_{core} [arcsec]	r_{cut} [arcsec]
G1	−57.07	−13.85	0.0	0.0	$91.5^{+14.8}_{-8.2}$	0.001	$1.4^{+1.2}_{-0.6}$
G2	−52.07	−18.52	0.0	0.0	$242.2^{+51.2}_{-34.4}$ ($219.2^{+5.3}_{-4.7}$)	0.001	$2.4^{+1.6}_{-1.4}$ ($3.2^{+0.4}_{-0.4}$)
9323 (EC)	−64.59	18.44	$0.36^{+0.27}_{-0.24}$	$103.9^{+41.6}_{-41.7}$	$107.4^{+24.7}_{-10.9}$ ($106.3^{+19.0}_{-10.1}$)	0.001	$3.6^{+4.3}_{-2.9}$ ($3.9^{+4.1}_{-3.0}$)
4104 (SG)	24.41	−0.51	0.0	0.0	$225.5^{+25.5}_{-28.2}$ ($213.2^{+6.4}_{-7.9}$)	0.01	$0.9^{+0.6}_{-0.3}$ ($1.5^{+0.2}_{-0.1}$)

Table 7.1: *Top:* Input dPIE parameters for the three cluster galaxies G1 (ID 3910), G2 (ID 2541) and Gal. 9323 of MACSJ1206 and the galaxy Gal. 4104 of A2163. Sky x, y coordinates are the offsets in arcsec from the reference BCG positions (MACSJ1206: $R.A. = 12^h06^m12^s.15$, $DEC = -8°48'03''.4$; A2163: $R.A. = 16^h15^m48^s.95$, $DEC = -06°08'41''.4$). The ellipticity e is defined as $e = \frac{a^2-b^2}{a^2+b^2}$, where a and b are the semi-major and semi-minor axis. The position angle θ is computed from East to North. σ_{LT} is the `LensTool` fiducial velocity dispersion, while r_{core} and r_{cut} are the core and truncation radii respectively. If a flat prior is assumed for a parameter, the interval for the prior is quoted in square brackets. Single values are fixed. *Bottom:* Optimized output parameters for the lens galaxies. We quote the median, and the 16^{th} and 84^{th} percentiles as errors. We use the measured velocity dispersion profiles of the two galaxies G2 and Gal. 4104, and the single aperture velocity dispersion of Gal. 9323, to further reduce the lens model degeneracy between σ_{LT} and r_{cut} (see Section 7.4). The σ_{LT} and r_{cut} values obtained in this way are in red.

In Fig. 7.6, we show a velocity map derived from the [OII] emission of the extended lensed images of the SG system. This map is obtained binning into 2x2 pixels the MUSE datacube and measuring the wavelength shift of the [OII] doublet. The final velocity map is re-interpolated on the MUSE spatial sampling (0.2″). To measure the wavelength shift, we exploit the MUSE Python Data Analysis Framework[1] (`mpdaf`). Since `LensTool` cannot perform a surface brightness reconstruction to constrain the lens model, we use the velocity map to identify the regions of the extended multiple images that are mapped into a same area of the source. In particular, we define the multiple image family 4 which is mapped into a source region with a velocity of -35 ± 3 km s^{-1}, and similarly family 5 corresponding to a velocity of 10 ± 3 km s^{-1}. We associate a single multiple image, 5b, to the [OII] emission 1.5″ NE of Gal. 4104 (see panel **b** of Fig. 7.3). The sky-coordinates of the five images belonging to families 4 and 5 and corresponding positional errors are in Table 7.2. The SG lens model is optimized using the positions of these five images, plus the 12 extra multiple images found by Rescigno et al. (2019, submitted) (note that we have redefined the multiplicity of family 4 in Rescigno's model using the velocity map).

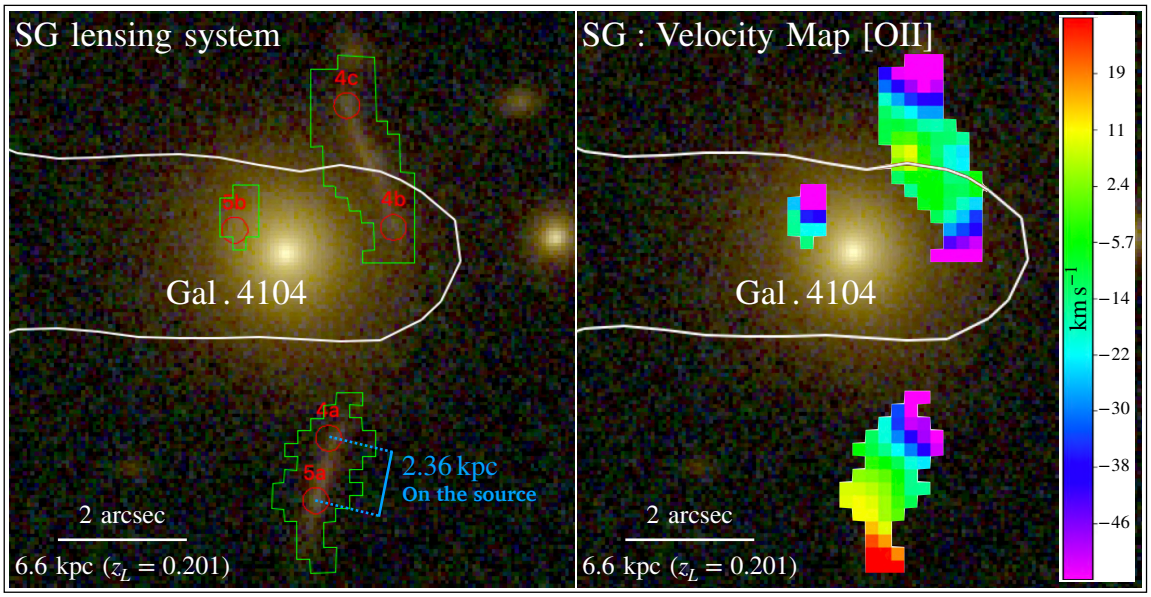


Figure 7.6: *Right:* HST color composite image of the SG system. Superimposed color-coded velocity maps derived form the [OII] emission of the lensed source at $z_S = 1.1628$. The white curve is the critical line (computed for the source redshift) from the best-fit lens model of the SG system. *Left:* Only the contours of the velocity maps are shown in green. In red we indicate the observed multiple image positions.

[1] https://mpdaf.readthedocs.io/

7.4 Strong lens model results

In the last section, we have described our strong lens models for the three systems: G1/G2, EC and SG. Unlike previous works (e.g. Grillo et al. 2014), we optimize the free parameters of the GSSLS lenses together with the complete mass distribution of clusters, including all member galaxies. Moreover, our models are optimized considering the whole sets of multiple images observed in the two clusters.
This procedure ensures an accurate and robust description of cluster mass distributions in the vicinity of the GSSLSs.

In this section, we summarize the main results obtained studying the G1/G2, EC and SG systems.

7.4.1 Multiple images positions and magnifications

In the fifth column of Table 7.2, we quote the root-mean-square separations between observed and model predicted multiple image positions for the three GSSLSs.
In particular, the ten compact images belonging to the G1/G2 system have total Δ_{rms} of $0.03''$ and $0.05''$ for the 14^{th} and 28^{th} families respectively. These values correspond roughly to the 46% and 77% of the position error assumed on observed positions (i.e. $\Delta x = 0.065''$). As recently pointed out by Birrer & Treu (2019), with such small Δ_{rms} the accuracy of the astrometric calibration of the HST images becomes also important.
For the Einstein cross system (EC), we obtain a total $\Delta_{rms} = 0.11''$ (55% of the observed position error), while families 4 and 5 of the SG system have root-mean-square separations of $\Delta_{rms} = 0.13''$ and $\Delta_{rms} = 0.15''$ respectively (i.e. the 65% and 75% of observed position errors).
In panel **b** of Figs. 7.1, 7.2 and 7.3, we show in red the observed multiple image positions with their errors, while the green crosses indicate the model predicted image positions. Note that the SG lens model predicts three multiple images without an observed counterpart (green squares in Fig. 7.3). Two of these images are close to the core of the lens galaxy (Gal. 4104) and are strongly de-magnified (see panel **d** of the same figure), while the third image is indistinguishable from the observed image 5*b* due to the MUSE PSF of about $1.0''$. The critical lines obtained from the best-fit lens models are shown in white.

In panel **a** of Figs. 7.1, 7.2 and 7.3, we indicate the source positions (green crosses) by ray-tracing the positions of the observed multiple images back to their source planes. The white curves are the caustic lines corresponding to the critical

Multiple images of the G1/G2 system					
ID	**R.A.**	**Dec.**	**Δx [arcsec.]**	**$\Delta_{\rm rms}$ [arcsec]**	**$-2.5\,{\rm Log}\left(\lvert\mu^{-1}\rvert\right)$**
14a	$12^h06^m15^s.9736$	$-8°48'16''.118$	0.065	0.04	2.34
14b	$12^h06^m15^s.9535$	$-8°48'17''.032$	0.065	0.03	1.92
14c	$12^h06^m15^s.9539$	$-8°48'18''.519$	0.065	0.03	2.49
14d	$12^h06^m15^s.9061$	$-8°48'22''.765$	0.065	0.04	2.29
14e	$12^h06^m15^s.7420$	$-8°48'27''.669$	0.065	0.02	2.28
Tot. (14)				**0.03**	
28a	$12^h06^m15^s.9851$	$-8°48'15''.821$	0.065	0.01	2.17
28b	$12^h06^m15^s.9547$	$-8°48'17''.264$	0.065	0.04	1.76
28c	$12^h06^m15^s.9578$	$-8°48'18''.311$	0.065	0.04	2.49
28d	$12^h06^m15^s.9014$	$-8°48'23''.184$	0.065	0.05	2.31
28e	$12^h06^m15^s.7620$	$-8°48'27''.277$	0.065	0.09	2.39
Tot. (28)				**0.05**	

Multiple images of the EC system					
ID	**R.A.**	**Dec.**	**Δx [arcsec.]**	**$\Delta_{\rm rms}$ [arcsec]**	**$-2.5\,{\rm Log}\left(\lvert\mu^{-1}\rvert\right)$**
29a	$12^h06^m16^s.5229$	$-8°47'45''.641$	0.2	0.07	1.91
29b	$12^h06^m16^s.4915$	$-8°47'45''.282$	0.2	0.12	1.16
29c	$12^h06^m16^s.5337$	$-8°47'44''.871$	0.2	0.05	1.33
29d	$12^h06^m16^s.4931$	$-8°47'44''.184$	0.2	0.15	1.67
Tot. (29)				**0.11**	

Multiple images of the SG system					
ID	**R.A.**	**Dec.**	**Δx [arcsec.]**	**$\Delta_{\rm rms}$ [arcsec]**	**$-2.5\,{\rm Log}\left(\lvert\mu^{-1}\rvert\right)$**
4a	$16^h15^m47^s.2653$	$-6°08'44''.788$	0.2	0.09	1.75
4b	$16^h15^m47^s.1980$	$-6°08'41''.527$	0.2	0.16	1.25
4c	$16^h15^m47^s.2464$	$-6°08'39''.642$	0.2	0.14	1.85
Tot. (4)				**0.13**	
5a	$16^h15^m47^s.2786$	$-6°08'45''.758$	0.2	0.20	1.75
5b	$16^h15^m47^s.3628$	$-6°08'41''.570$	0.2	0.06	-0.88
Tot. (5)				**0.15**	

Table 7.2: Tables show the observed positions of the multiple images belonging to the galaxy scale strong lensing systems G1/G2, EC and SG. Δx are isotropic errors on observed positions. In the fifth column, we show the root-mean-square separation, $\Delta_{\rm rms}$, between observed and model predicted image positions. The total $\Delta_{\rm rms}$ of every image family is quoted in the last line of each sub-table. In the last column, we show the absolute magnifications of the multiple images (expressed in magnitudes).

lines in **b** panels.

As already mentioned, the ten multiple images of the G1/G2 system come from two close compact sources at $z_S = 3.753$, which are $0.071''$, or $\sim 500\,{\rm pc}$ apart on the

source plane, based on our bet-fit lens model. A similar GSSLS belonging to the cluster MACSJ0416 was studied by Vanzella et al. (2017). In that case, two cluster member lenses produce four multiple images of two close ($\sim 300\,$pc apart) compact sources at a redshift $z_S = 3.2222$.

Similarly, by mapping families 4 and 5 of the SG system onto the source plane, we obtain a size of the rotating lensed source of $\sim 0.3''$, i.e. 2.4 kpc at the source redshift $z_S = 1.1628$.

Finally, the **d** panels of Figs. 7.1, 7.2 and 7.3 show the magnification maps obtained for our GSSLSs. The maps are color-coded according to $-2.5\,\mathrm{Log}\,(|\mu^{-1}|)$, with the magnification μ computed as in Eq. 3.21, so to express the magnification directly in magnitudes. In the last column of Table 7.2, we show the magnification of every multiple image belonging to the GSSLSs. The magnifications are computed from the best-fit lens models. For the G1/G2, EC and SG systems we find an average magnification for the observed multiple images of 2.24 magnitudes ($\mu = 7.90$), 1.52 magnitudes ($\mu = 4.05$) and 1.65 magnitudes ($\mu = 4.57$), respectively (only the HST detected images of the SG system, i.e. 4a, 4b, 4c and 5a, are considered to compute the average magnification).

7.4.2 Lens mass parameters and inclusion of stellar velocity dispersion measurements

The marginalized posterior distributions of the lens model free parameters for G1, G2, Gal. 9323 and Gal. 4104 are showed in Fig. 7.7, 7.8 and 7.9 (black contours and histograms), where we also indicate the medians and the 16^{th} and 84^{th} percentiles of these distributions.

Despite the significant number of multiple images around the GSSLSs, a large uncertainty in the determination of galaxy velocity central dispersion values and truncation radii is still present. This is essentially due to the σ_0-r_{cut} degeneracy discussed in Sec. 3.7.1. In Chapter 6, we demonstrated how a similar degeneracy between the reference scaling relation parameters σ_{ap}^{ref} and r_{cut}^{ref} can be partially removed introducing priors derived from the measured cluster member stellar kinematics. Here instead, the single measured stellar velocity dispersion of the galaxy Gal. 9323 and the σ profiles of G2 and Gal. 4104 (see Section 7.2) are used to select, a posteriori, the `LensTool` MCMC chains that are compatible with the galaxy stellar kinematics (within a certain confidence region).

The agreement between measured stellar velocity dispersions and lens model

results is quantify by the following χ^2_{kin} function:

$$\chi^2_{kin} = \sum_{i}^{N_{ring}} \frac{\left[\sigma_i^{lens}(R_i^{in}, R_i^{ex}) - \sigma_i^{pPXF}(R_i^{in}, R_i^{ex})\right]^2}{\left[\delta\sigma_i^{pPXF}(R_i^{in}, R_i^{ex})\right]^2}. \tag{7.1}$$

$\sigma_i^{pPXF}(R_i^{in}, R_i^{ex})$ is the measured stellar velocity dispersion within a radial bin of internal radius R_i^{in} and external radius R_i^{ex} and $\delta\sigma_i^{pPXF}(R_i^{in}, R_i^{ex})$ is its error. $\sigma_i^{lens}(R_i^{in}, R^{ex})$ is the predicted velocity dispersion inside the same radial bin, computed using Eq. 2.64, for every combination of `LensTool` MCMC chain parameters.

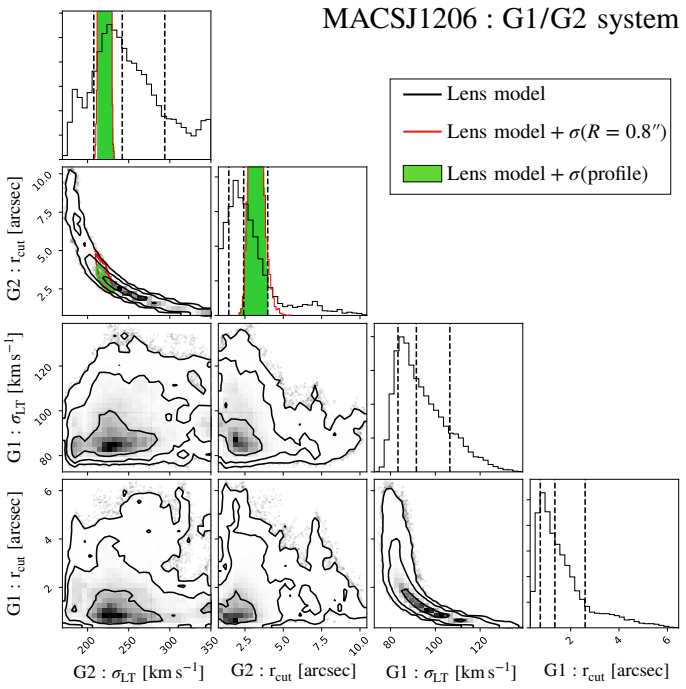


Figure 7.7: The marginalized posterior distributions for the lens model parameters (σ_{LT}, r_{cut}) of the two galaxies G1 and G2 are shown in black. The vertical dashed lines indicate the 16^{th}, 50^{th} and 84^{th} percentiles of the distributions. The green contours and distributions (re-normalized for clarity) are obtained by selecting only the `LensTool` MCMC chains which are consistent, a 95% confidence level, with the measured stellar velocity dispersion profile of G2. For comparison, we plot in red the posterior distributions which are 95%-compatible with the measured stellar velocity dispersion within a single circular aperture of $R = 0.8''$ centered on G2 (see Section 7.4.2).

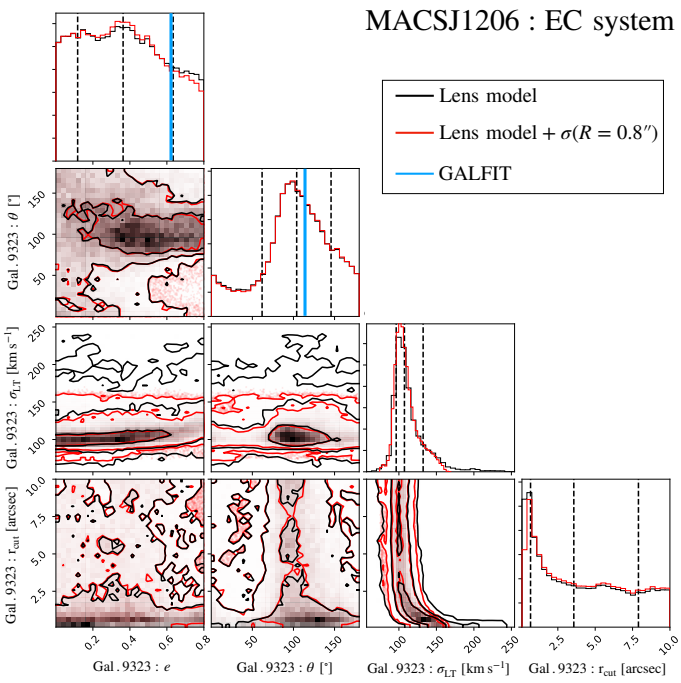


Figure 7.8: Same as Fig. 7.7 for the member Gal. 9323 (Einstein cross system) in MACSJ1206, modeled as an elliptical dPIE. The vertical blue lines are the ellipticity and the position angle values obtained by fitting the light profile of the galaxy with the public software GALFIT (Peng et al. 2002, 2010), in the F160W band.

The G2 σ profile is sampled considering three annular apertures (see Section 7.2) from the MUSE data. By requiring that the `LensTool` model for G2 is consistent with these measurements, we select those MCMC chains for which $\chi^2_{kin} < 7.815$, using Eq. 7.1. The value 7.815 corresponds to the 95% confidence level of a χ^2 random variable with 3 degree of freedom (the number of measurements), so that the probability to obtain a $\chi^2_{kin} > 7.815$ is $P(\chi^2_{kin} > 7.815) = 0.05$. By selecting only those lens models which are consistent with the observed spatially resolved kinematics, we can substantially break the σ_{LT}-r_{cut} degeneracy for G2.
Of the 77470 MCMC chains of the G1/G2 lens model, only 9092 are consistent with the kinematics measurements. These restricted posterior distributions are plotted in green in Fig. 7.7. As comparison, we also show in red the posteriors obtained considering only a single velocity dispersion measured within a central aperture with $R^{in} = 0''$. In this case, the 95% critical value for the χ^2 reduces to 3.841 ($P(\chi^2_{kin} > 3.841) = 0.05$, for one degree-of-freedom).
Similar corner plots are shown in Figs. 7.8 and 7.9 for the EC and SG systems. For the EC system, we select the `LensTool` MCMC chains (72140 out of 77470) which are consistent with a single aperture velocity dispersion measurement of Gal.9323.

For the SG system, we utilize the measured velocity dispersion profile of the member Gal. 4104 in 3 radial bins (see Section 7.2) and adopt a larger, 99.9%, compatibility level ($P(\chi^2_{kin} > 16.266) = 0.05$). Thus, we find only 633 chains out of a total 319020 from the lens models, which are consistent with the observed kinematics.
The values (median, 16^{th}, 84^{th} percentiles) of σ_{LT} and r_{cut} obtained with this combination of lens models with lens galaxy kinematics for the three GSSLSs are quoted in red in Table 7.1.

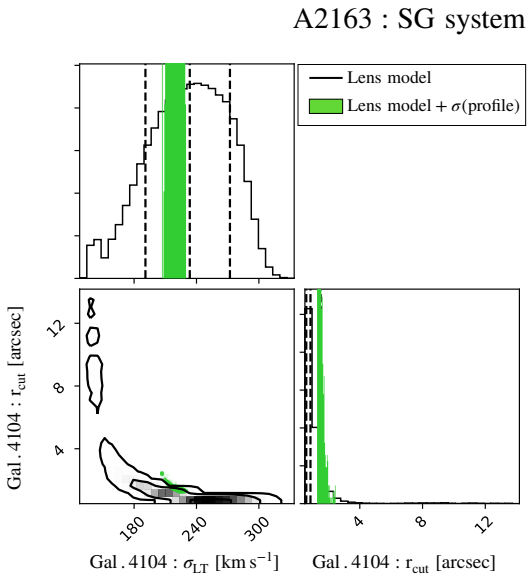


Figure 7.9: Same as Fig. 7.7 for the strong lensing system in A2163 (Gal. 4104).

In Fig. 7.10, we show the LoS velocity dispersion profiles for the two galaxies G2 and Gal. 4104. The grey areas are the `LensTool` predicted profiles computed considering all the MCMC chains of the lens model, while the green narrower regions are obtained only the chains consistent with the kinematics measurements at at three different radii. We note that there is some marginal tension between the SG lens model and the σ measurements. This would require additional analysis, testing a different mass density profile than a dPIE with vanishing core radius.

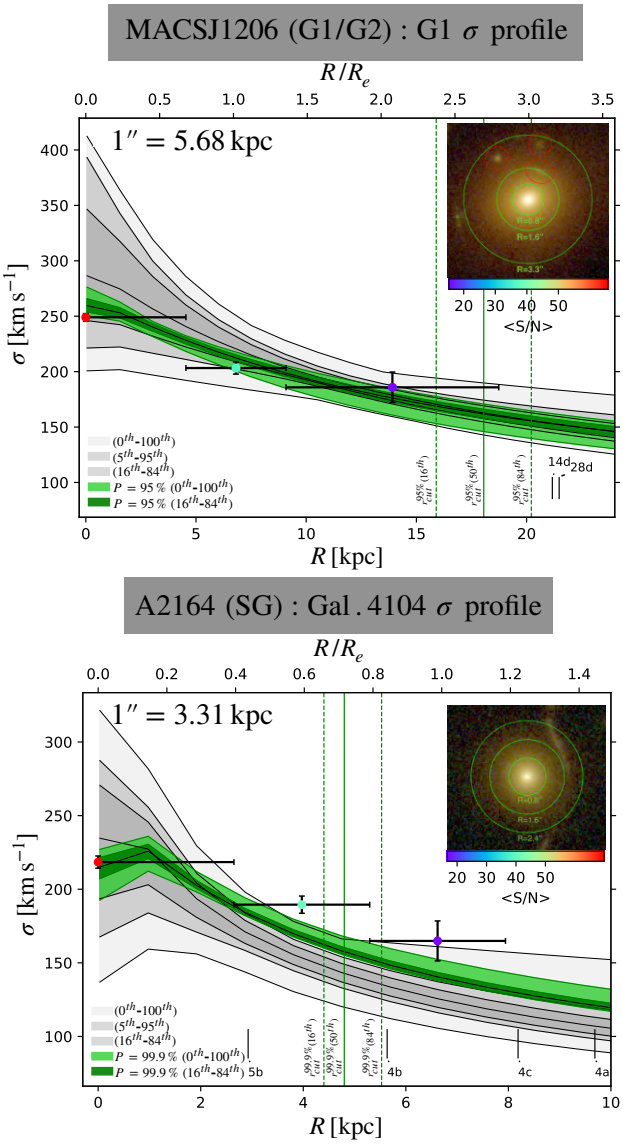


Figure 7.10: Line-of-sight velocity dispersion profile for the two GSSLS lenses G2 and Gal. 2104. The top axis shows the distance from the galaxy centers in units of the effective radius of each galaxy. The data points are the measured stellar velocity dispersions with MUSE within the three annuli, shown in green in the color cutouts. Points are color-coded according to the $\langle S/N \rangle$ of the spectra, while the horizontal error bars correspond to the radial bin sizes. Contaminating sources that were masked out in the spectral extractions are marked in red in the cutouts. The grey areas are computed from the lens model MCMC chains using the Eq. 2.61, while the green regions are obtained considering only the chains consistent with the kinematic measurements at 95% and 99.9% confidence level for G2 and Gal. 2104 respectively (see text). The vertical solid and dashed green lines are the median value, with the 16^{th} and 84^{th} percentiles, of the truncation radius, r_{cut}, obtained from the combination of the posterior distributions of lensing and kinematics. The vertical black segments indicate the projected distances of the multiple images from the galaxy centers (within the displayed range).

A more appropriate statistical approach to increase the sampling (the number of MCMC chains) in the region of the parameters space compatible with the kinematic measurements is the 'important sampling' method (Lewis & Bridle 2002). We plan to apply this technique in our future analyses.

7.4.3 Mass profiles and dark matter fractions

In Fig. 7.12, we plot the projected cumulative total mass profiles of all galaxies involved in the GSSLS analyzed (G1, G2, Gal. 9323, Gal. 4104) as a function of the projected distance from their centers. The green curves are cumulative mass profiles obtained from the `LensTool` MCMC chains which are compatible with kinematic measurements. While the single aperture velocity dispersion for Gal. 9323 only slightly reduces the uncertainty on the galaxy mass, the σ-profiles for G2 and Gal. 4104 reduce the 1σ uncertainty on the galaxy masses (within their effective radii) by factors between [1/9 - 1/4] for G2 and [1/7 - 1/3] for Gal. 4104.
Note that the values of the underlying cluster halo mass (grey curves) at the positions of the multiple image positions are comparable, or even larger, than the values of the masses of the GSSLS lenses. This shows how critical it is to model simultaneously the cluster mass distribution with a high-precision lens model (with a large number of independent constraints) together with the galaxy scale system (see Section 7.3).

As a final result, we show an estimate of the DM fraction of G1, G2 and Gal. 9323 galaxies. Since the total mass of a galaxy, M_{tot}, can be considered as the sum of its DM and stellar mass contents, we compute galaxy DM fractions as:

$$f_{DM} = \frac{M_{DM}}{M_{tot}} = 1 - \frac{M_L}{M_{tot}}. \tag{7.2}$$

The stellar masses (M_L) of G1, G2 and Gal. 9323 are measured by fitting at the same time the SED obtained from the 12 optical+NIR HST broadband images provided from CLASH and the MUSE spectrum, using the method developed by Gobat et al. (2008). We assume complex stellar populations with delayed exponential star formation histories, a Salpeter (Salpeter 1955) stellar initial mass function (see Section 5.1), and the Calzetti et al. (2000) law to account for the presence of dust. We adopt the updated 2016 version of the Bruzual & Charlot (2003) stellar population models, which uses the MILES stellar library (Sánchez-Blázquez et al. 2006).
The measured stellar masses, M_L, are in the fourth column of Table 7.3.

The total masses (M_{tot}) of the G1, G2 and Gal. 9323 galaxies are computed from the lens models describe in Section 7.3 using the Eq. 2.57 (see the third column of Table 7.3).
For the G2 galaxy we obtain the lowest DM fraction, $f_{DM}^{G2} = 0.66^{+0.09}_{-0.28}$, while for G1 and Gal. 9323 we obtain $f_{DM}^{G1} = 0.86^{+0.05}_{-0.18}$ and $f_{DM}^{\text{Gal. 9323}} = 0.93^{+0.05}_{-0.24}$.

In Fig. 7.11, we show an updated version of the Fig. 8 in Grillo et al. (2014). The plot shows the variation of the stellar mass of galaxies (M_L) as a function of their central velocity dispersions (σ_0). The red dots are our results, based on our lensing+kinematics analysis, while previous finding (for G1 and G2) by Grillo et al. (2014) based on lensing alone are in black. The plot also includes (in grey) the values obtained from a large sample of galaxy lenses selected from the Sloan Lens ACS Survey (SLACS; Bolton et al. 2006; Treu et al. 2006; Auger et al. 2010).
While for G1 we obtain σ_0 and M_L values in agreement with Grillo et al. (2014) results (within the errors), we derive a slightly higher central velocity dispersion for galaxy G2. However, all three G1, G2, and Gal. 9323 galaxies lie on the extrapolation of the best-fit scaling relation based on the SLACS galaxies only.

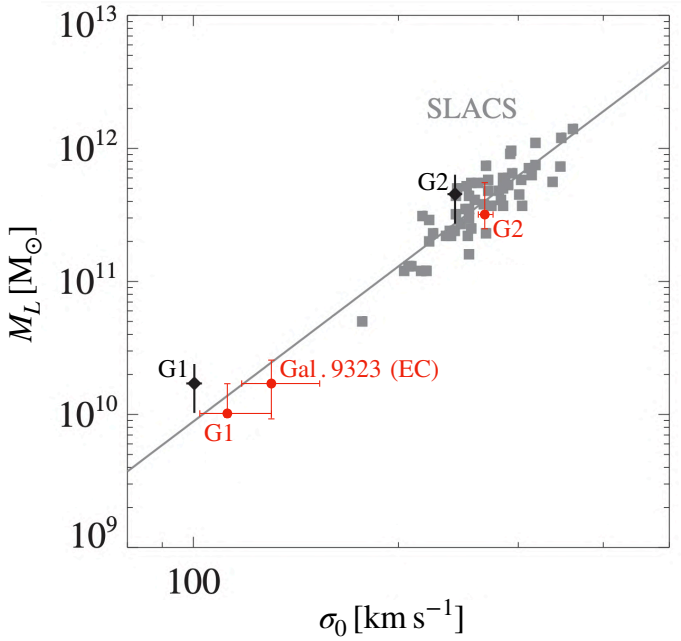


Figure 7.11: Luminous mass M_L as a function of the central velocity dispersion σ_0. The red dots are our results for the galaxies G1, G2 and Gal. 9323, while the black diamonds are previous findings by Grillo et al. (2014). The 1σ error bars are plotted. The grey squares are the SLACS galaxies. The solid line shows the best-fit relation based on the SLACS sample only. (Original figure from Grillo et al. 2014)

ID	m_{F160W}	M_{tot} [$\times 10^{10} M_\odot$]	M_L [$\times 10^{10} M_\odot$]	f_{DM}
G1	20.6	$7.00^{+4.21}_{-1.79}$	$1.00^{+0.67}_{-0.03}$	$0.86^{+0.05}_{-0.18}$
G2	17.7	$94.73^{+8.65}_{-7.69}$	$31.90^{+22.3}_{-7.0}$	$0.66^{+0.09}_{-0.28}$
Gal. 9323	20.7	$25.56^{+24.35}_{-17.45}$	$1.70^{+0.81}_{-0.79}$	$0.93^{+0.05}_{-0.24}$

Table 7.3: F160W magnitudes, total masses, luminous masses and dark matter fractions for the three lens galaxies G1, G2 and Gal. 9323.

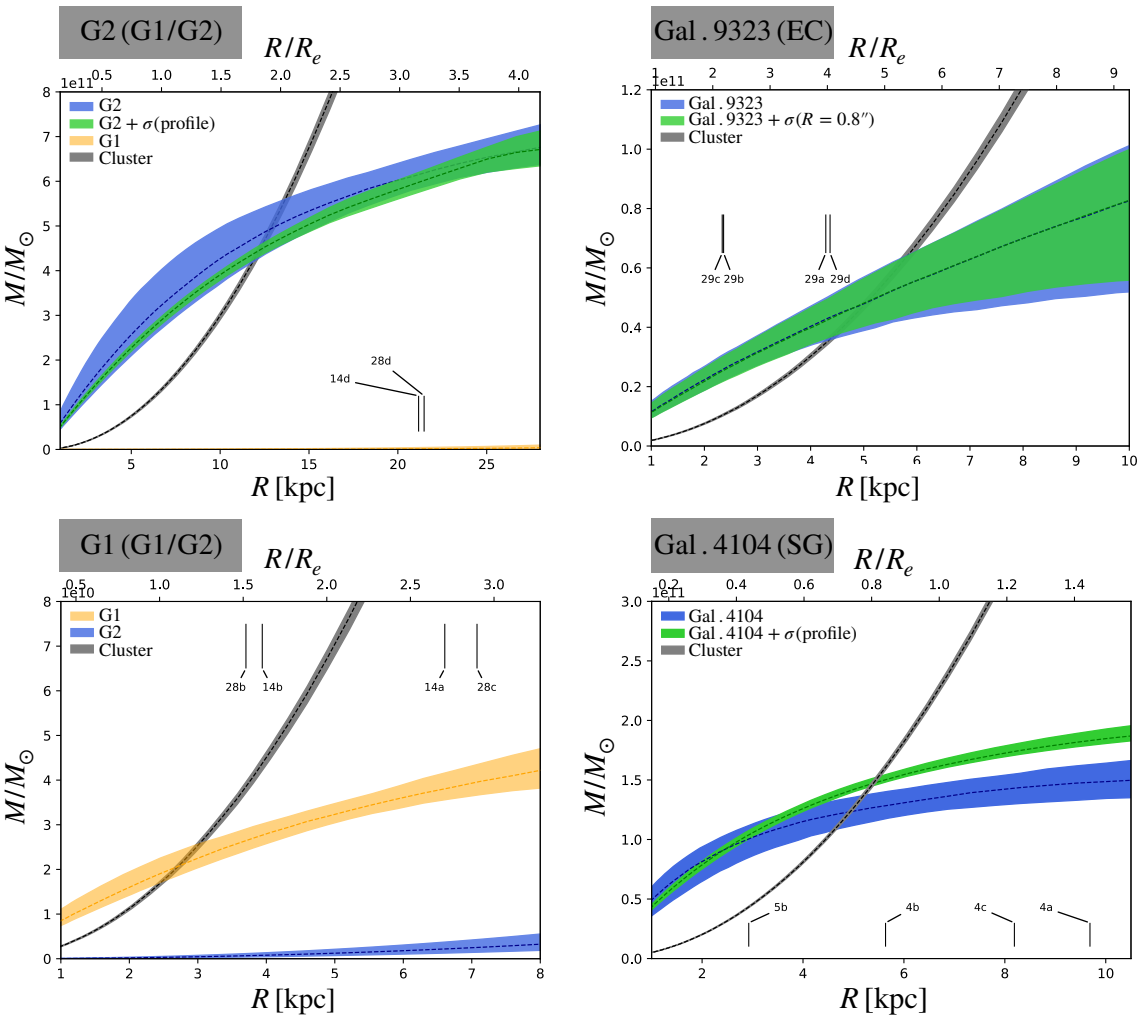


Figure 7.12: Projected cumulative mass profiles for the G1, G2, Gal. 9323 and Gal. 4104 galaxies, as a function of the projected distance from their centers. The colored regions encompass the 16^{th} and the 84^{th} percentiles; colored dashed lines are the median values. Different colors correspond to different cluster mass components. The green curves are obtained considering only the `LensTool` MCMC chains consistent with the kinematic measurements. The multiple image projected distances from the galaxy centers are marked with vertical black segments.

7.5 Conclusions

In this chapter, we have shown a preliminary strong lensing analysis of three galaxy cluster galaxies in the central regions of MACSJ1206 and A2163, which contribute to produce several multiple images within a few arcseconds (GSSLSs: galaxy scale strong lensing systems). Thanks to the underlying cluster lensing potential, the probability to have such an event is significantly higher than in the field for a given galaxy mass. A a result, one can explore the mass distribution of early-type galaxies in clusters to significantly lower masses than in the field (more than one order of magnitude). Our aim is to find a technique to robustly constrain the mass distribution of cluster galaxies, in particular the size of their associated sub-halos, over a large mass range. A comparison of the halo sizes of early-type galaxies, for a given mass, in cluster and field environments can shed light on the stripping mechanisms and other dynamical processes included in structure formation simulations.

Extending previous studies, we model the mass distribution of these early-type galaxies, by modeling at the same time the entire cluster mass distribution. For the latter, we use the high-precision lens model of MACSJ1206 described in Chapter 6 and the refined lens model for A2163 by Rescigno et al. (2019, submitted).

The modeling of the GSSLSs, in combination with a high-precision cluster lens model, is important to avoid possible biases in the mass distribution of the GSSLS lenses, specifically the mass and size of their associated sub-halos.

We then make use of the kinematic measurements with VLT/MUSE on the lens galaxies to improve the mass modeling of the GSSLSs. We particularly take advantage of the possibility to measure the stellar velocity dispersion profiles for two bright lens galaxies (MACSJ1206/G2 and A2163/Gal. 4104). By selecting only those lens models (i.e. the corresponding MCMC chains) which are consistent with the velocity dispersion measurements at different radii, we show how the σ_0-r_{cut} degeneracy can be significantly reduced. In particular, one obtains a robust estimate of the truncation radius of these member galaxies, which straddle more than a factor of 10 in mass.

In Fig. 7.13, we compare the statistical r_{cut}-σ_0 relation derived in Chapter 6 (Eq. 6.8) with the specific σ_0, r_{cut} values obtained in this chapter for the GSSLSs. Other estimates of σ_0, r_{cut} of cluster galaxies from the literature, using different methods and data sets, are included. While the G2 galaxy perfectly lies on the statistical relation, the other three G1, Gal. 9323 and Gal. 4104 cluster members scatter around the mean r_{cut}-σ_0 curve. This scatter could be explained considering

different stripping histories for galaxy halos (Warnick et al. 2008; Limousin et al. 2009; Natarajan et al. 2002), or simply admitting that the uncertainty of our empirical r_{cut}-σ_0 is not realistic for the poorly sampled low-mass galaxies. Clearly, a significantly larger sample of GSSLS in clusters, with high quality lensing and kinematic data, are needed to further investigate the mass-size relation of cluster sub-halos, and its scatter. This will be pursued in the near future with the cluster sample already in hand (Caminha et al. 2019).

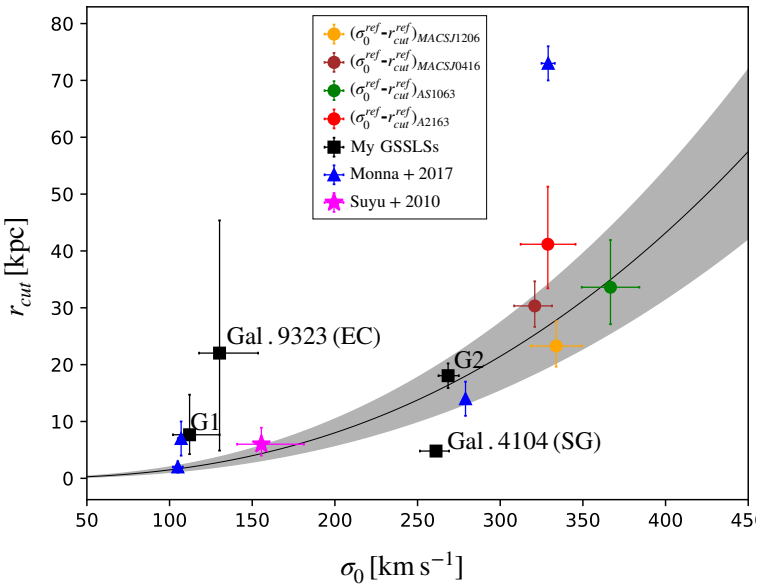


Figure 7.13: Comparison between the σ_0-r_{cut} scaling relation in Eq. 6.8 (grey area) and results from lensing+kinematics modeling of the GGSLS developed in this chapter (in black). The colored data points are the optimized reference values of the cluster member scaling relations of the lens models developed in Chapter 6 (for MACSJ1206, MACSJ0416 and AS1063) and in Chapter 7 (for A2163). As a comparison, we plot previous results from the analyses of galaxy scale systems by Suyu & Halkola (2010) (magenta star) and Monna et al. (2017) (blue triangles). In particular, Suyu & Halkola (2010) studied the mass distribution of a satellite galaxy of the group-scale lens SL2S J085440121 modelling the surface brightness distribution of extended lensed images around the galaxy. Instead, Monna et al. (2017) determined the mass profile of five cluster members of Abell 611 (one, with $\sigma_0 \sim 30$ km s^{-1}, outside the plot) through the surface brightness reconstruction of a giant arc.

Chapter 8

Comparing lensing results with simulations

In Chapter 6, we developed accurate lens models with kinematic priors for the three galaxy clusters MACSJ1206, MACSJ0416 and AS1063. The inclusion of the extra independent kinematic information has two main effects on cluster lens models (Section 6.5): firstly, it robustly constrains the sub-halo component of galaxy clusters, reducing inner degeneracies with other mass terms (Eq. 3.27); secondly, it breaks the σ_0^{ref}-r_{cut}^{ref} degeneracy between normalization parameters of cluster member scaling relations in Eqs. 3.28 and 3.29 (see Section 3.7.1).

In this chapter, I present my contribution to the paper by Meneghetti et al., currently submitted to *Science*. This paper reveals a one-order-of-magnitude mismatch between the probability to observe galaxy-scale strong lensing systems (GSSLS) in real clusters, and the same probability predicted by state-of-the-art N-body and hydrodynamical simulations.
Since our lensing-plus-kinematics models, unlikely most of models in the literature, ensure an accurate description of the sub-halo component of clusters, they are chosen as the reference observed data sets by Meneghetti et al., submitted. Indeed, a robust determination of sub-halo mass distribution, sometimes generally referred to as "sub-structures" is critical for a correct estimation of the GSSLS probability.

8.1 Hydrodinamical and N-body simulations of galaxy clusters: AGN, MOKA and Illustris

In this section, we briefly describe the three different sets of N-body and hydrodynamical cosmological simulations considered by Meneghetti et al., submitted.

8.1.1 AGN simulations

With the term "AGN", we refer to a subset of 25 highly resolved simulated galaxy clusters, extracted from the DIANOGA[1] simulations (Planelles et al. 2014), which include several mechanisms of energy feedback, from Supernova explosions to Active Galactic Nuclei (AGN). A flat ΛCDM Universe with $\Omega_{m,0} = 0.24$, $\Omega_{b,0} = 0.04$ and $H_0 = 72\,\mathrm{km/s/Mpc}$ is assumed. Simulated galaxy clusters were extracted from a lower resolution initial simulated box with a comoving size of 1 Gpc/h and containing 1024^3 collision-less particles. To identify the clusters in the simulated box a *Friends-of-Friends* algorithm was applied to the snapshot at $z = 0$. The identified clusters were re-simulated at higher resolution, starting from initial conditions derived by the *Zoomed Initial Conditions* code (Tormen et al. 1997; Bonafede et al. 2011), using the TreePM-SPH `GADGET-3` software by Springel (2005).
`GADGET-3` assumes collisionless DM particles with a mass of $8.47 \times 10^8\, h^{-1}\mathrm{M}_\odot$, while gas particles have a mass of $1.53 \times 10^8\, h^{-1}\mathrm{M}_\odot$. The code includes the description of the radiative gas heating/cooling caused by the cosmic microwave background and by a UV/X-ray time-dependent uniform ionizing background (Haardt & Madau 2001; Wiersma et al. 2009). It also adopts the Springel & Hernquist (2003) and Steinborn et al. (2015) models for the star formation and the Super-Massive-Black-Hole evolution, respectively.
Meneghetti et al. (submitted) consider six simulation snapshots in the redshift range [0.25 - 55]. All the 25 simulated clusters have $M_{200} \geq 5 \times 10^{14} h^{-1}\mathrm{M}_\odot$ and three independent lines-of-sight projections of each halos are considered as a different cluster to increase sample statistics.

[1] http://adlibitum.oats.inaf.it/cosmcomp/

8.1.2 MOKA simulations

The semi-analytic code MOKA (Giocoli et al. 2012a,b) was developed to analytically produce mock galaxy clusters by combining several mass components with realistic density profiles. In particular, MOKA adopts the same mass decomposition in Eq. 3.27. Smooth cluster-scale halos are usually parameterized as Navarro-Frenk-White (NFW) mass distributions, while cluster member subhalos consist in truncated SIS profiles (Hernquist 1990). The BCGs are modelled differently respect to cluster galaxies as Metcalf & Madau (2001) profiles. In real clusters, the baryonic component of the BCG causes an adiabatic contraction of the cluster central DM halo. MOKA takes into account this effect as described in Keeton & Madau (2001). Moreover, the sub-halo abundance relation by Giocoli et al. (2010b) is also implemented.
Since the total mass distribution of simulated clusters is simply the sum of several analytic mass profiles, MOKA clusters are characterized by an "infinite mass resolution".

8.1.3 Illustris simulations

The final set of simulated galaxy clusters is extracted from the *Illustris* simulations. The Illustris project[2] consists in a series of large-scale hydrodynamical simulations (eight in total) covering a volume of $[106.5\,\mathrm{Mpc}]^3$ each (see Fig. 8.1). Illustris-(1,2,3) take into account all the astrophysical processes known to be crucial for galaxy formation and evolution: e.g. primordial and metal-line cooling with self-shielding corrections; stellar evolution and stellar feedback; chemical enrichment; supermassive black hole growth; and feedback from active galactic nuclei. Conversely, Illustris-Dark-(1,2,3) are DM-only simulations, while Illustris-NR-(1,2) do not consider any energy feedback, nor cooling processes (Vogelsberger et al. 2014b). In Fig. 8.1, the main input parameters of different types of Illustris simulations are shown.
Illustris-1 has a DM mass resolution of $6.26 \times 10^6\,\mathrm{M}_\odot$ and an initial baryonic mass resolution of $1.26 \times 10^6\,\mathrm{M}_\odot$. At $z = 0$, the smallest hydrodynamical gas cell has a size of 48 pc and a mass of $\sim 10^4\,\mathrm{M}_\odot$. In the simulations the dynamic evolution of $\sim 12 \times 10^9$ DM particles is taken into account.
Illustris simulations adopt the (WMAP)-9 (Hinshaw et al. 2013) measured values for the cosmological parameters: $\Omega_m = 0.2726$, $\Omega_\Lambda = 0.7274$, $\Omega_b = 0.0456$ and $H_0 = 70.4\,\mathrm{ks/s/Mpc}$. These and the other initial conditions are implemented as in Vogelsberger et al. (2014a).

[2] http://www.illustris-project.org

All galaxy clusters in the Illustris-1 simulations have $M_{200} < 2 \times 10^{14}\, h^{-1}\mathrm{M}_{\odot}$.
For more details on the Illustris project see Vogelsberger et al. (2014b,a); Genel et al. (2014); Sijacki et al. (2015).

Name	Volume $(\mathrm{Mpc})^3$	DM particles/hydro cells/ MC tracers	$\epsilon_{\mathrm{baryon}}/\epsilon_{\mathrm{DM}}$ (pc)	$m_{\mathrm{baryon}}/m_{\mathrm{DM}}$ $(10^5\,\mathrm{M}_{\odot})$	$r_{\mathrm{cell}}^{\mathrm{min}}$ (pc)	$m_{\mathrm{cell}}^{\mathrm{min}}$ $(10^5\,\mathrm{M}_{\odot})$	Description
Illustris-1	106.5^3	$3 \times 1820^3 \cong 18.1 \times 10^9$	710/1420	12.6/62.6	48	0.15	Full physics
Illustris-2	106.5^3	$3 \times 910^3 \cong 2.3 \times 10^9$	1420/2840	100.7/501.0	98	1.3	Full physics
Illustris-3	106.5^3	$3 \times 455^3 \cong 0.3 \times 10^9$	2840/5680	805.2/4008.2	273	15.3	Full physics
Illustris-Dark-1	106.5^3	1×1820^3	710/1420	−/75.2	–	–	DM only
Illustris-Dark-2	106.5^3	1×910^3	1420/2840	−/601.7	–	–	DM only
Illustris-Dark-3	106.5^3	1×455^3	2840/5680	−/4813.3	–	–	DM only
Illustris-NR-2	106.5^3	$2 \times 910^3 \cong 1.5 \times 10^9$	1420/2840	100.7/501.0	893.8	6.6	No cooling/SF/feedback
Illustris-NR-3	106.5^3	$2 \times 455^3 \cong 0.2 \times 10^9$	2840/5680	805.2/4008.2	2322.8	39.4	No cooling/SF/feedback

Figure 8.1: Details of the Illustris simulation suite. Illustris-(1,2,3) are hydrodynamical simulations including the model by Vogelsberger et al. (2014b) for galaxy formation physics. Illustris-Dark-(1,2,3) are DM-only versions of the original Illustris simulations. They have the same initial conditions but do not include baryonic matter. Illustris-NR-(2,3) include baryons, but do not account for any feedback or cooling processes (non-radiative). Illustris-1 follows the evolution of 12 057 136 000 DM particles and hydrodynamical cells in total, with the smallest fiducial cell size (r_{cell}^{min}) below 50 pc and the least massive cells having masses of a few times $10^4\,\mathrm{M}_{\odot}$ (m_{cell}^{min}). In addition, it follows the evolution of 1820^3 Monte Carlo tracer particles (MC tracers). (Figure from Vogelsberger et al. 2014b).

8.2 GSSLS probability: Lens models vs simulations

The probability to observe a Galaxy Scale Strong Lensing (GSSL) event in a cluster is computed, in Meneghetti et al. (submitted), through the following equation:

$$P_{GSSLS}(z_S) = \frac{\sum_i \sigma_{cs,i}(z_S)}{A_S(z_S)} = \frac{\sigma_{GSSLS}(z_S)}{A_S(z_S)}, \tag{8.1}$$

where $A_S(z_s)$ is the observed area of the cluster (e.g. the area covered by the HST or MUSE pointings) mapped, by the lens equation, on the source plane at redshift z_S, and $\sigma_{GSSLS}(z_S) = \sum_i \sigma_{cs,i}(z_S)$ is the GSSLS total cross section computed for the redshift z_S. The latter quantity is defined as the area enclosed by all the galaxy-scale secondary tangential caustics produced by cluster sub-halos and not connected to the main cluster caustic (see Figure 8.2).
The probability $P_{GSSLS}(z_S)$ can be easily computed from the deflection angle map

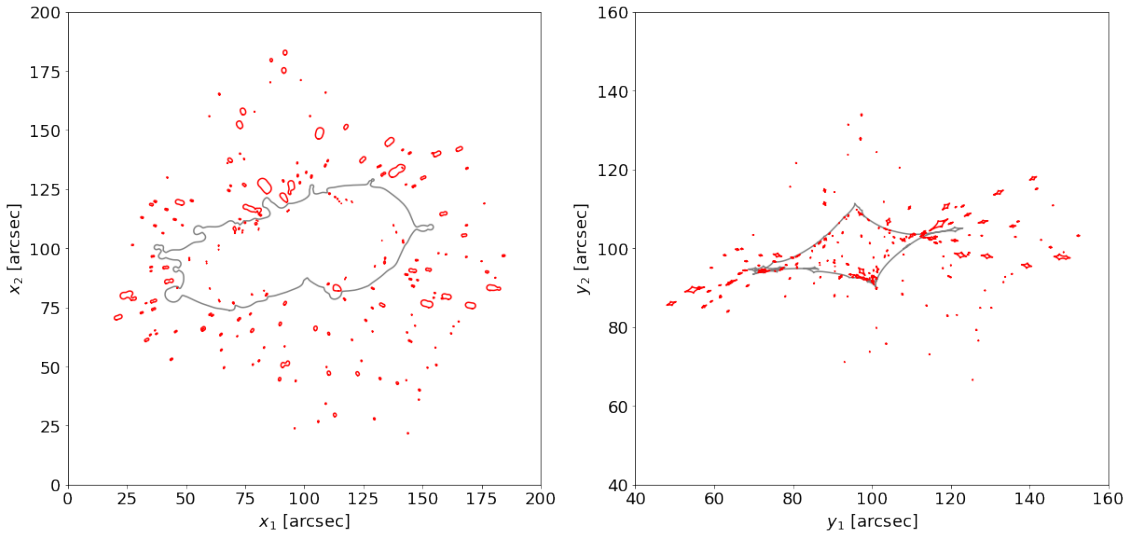


Figure 8.2: Identification of primary and secondary tangential critical lines of the galaxy cluster MACSJ1206 (left panel). The primary critical lines are shown in gray. The secondary critical lines are shown in red. In the right panel, we display the corresponding caustics. The cross section for GSSLS is computed by summing up the areas enclosed by each secondary caustic. The results in these figure refer to a source redshift $z_S = 7$ (Figure from Meneghetti et al., submitted).

of cluster lens models, or from the projected mass distribution of simulated galaxy clusters.

From the Eq. 8.1 we can also derive an expression for the expected number of observed GSSL events in each cluster:

$$N_{GSSLS} = \int_{S_{lim}}^{\infty} \int_{z_L}^{\infty} n(S, z_S) \sigma_{GSSLS}(z_S) dz_S dS, \tag{8.2}$$

where S_{lim} is a lower limit on the observable surface brightness, and $n(S, z_S)$ is the number density of sources, at redshift z_S, with surface brightness between S and $S + dS$.

In the upper panel of Fig. 8.3, we show the variation of $P_{GSSLS}(z_S)$ as a function of the source-plane redshift z_S. The red curve is the GSSLS probability computed for galaxy clusters in AGN simulations, while the other colored curves refer to the following samples of real cluster lens models:

- **MACSJ1206+MACSJ0416+AS1063:** This sample contains the cluster lens models with kinematic priors developed in Chapter 6. Our models take

into account both the DM and the hot-gas contents of the three clusters MACSJ1206, MACSJ0416 and AS1063. Moreover, the inclusion of kinematic priors on reference parameters of cluster member scaling relations ensures an accurate description of galaxy masses down to $\sim 10^{10}\,M_{\odot}$. For these reasons Meneghetti et al. chose these three clusters as the reference sample for the comparison with cosmological simulations.

- **Frontier Fields:** These are `LensTool` lens models[3], developed by the CATS[4] and the Johnson-Sharon[5] teams, of the six HFF clusters: Abell 2744, MACSJ0416, MACSJ1149, AS1063, MACSJ0717, and Abell 370 (see Fig. 4.4).

- **CLASH "Gold" (Caminha+19):** This sample contains lens models, developed by Caminha et al. (2019), for the eight CLASH clusters: RXJ2129, MACSJ1931, MACSJ0329, MACSJ2129, MACSJ1115, MACSJ0429, RXJ1347 and MACSJ1311 (see Fig. 4.3).

Finally, the dash blue curve in the plot is the $P_{GSSLS}(z_S)$ value computed from our MACSJ1206 lens model with kinematic prior alone.
Despite the differences among cluster lens models, they predict similar values for the GSSLS probability. Instead, Fig. 8.3 shows an evident discrepancy, of about one order of magnitude, between $P_{GSSLS}(z_S)$ values computed from lens models and AGN simulations. This mismatch is even larger (about two order of magnitude) considering our robust MACSJ1206 model alone.

To test if the resulting P_{GSSLS} values predict the correct number of GSSLSs observed in a real galaxy cluster we use the Eq. 8.2. To determine N_{GSSLS}, we assume a number density of sources, $n(S, z_S)$, reproducing the distribution of galaxies observed by the Hubble-Ultra-Deep-Field (HUDF) survey (Beckwith et al. 2006). We consider the surface brightness of HUDF galaxies in the F850LP filter, while galaxy photometric redshifts were determined by Rafelski et al. (2015) using two different codes: the Bayesian Photometric Redshift (BPZ) algorithm (Benítez 2011) and the EAZY software (Brammer et al. 2008).

[3]https://archive.stsci.edu/prepds/frontier/lensmodels/

[4]Clusters As TelescopeS team. P.I.s: J.-P. Kneib & P. Natarajan.

[5]P.I.: Keren Sharon (University of Michigan)

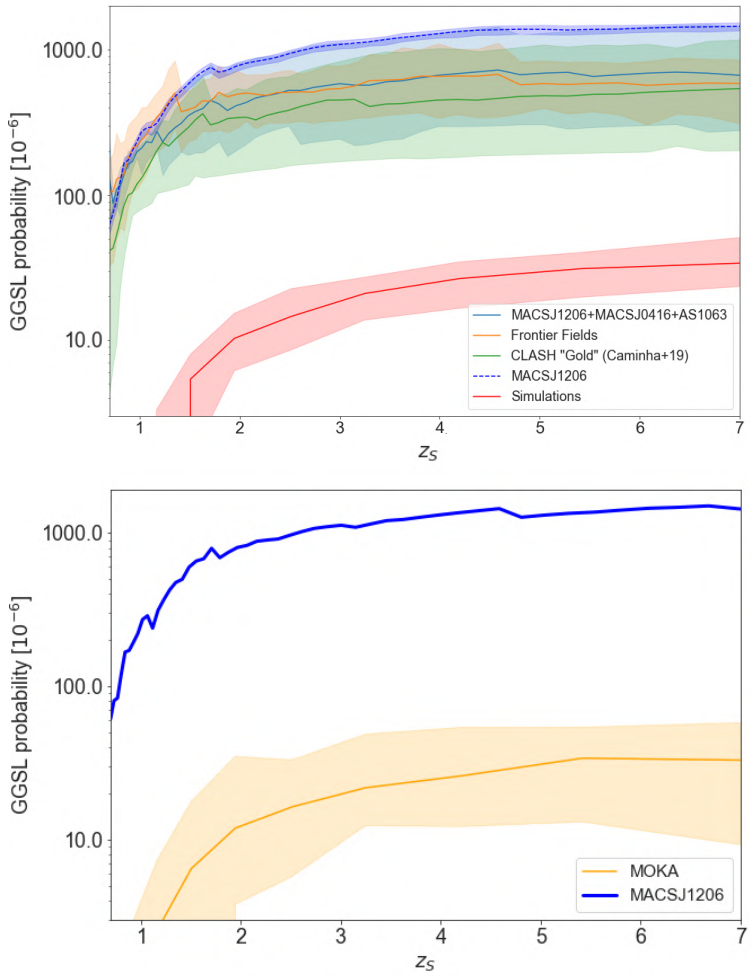


Figure 8.3: *Top*: Comparison between GSSLS probabilities of simulated and real galaxy clusters. The dashed dark blue line shows the GSSLS probability as a function of the source redshift for MACSJ1206. The solid light blue curve show the GSSLS probability for the reference sample composed of the three clusters MACSJ1206, MACSJ0416, and AS1063 (see Chapter 6). The orange solid line shows the median GSSLS probability of the six frontier fields, based on the available public `LensTool` models from the CATS and Johnson-Sharon teams. The green solid line refers to the CLASH Gold sample (see Section 8.1). The red line refers to the AGN simulations. The colored bands show the 99.9% confidence intervals around the median, computed using bootstrapping. *Bottom*: Probability for GSSLS in MACSJ1206 (blue curve) and in equivalent lens models generated with MOKA (orange curve). The red-colored band shows the 3σ uncertainty. (Figure from Meneghetti et al., submitted)

In the upper panel of Fig. 8.4, we show the number of GSSLSs predicted by our MACSJ1206 lens model with kinematic prior, as a function of the limiting surface brightness, S. For $S < 10^{-1}\,\mu\mathrm{Jy\,arcsec^{-2}}$ we expect about three GSSLSs (up to $z_s = 7$) in the field of view shown in the left panel of Fig. 8.2. This is exactly the

number of observed systems in MACSJ1206 (the three GSSLSs are shown in the lower panel of Fig. 8.4). Two of these GSSLSs, i.e. the G1/G2 and the EC systems, were studied in details in Chapter 7, while the third system is an almost complete Einstein ring originated by a source at $z_S = 1.42$ (see Caminha et al. 2017b).

8.3 Discussion

Meneghetti et al. (submitted) discuss several possible explanation to justify the mismatch between GSSLS probability values computed from cluster lens models or cosmological simulations. In this section, we summarize the main results obtained.

8.3.1 Selection effects in galaxy cluster samples

MACSJ1206 and all the other CLASH clusters were not lensing selected (see Meneghetti et al. 2014 and Section 4.2.1), thus the existence of selection biases on these clusters is unlikely. Conversely, selection effects could be present in the HFF sample since its clusters were chosen on the basis of their strong lensing strength (see Section 4.2.3).
To test the impact of selection effects on the P_{GSSLS} value, Meneghetti et al., submitted compare the GSSLS cross sections computed for real and simulated clusters with similar values of Einstein radii. Even comparing observed clusters with simulated analogous with similar lensing strength, the GSSLS cross section obtained from lens models is dramatically larger. This is due to an higher number of critical sub-halos with larger caustic lines in real clusters. If we account for a possible higher surface mass density of the observed clusters, we still found that the simulations underestimate the GSSLS cross section by more than one order of magnitude.
Thus, the mismatch in Fig. 8.3 is most likely not due to selection biases.

8.3.2 Resolution issues and baryonic effects in simulations

The AGN simulations have a mass resolution of about $10^8\,M_\odot$ (see Section 8.1.1). Since a simulated halo is well sampled only if it contains at least 100 particles, the AGN simulations accurately describe cluster sub-halos with total masses larger than $10^{10}\,M_\odot$.
Our lens models in Chapter 6 only consider cluster galaxies with $m_{F160W} < 24$. Since this magnitude corresponds to a total sub-halo mass of about $10^{10}\,M_\odot$ (Caminha

et al. 2017b), this suggests that the AGN simulations have enough resolution to resolve the smallest substructures considered in lens models.

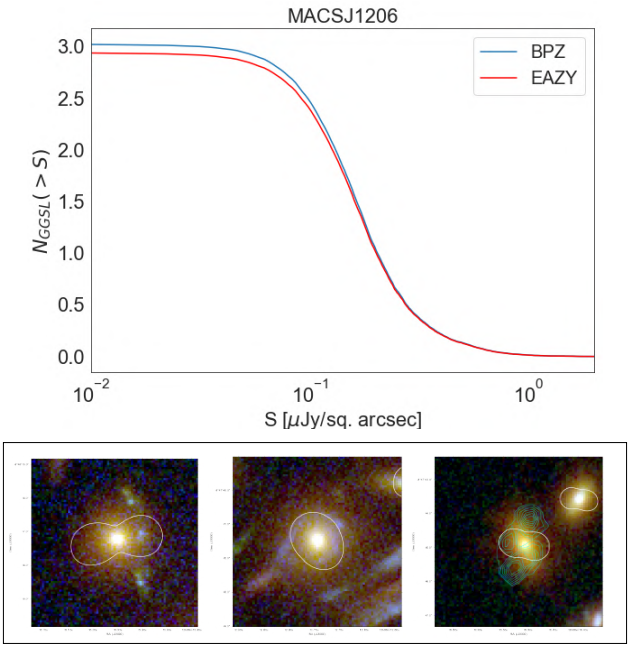


Figure 8.4: *Top*: Expected number of GSSL events in the field of MACSJ1206. The number is calculated using Eq. 8.2, assuming the source number density $n(S, z_S)$ measured in the HUDF, based on the photometric redshift catalogs of Rafelski et al. (2015). The blue and the red lines show the results based on the photometric redshift estimates obtained with two different codes, namely BPZ (Benítez 2011) and EAZY (Brammer et al. 2008). The estimated number of events is shown as a function of the minimal source surface brightness. *Bottom*: Three GSSL events detected in the core of MACSJ1206. From the left to the right: the GSSLS system studied in Chapter 7 and by Grillo et al. (2014) ($z_S = 3.75$), the Einstein ring ($z_S = 1.42$), and the Einstein cross studied in Chapter 7 ($z_S = 4.99$). In all panels, the critical lines for the corresponding source redshifts are displayed in white color. The cyan contours in the right panel, showing the source Lyα emission, are extracted from a stack of the MUSE data-cube in the wavelength range [7288-7293]Å. (Figure from Meneghetti et al., submitted)

As an additional verification that the mismatch in the P_{GSSLS} value is likely not due to mass resolution issues, we plot in the lower panel of Fig. 8.3 the MOKA predicted GSSLS probability. Although the MOKA simulations are not limited in mass resolution (owing to an analytical description of cluster mass distributions; see Section 8.1.2), we still observe a mismatch similar to that obtained using the AGN simulations.

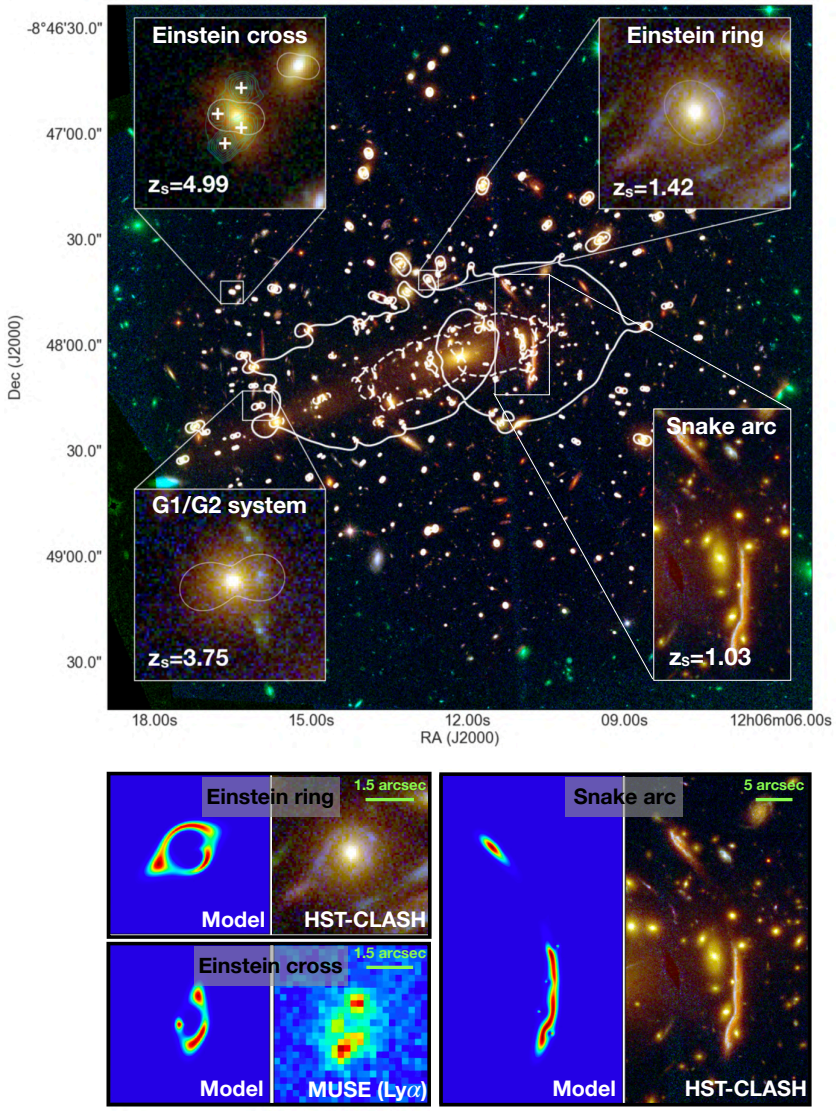


Figure 8.5: *Top*: Color-composite image of the central region of the galaxy cluster MACSJ1206 from 5 HST bands. The dashed and the solid lines show the lens critical lines for source redshifts 1 and 3, respectively. As the source redshift increases, the area enclosed by the critical lines grows in size. Three detected GSSL events involving sources at redshifts 3.75, 1.42, and 4.99 are shown in the insets, with the corresponding lens critical lines (in white). The bottom-right inset shows a previously known giant "snake" arc at $z = 1.03$. *Bottom*: Comparison between the lensed images of the systems Einstein ring, Einstein cross, and snake arc and the corresponding model-predicted images. The latter are obtained by lensing a simple gaussian source with the lens model of MACSJ1206 obtained by Caminha et al. (2017b), with a mass distribution very similar to our MACSJ1206 model with kinematic priors (see Section 6.5). This comparison shows the ability of our lens models to accurately reproduce different strong lensing features, determined by galaxy-to-cluster scale masses. (Figure adapted from Meneghetti et al., submitted)

The impact of the baryonic physics on the lensing signal was studied in details by Puchwein et al. (2005); Mead et al. (2010); Killedar et al. (2012). These and other studies demonstrated that the inclusion of stellar evolution processes (such as gas cooling, star formation and the stellar feedback) produces more concentrated and steepened mass profiles for the cluster galaxies with a consequent increase of the GSSLS cross section up to one order of magnitude. Conversely, the inclusion of the AGN feedback, necessary to solve problems such as over-cooling[6] (Sijacki et al. 2007; Fabjan et al. 2010; van de Voort et al. 2011; Rasia et al. 2015), reduces the GSSLS cross section because it smooths the cluster member profiles making the cluster subhalos puffier. Since a description of cluster subhalos without the AGN feedback is unrealistic, we finally conclude that baryonic processes, at least as implemeted to date, are not sufficient to justify the mismatch in the P_{GSSLS} value.

8.3.3 Lens models accuracy and LoS structures.

The ability of `LensTool` to determine the correct cluster sub-halo mass distributions was successfully tested, in the framework of the *Frontier-Fields Lens Modeling Comparison Project* (Meneghetti et al. 2017), using two realistic mock clusters. An exhaustive analysis presented by Meneghetti et al. (submitted) (here not reported) suggests that the uncertainties in the mass reconstructions are not enough to justify the one-order-of-magnitude mismatch in the P_{GSSLS} value. Moreover, it is reasonable to expect that our lens models with kinematic priors are even more accurate, i.e. less prone to systematics (see Chapter 6), on small scales than the models in Meneghetti et al. (2017).
A further confirmation of the accuracy of the MACSJ1206 lens model can be obtained by studying the model-predicted shapes of three extended multiple images, i.e.: the Einstein cross system (named EC in Chapter 7); the Einstein ring in Fig. 8.4; and the cluster-scale "snake" arc studied by Eichner et al. (2013). Our MACSJ1206 best-fit lens+kinematics model is able to reproduce the geometrical configurations, the shapes and the sizes of these extended images (simply ray-tracing sources with Gaussian surface brightness profiles on to the lens-plane) despite they are not used as model constrains (see Fig. 8.5 and Meneghetti et al., submitted for more details).

The mass distributions along the LoS are not considered by our cluster lens models. Only in the MACSJ0416 model we parameterize a foreground galaxy, close to the positions of several multiple images, as an extra dPIE profile (see Table 6.3).

[6] overestimation of the gas mass and stellar mass in the cluster core and at the center of each sub-halo

In Meneghetti et al. (submitted), we study the impact of LoS structures on the GSSLS cross section using simulations. In particular, we draw random samples of DM halos along the whole light-cones surrounded by cluster fields of view. For the halos we assume the mass function developed by Sheth & Tormen (2002). Finally, the code `GLAMER` (Metcalf & Petkova 2014; Petkova et al. 2014) is used to produce multi-plane ray-tracing simulations. These simulations demonstrate that the secondary critical lines produced by cluster sub-halos are only marginally affected by the presence of LoS structures. Thus, we conclude that the impact of LoS structures on the GSSLS cross section is negligible.

8.3.4 Sub-halos mass density profiles

The compactness of cluster member mass density distributions has a large impact on the value of the GSSLS cross section. In this subsection, we discuss the main differences between observed and simulated sub-halo profiles and test whether this can justify the mismatch in the P_{GSSLS} value.

As a first step, we compare the sub-halo maximum circular velocities, V_{circ} (see Eq. 2.27).
To measure the maximum circular velocity of simulated sub-halos we exploit the public code `SUBFIND` (Springel et al. 2001). This code decomposes simulated clusters into disjoint sub-halos and provides a measurement of their total bound mass and maximum circular velocity.
Conversely, `LensTool` parameterize cluster members as dPIE profiles and their circular velocities are simply given by the Eq. 2.59. As one can easily verify, for the full range of r_{core} and r_{cut} values assumed in our models the dPIE maximum circular velocity is well approximated (within 3%) by $V_{circ} = \sqrt{2}\sigma_0$, i.e. the circular velocity of the SIS profile (see Eq. 2.50 and Fig. 2.7).

In Chapter 6, we found a relation between the total mass of a cluster galaxy and its central velocity dispersion given by the Eq. 6.9. Replacing in this equation the Eq. 2.50, we obtain the following expression linking the total mass of a sub-halo to its maximum circular velocity:

$$M_{sub} = 3.5^{(4.6)}_{(2.6)} \times 10^{11}\,\mathrm{M}_{\odot} \left(\frac{\mathrm{V_{circ}}/\sqrt{2}}{220\,(\mathrm{km\,s^{-1}})} \right)^{4.43^{(4.45)}_{(4.38)}}, \tag{8.3}$$

where the parameter range refers to the 16^{th} and 84^{th} percentiles.

In the upper panel of Fig. 8.6, we show the cumulative distributions of the

sub-halo maximum circular velocities computed from our cluster lens+kinematics models, and from the AGN simulations. Only sub-halos with a total mass $M_{sub} > 10^{10}\,M_{\odot}\,h^{-1}$ and within the 15% of the projected virial radius of the host cluster are considered. As showed in Chapter 6, this is the fraction of the virial radius covered by the HST field of view where we have a highly complete sample of member galaxies in MACSJ1206, MACSJ0416 and AS1063. The blue, red and green curves in Fig. 8.6 are obtained by counting sub-halos in each cluster and using the lens model best-fit scaling relations (with $V_{circ} = \sqrt{2}\sigma_0$). Instead, the grey lines are computed from the AGN simulations, considering three projections of each simulated halo and stacking all the snapshots between redshifts 0.25 and 0.55.
The figure shows that simulated sub-halos have in general smaller maximum circular velocities than sub-halos in lens models, for a given mass. This result corroborates the previous findings by Grillo et al. (2015) and Bonamigo et al. (2018). Despite such a mismatch of the circular velocity functions, the bottom panel of Fig. 8.6 shows that the cumulative sub-halo mass functions derived from the lens models is in good agreement with the AGN and Illustris-1 simulations (see also Natarajan et al. 2017). Similarly, a very good agreement is found by comparing the differential sub-halo mass functions of simulated and observed sub-halos, as shown in Fig. 8.7.
To shed more light on these results we plot, in the upper panel of Fig. 8.8, the variation of the sub-halo maximum circular velocity as a function of the sub-halo total mass. The black line in the plot corresponds to the average relation in Eq. 8.3, while the colored dots are the AGN sub-halos within the virial radius of their host clusters. The blue dashed curve is the best-fit of those simulated sub-halos within the 15% of the projected virial radius.
Over the whole mass range, sub-halos in lens models have larger circular velocities than sub-halos in the AGN simulations. This translates into steeper and much more compact mass profiles for the observed galaxies that may explain the mismatch in the GSSLS cross section. This result is also confirmed by the Illustris-1 simulations. The lower panel of Fig. 8.6 shows that the AGN and Illustris-1 simulations predict similar sub-halo distributions in the V_{circ}-M_{sub} plane. Moreover, the higher mass resolution of the Illustris-1 simulations permits to extend this analysis over a wider range in mass, from about $10^{7}\,M_{\odot}\,h^{-1}$ up to $10^{13}\,M_{\odot}\,h^{-1}$.

A last quantity that can strongly affect the P_{GSSLS} value is the sub-halo radial distribution inside galaxy clusters. In Fig. 8.9, we plot the cumulative projected radial distributions of sub-halo cluster-centric distances scaled by the virial radius of the host cluster. Only galaxies with $M_{sub} > 10^{10}\,M_{\odot}\,h^{-1}$ and within 15% of the projected virial radius are considered. Red and grey lines are the average cumulative distributions for sub-halos in our lens+kinematics models and in the AGN simulations respectively.

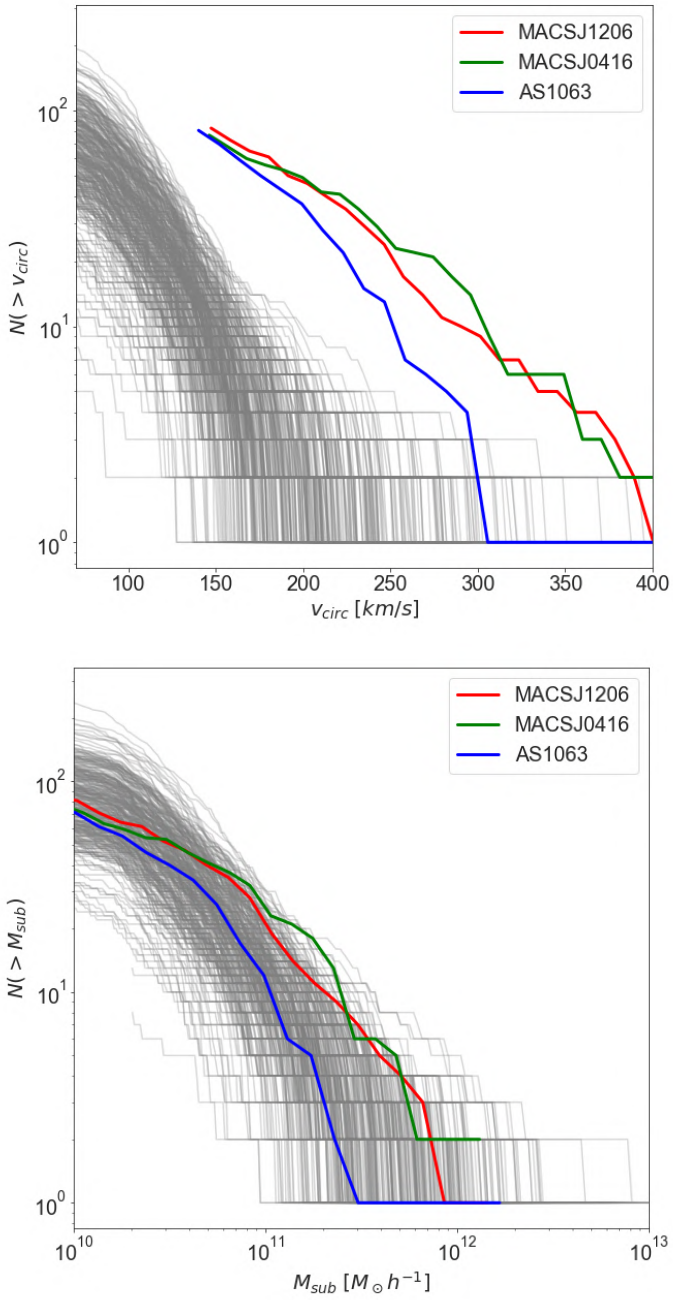


Figure 8.6: *Top*: Cumulative distributions of the sub-halo maximum circular velocities. *Bottom*: Cumulative distributions of the sub-halo masses. In both the top and the bottom panels the grey lines show the results for three projections of each cluster in the AGN dataset, stacking all snapshots corresponding to $0.25 \leq z_L \leq 0.55$. The red, blue, and green curves show the results from the lens models of MACSJ1206, MACSJ0416, and AS1063 developed in Chapter 6. (Figure from Meneghetti et al., submitted)

The plot shows that observed cluster galaxies are more concentrated towards cluster centers than sub-halos in the AGN simulations. Since the observed cluster members are also closer to the cluster main critical lines, they receive a boost of their strong lensing cross sections.

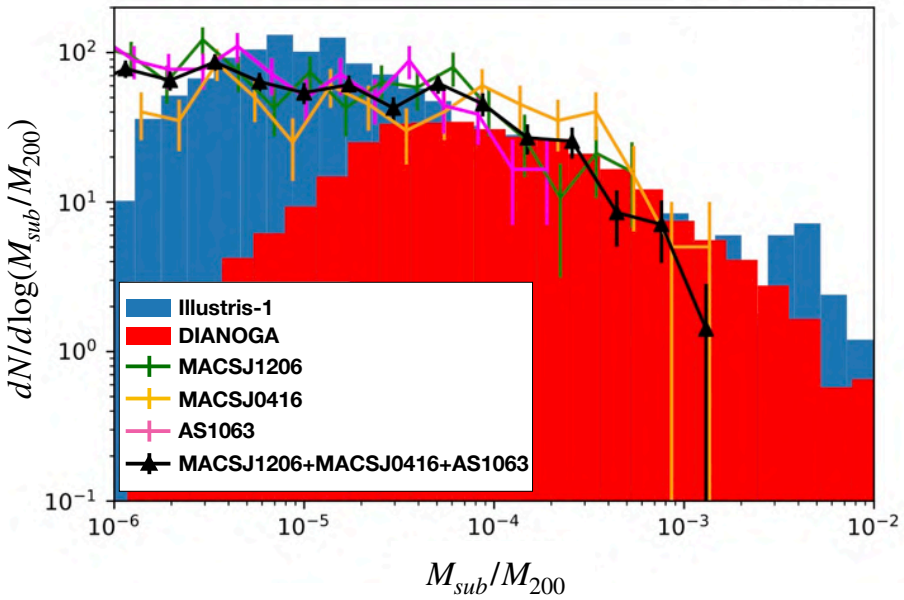


Figure 8.7: Sub-halo mass functions from different data sets of observed and simulated clusters. The sub-halo masses are normalized to the M_{200} mass of the host cluster. The results from the whole set of simulated clusters in the DIANOGA (see Planelles et al. 2014 and Section 8.1.1) and Illustris-1 simulations are shown in red and blue, respectively. Only clusters with $M_{200} > 2 \times 10^{13}\, h^{-1}\mathrm{M}_{\odot}$ and sub-halos within 16% of the projected virial radius of the host cluster are taken into account. Three projections of each simulated halo are treated as different clusters to increase statistics. The sub-halo mass functions obtained from the lens models with kinematic priors for the three clusters MACSJ1206, MACSJ0416 and AS1063 (see Chapter 6) are shown in green, yellow and magenta. The median distribution of the observed datasets is in black. The mass function from the DIANOGA simulations drops below $\sim 10^{-4}\, M_{sub}/M_{200}$, while the Illustris-1 mass function is complete down to $\sim 10^{-5}\, M_{sub}/M_{200}$ owing to the higher mass resolution. (Figure from Giocoli et al. (2019), in prep.)

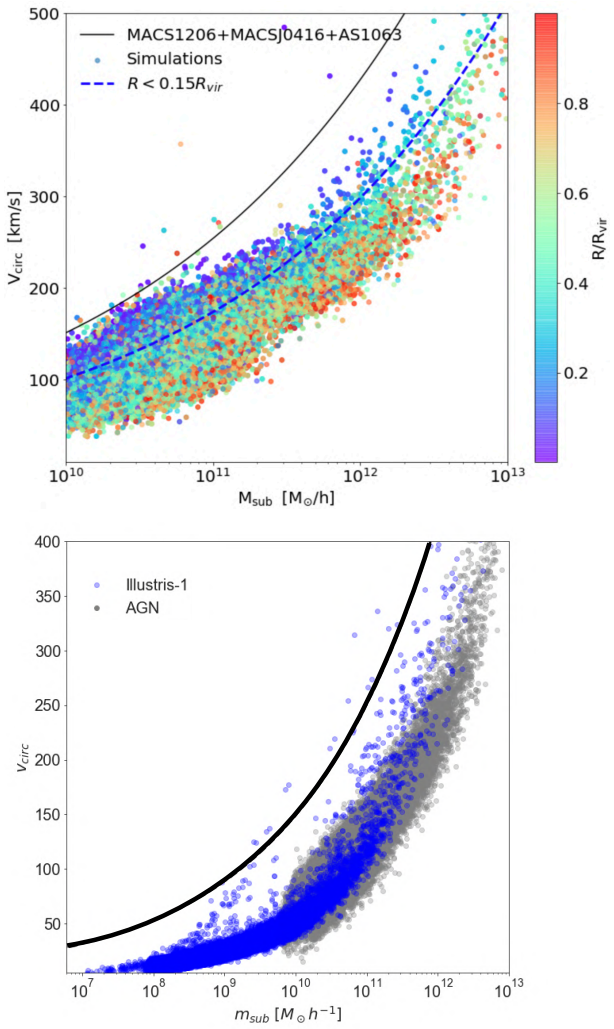


Figure 8.8: *Top*: Substructure circular velocity versus substructure mass. The circular velocity is a proxy of the substructure concentration. The solid line shows the average relation in Eq. 8.3 with the assumption that $V_{circ} = \sqrt{2}\sigma_0$. The colored circles indicate the results for sub-halos in the AGN dataset. The colors encode the projected distances of the sub-halos from the host halo center in units of the virial radius R_{vir}. The dashed curve shows the best-fit to the distribution of sub-halos within a cluster-centric distance of $0.15\,R_{vir}$. The observed relation is always above the simulated datapoints, indicating that observed substructures are more compact than the simulated ones. *Bottom*: Same as the upper panel but, for comparison, we also show in blue the results for the sub-halos extracted from the 11 most massive halos in the Illustris-1 simulation at $z_L = 0.4$ and $z_L = 0.47$. The solid line shows the average relation in Eq. 8.3 as before. The grey circles indicate the results for sub-halos in the AGN dataset. (Figure from Meneghetti et al., submitted)

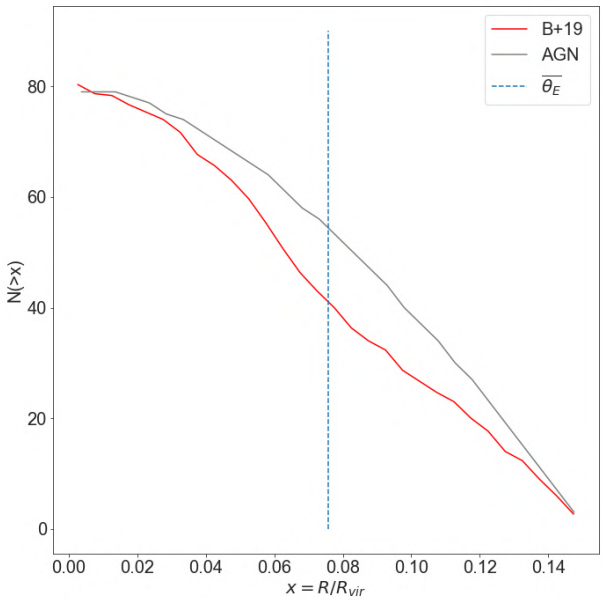


Figure 8.9: Cumulative projected distribution of sub-halo cluster-centric distances. We limit the analysis to sub-halos located within $0.15\,R_{vir}$ of their host and with mass $M_{sub} > 10^{10}\,\mathrm{M_\odot\,h^1}$. The grey and the red lines show the average cumulative distribution of sub-halos in the AGN and in the observed dataset (i.e. the lens models with kinematic priors in Chapter 6). (Figure from Meneghetti et al., submitted)

8.4 Conclusions

Fig. 8.3 shows a one-order-of-magnitude mismatch between the P_{GSSLS} values computed from lens models and from the state-of-the-art cosmological simulations performed in the ΛCDM framework.
From the discussion in the previous section emerges that this mismatch is likely due to differences in the internal structure and spatial distribution of the cluster sub-halos. Indeed, the observed sub-halos are more compact and more concentrated towards the cluster center than to those in simulations.

We still do not understand whether the origin of the lack of compact simulated substructures has to be ascribed to numerical and resolution effects in the simulations (e.g. a over-efficient tidal stripping, resolution limits, inadequate force softening), or to a different DM prescription respect to the CDM framework, or even to a missing piece of physics affecting the formation and the evolution of cluster galaxies. In any case, new high-resolution simulations are necessary to exclude numerical issues. Moreover, the development of accurate lens models for a larger number of GSSLS (such as the models in Chapter 7) will be fundamental for a detailed reconstruction

of cluster member mass profiles that may be compared with simulation results. For further details see the reference paper by Meneghetti et al., submitted.

Chapter 9

Accurate Cluster Lensing with Hierarchical Inference

In almost all the cluster lens models developed until now (including models in the previous chapters) cluster galaxies are modeled as belonging to zero-scatter scaling relations (in Eqs. 3.28, 3.29), typically with hyperparameters that are imposed externally from theoretical considerations (e.g. Limousin et al. 2007; Halkola et al. 2007; Eichner et al. 2013; Grillo et al. 2015; Monna et al. 2015; Caminha et al. 2019, 2017b,a, 2016; Jauzac et al. 2015; Kawamata et al. 2016; Lagattuta et al. 2017; Bonamigo et al. 2018) or inferred from kinematic measurements (as in Chapter 6). Any departures of individual galaxies from the prescribed scaling relations were determined heuristically, on a galaxy-by-galaxy basis, as we did in Chapter 7 to study the GSSLSs.

We then aim at cluster lensing models where the parameters of all the cluster-member galaxies should be let free to vary around some finite-scatter scaling relations, and deviate significantly from them if and only if the data require it. Moreover the scaling relation parameters have to be determined directly from the observed data of each given cluster. We achieve this, developing the *Hierarchical Bayesian* inference formalism show in this chapter, where each galaxy has its own associated parameters, and the parameters of all the galaxies are posited to be drawn from common relations with *hyperparameters*, including intrinsic scatter, to be determined through lens modeling and (if given) auxiliary stellar kinematic information. Since only one of the contributions to the inference is from lensing, and in order to ensure a fair comparison with state-of-the art technology, we build our inference as a modular wrapper, that we call `BayesLens`, that relies on common (external) lensing codes for the lensing likelihood and samples the posterior

on parameters and hyperparameters. This also ensures that further constraints, e.g. from time delays, flux-ratios and shapes of background sources, can be easily included in the inference.

The chapter is structured as follows. Our hierarchical inference is detailed in Section 9.1. Section 9.2 covers some technicalities inherent to our code implementation of the models. In Section 9.3, we perform functional tests and a comparison of our models to the current benchmark, on a simple toy model of a galaxy cluster. Results are discussed in Section 9.4, and we conclude in Section 9.5.

The results of this chapter are independent of the adopted cosmology. In fact, if needed, cosmological parameters may be sampled as additional hyperparameters in our inference scheme.

9.1 A Hierarchical Lensing Model

9.1.1 Measured kinematics

Our purpose-built `BayesLens` wrapper uses the available measured velocity dispersions of the cluster galaxies, $\sigma_m^{gal} \pm \delta\sigma_m^{gal}$, to infer the hyperparameters ($\hat{\sigma}^{ref}$, $\hat{\alpha}$ and $\hat{r}_{core}^{ref}$) of the σ - mag and r_{cut} - mag scaling relations (in Eqs. 3.28 and 3.29). β_{cut} is obtained from Eq. 6.1. A third hyperparameter, $\hat{\Delta\sigma}^{ref}$, quantifies the scatter of the measured galaxies around the σ - mag scaling relation. Gaussian priors, centered on the measured σ_m^{gal} and with standard deviations equal to the measured errors $\delta\sigma_m^{gal}$, are adopted for the measured kinematics of galaxies inside the lens model. For galaxies without a measured velocity dispersion, we assumed Gaussian priors centered on the inferred σ - mag scaling relation and with a standard deviation equal to $\hat{\Delta\sigma}^{ref}$. Unless otherwise stated, all model hyper-parameters are left free to vary, so as to fully explore the model posterior probability as described below.

9.1.2 Fitting it all together: the posterior probability distribution

In our models, we use the measured velocity dispersions of N_m^{gal} cluster galaxies, $\sigma_m^{gal} \pm \delta\sigma_m^{gal}$, together with the positions $\mathbf{x}_{im}$ of N^{im} multiple images, from N^{fam} different sources with positions $\mathbf{x}_{sou}$, as observational constraints to the lens model free parameters. Hereafter, these free parameters will be marked with a hat symbol.

In order to explore the lens models, we sample the total posterior probability function p_{tot}, expressed as the product of five factors:

$$\begin{aligned} p_{tot}\left(\hat{\sigma}^{ref}, \hat{\alpha}, \hat{\Delta\sigma}^{ref}, \hat{r}_{cut}^{ref}, \hat{\sigma}_m^{gal}, \hat{\sigma}^{gal}, \hat{\phi}^h \mid m^{gal}\sigma_m^{gal}, \delta\sigma_m^{gal}, \mathbf{x}_{im}\right) \\ \propto p_{sr}\left(\hat{\sigma}^{ref}, \hat{\alpha}, \hat{\Delta\sigma}^{ref}, \hat{r}_{cut}^{ref} \mid m^{gal}\sigma_m^{gal}, \delta\sigma_m^{gal}\right) \\ \times p_{mg}\left(\hat{\sigma}_m^{gal} \mid \sigma_m^{gal}, \delta\sigma_m^{gal}\right) \times p_g\left(\hat{\sigma}^{gal} \mid \hat{\sigma}^{ref}, \hat{\alpha}, \hat{\Delta\sigma}^{ref}\right) \\ \times p_h\left(\hat{\phi}^h\right) \times p_{im}\left(\mathbf{x}_{im} \mid \hat{\sigma}^{ref}, \hat{\alpha}, \hat{\Delta\sigma}^{ref}, \hat{r}_{cut}^{ref}, \hat{\sigma}_m^{gal}, \hat{\sigma}^{gal}, \hat{\phi}^h, \mathbf{x}_{sou}\right). \end{aligned} \quad (9.1)$$

Each of the five factors at the right-hand-side is discussed below. In the following, the quantities referring to cluster galaxies with measured velocity dispersions are marked with the subscript "m".

Scaling relation p_{sr}

This factor is responsible for the σ - mag and r_{cut} - mag scaling relation hyper-parameters, given the set of N_m^{gal} measured cluster galaxies. For the galaxy velocity dispersions we consider here a scaling relation of the same form of Eq. 3.28 parameterized by the reference measured velocity $\hat{\sigma}^{ref}$ and slope $\hat{\alpha}$ plus an intrinsic scatter $\hat{\Delta\sigma}^{ref}$ in measured velocity dispersions. This in turn can be expressed as:

$$\begin{aligned} p_{sr}\left(\hat{\sigma}^{ref}, \hat{\alpha}, \hat{\Delta\sigma}^{ref}, \hat{r}_{cut}^{ref} \mid m^{gal}, \sigma_m^{gal}, \delta\sigma_m^{gal}\right) \propto \\ p_{sr}\left(\sigma_m^{gal} \mid m^{gal}, \delta\sigma_m^{gal}, \hat{\sigma}^{ref}, \hat{\alpha}, \hat{\Delta\sigma}^{ref}\right) p_{sr}\left(\hat{\sigma}^{ref}, \hat{\alpha}, \hat{\Delta\sigma}^{ref}, \hat{r}_{cut}^{ref}\right), \end{aligned} \quad (9.2)$$

where the prior:

$$\ln\left\{p_{sr}\left(\hat{\sigma}^{ref}, \hat{\alpha}, \hat{\Delta\sigma}^{ref}\right)\right\} = \begin{cases} -\ln(\hat{\Delta\sigma}^{ref}), & \text{if } \sigma_{min}^{ref}<\hat{\sigma}^{ref}<\sigma_{max}^{ref} \\ & \text{and } \alpha_{min}<\hat{\alpha}<\alpha_{max} \\ & \text{and } \Delta\sigma_{min}^{ref}<\hat{\Delta\sigma}_m<\Delta\sigma_{max}^{ref} \\ & \text{and } r_{cut,min}^{ref}<\hat{r}_{cut}^{ref}<r_{cut,max}^{ref} \\ -\infty, & \text{otherwise} \end{cases} \quad (9.3)$$

limits the σ - mag plus scatter scaling relation parameters to lie within the boundaries $\sigma_{min(max)}^{ref}$, $\alpha_{min(max)}$ and $\Delta\sigma_{min(max)}^{ref}$. In principle, this term would also include a prior on the inferred reference truncation radius $\hat{r}_{cut}^{ref}$ appearing in Eq. 3.29.

The log-likelihood

$$\ln\left\{p_{sr}\left(\sigma_m^{gal} \mid m^{gal}, \delta\sigma_m^{gal}, \hat{\sigma}^{ref}, \hat{\alpha}, \hat{\Delta\sigma}^{ref}\right)\right\} = \\ = -\frac{1}{2}\sum_{i=1}^{N_m^{gal}}\left[\frac{\left(\sigma_{m,i}^{gal} - \hat{\sigma}_{m,i}^{sr}\right)^2}{\left(\delta\sigma_{m,i}^{gal}\right)^2 + \left(\hat{\Delta\sigma}^{ref}\right)^2} + \ln\left\{2\pi\left[\left(\delta\sigma_{m,i}^{gal}\right)^2 + \left(\hat{\Delta\sigma}^{ref}\right)^2\right]\right\}\right], \quad (9.4)$$

determines the values of scaling relation parameters that better reproduce the measured velocity dispersions. In Eq. 9.4 we define $\hat{\sigma}_{m,i}^{sr}$ as:

$$\hat{\sigma}_{m,i}^{sr} = \hat{\sigma}^{ref} 10^{0.4\,\hat{\alpha}\left(m^{ref} - m_i^{gal}\right)}, \quad (9.5)$$

where m^{ref} corresponds to the reference luminosity L_0 in Eq. 3.28.
The p_{sr} term is totally analogous to the Eq. 6.3 in Chapter 6.

Regarding the r_{cut} - mag scaling relation, in the current version of our models we optimize only the reference value $\hat{r}_{cut}^{ref}$ while the slope β_{cut} is determined using Eq. 6.1 from the inferred $\hat{\alpha}$ and assuming a fixed mass-to-light scaling for the cluster galaxies. No scatter around this relation is considered.

Prior on 'measured' galaxies p_{mg}

This term applies only on the N_m^{gal} galaxies with measured velocity dispersions. p_{mg} alone consists in Gaussian priors centering the cluster-member aperture-average velocity dispersions $\hat{\sigma}_m^{gal}$ on their kinematic values σ_m^{gal}. We choose the standard deviations of the Gaussian priors equal to the errors on the kinematic measurements $\delta\sigma_m^{gal}$. This term is given by:

$$\ln\left\{p_{mg}\left(\hat{\sigma}_m^{gal} \mid \sigma_m^{gal}, \delta\sigma_m^{gal}\right)\right\} = \\ = -\frac{1}{2}\sum_{i=0}^{N_m^{gal}}\left\{\frac{\left(\hat{\sigma}_{m,i}^{gal} - \sigma_{m,i}^{gal}\right)^2}{\left(\delta\sigma_{m,i}^{gal}\right)^2} + \ln\left[2\pi\left(\delta\sigma_{m,i}^{gal}\right)^2\right]\right\}. \quad (9.6)$$

In other words, this term attributes to the galaxies their measured velocity dispersions, unless a deviation from these values produces a significant improvement of the lensing model.

Prior on unmeasured galaxies p_g

This term is a collection of Gaussian priors on the velocity dispersion values for the $N^{gal} - N_m^{gal}$ galaxies without kinematics measurements. Its form is such that the final posterior prefers lens models in which the unmeasured galaxies lie on the σ - mag scaling relation inferred by p_{sr}, unless otherwise required by the lensing data:

$$\ln\left\{p_g\left(\hat{\sigma}^{gal} \mid \hat{\sigma}^{ref}, \hat{\alpha}, \hat{\Delta\sigma}^{ref}\right)\right\} = \\ = -\frac{1}{2}\sum_{i=1}^{N^{gal}-N_m^{gal}}\left\{\frac{\left(\sigma_i^{gal} - \hat{\sigma}_i^{sr}\right)^2}{\left(\hat{\Delta\sigma}^{ref}\right)^2} + \ln\left[2\pi\left(\hat{\Delta\sigma}^{ref}\right)^2\right]\right\}. \quad (9.7)$$

The $\hat{\sigma}^{sr}$ are computed through Eq. 9.5, but now considering galaxies without measured velocity dispersions. In our code we include also the following log-prior, summed to the previous one:

$$\ln\left\{p_g\left(\hat{\sigma}^{gal}\right)\right\} = \begin{cases} 0, & \text{if } \sigma_{min}^{gal} < \hat{\sigma}^{gal} < \sigma_{max}^{gal} \\ -\infty, & \text{otherwise} \end{cases}. \quad (9.8)$$

This additional flat prior is used to further reduce the dimensionality of the parameter space, avoiding that unconstrained galaxies (e.g. those that do not contribute substantially to the mass budget, or to the lensing likelihood) deviate too far from any model scaling relation. The prior boundaries are given by:

$$\sigma_{min(max)}^{gal} = \sigma_{min(max)}^{ref} 10^{0.4\,(m^{ref} - m_i^{gal})\,\alpha_{min(max)}}, \quad (9.9)$$

with appropriate values for $\sigma_{min(max)}^{ref}$ and $\alpha_{min(max)}$.

Prior on halo parameters p_h

This term consists in flat priors on the smooth cluster-scale halos parameters collectively indicated as ϕ^h:

$$\ln\left\{p_h\left(\hat{\phi}^h\right)\right\} = \begin{cases} 0, & \text{if } \phi_{min}^h < \hat{\phi}^h < \phi_{max}^h \\ -\infty, & \text{otherwise} \end{cases} \quad (9.10)$$

If (as follows) these halos are parameterized as PIEMD profiles (see Eq. 2.44), p_h is a prior on the sky coordinates x^h, on the ellipticity e^h, position angle θ^h, core radius r_{core}^h and central velocity dispersion σ_0^h.

Multiple-image likelihood p_{im}

This final term is a likelihood function, which quantifies the agreement between observed and predicted multiple-image positions.

Given N^{fam} sources with N_i^{im} multiple images associated to the same i^{th} source, we follow Jullo et al. (2007) and express the likelihood as:

$$p_{im}\left(\mathbf{x}_{im} \mid \hat{\sigma}^{ref}, \hat{\alpha}, \hat{\Delta\sigma}^{ref}, \hat{r}_{cut}^{ref}, \hat{\sigma}_m^{gal}, \hat{\sigma}^{gal}, \hat{\phi}^h, \mathbf{x}_{sou}\right) = \\ = \prod_{i=0}^{N^{fam}} \frac{\mathrm{e}^{-\chi_i^2/2}}{\prod_j \Delta x_{i,j}^2 \sqrt{2\pi}}, \quad (9.11)$$

where the χ_i^2 associated to the i^{th}-family is:

$$\chi_i^2 = \sum_{j=1}^{N_i^{im}} \frac{\left\| \mathbf{x}_{i,j}^{obs} - \mathbf{x}_{i,j}^{pred} \right\|^2}{\Delta x_{i,j}^2}, \quad (9.12)$$

with $\mathbf{x}_{i,j}^{obs}$ the observed positions of the multiple images on the lens plane, $\Delta x_{i,j}$ are the uncertainties on these positions and $\mathbf{x}_{i,j}^{pred}$ are the model predicted positions, given the inferred set of model parameters: $\hat{\sigma}^{ref}, \hat{\alpha}, \hat{\Delta\sigma}^{ref}, \hat{r}_{cut}^{ref}, \hat{\sigma}_m^{gal}, \hat{\sigma}^{gal}, \hat{\phi}^h$.

9.2 Technicalities

To sample the complete posterior in Eq. 9.1, we use the Affine-Invariant sampling as originally introduced by Goodman & Weare (2010), which is especially suited to our highly-dimensional (and possibly degenerate) parameter space. In particular, to enable full portability and reproducibility, we use the latest python release[1] of `emcee` (Foreman-Mackey et al. 2013).

The first four terms in equation Eq. 9.1 are directly implemented in our code, while to compute the multiple-images likelihood p_{im} and the flat priors on the two cluster-scale halo parameters (e^h, θ^h) we exploit `LensTool`. This is also done in order to have a benchmark comparison with a commonly used lensing code.

[1] https://github.com/dfm/emcee

The synergy between our code and `LensTool` requires some technicalities, as described in the following subsections. In any case, our code is fully modular, so that `LensTool` can be replaced by any other parametric lensing software by changing only a few python lines.

9.2.1 Calls to `LensTool`: Input and output files

For each parameters combination, corresponding to a given walker position inside the parameters space, `BayesLens` silently calls `LensTool` to compute the p_{im} term of the total posterior. Every `LensTool` call needs a different input file generated in our code by a specific python-function. Another function reads the resulting likelihood, computed by `LensTool`, from an output file. Since all these output files are saved on the disk using the same name, we create folders with unique (random) names to differentiate each `LensTool` call. These folders are deleted at the end of every p_{im} computation.

To sample the posterior p_{tot}, millions of walker positions are required. So millions of input and output `LensTool` files are quickly created and deleted on the disk during this process. To avoid the disk wear and the bottle neck represented by the process of writing/reading files, part of our computer RAM is reserved to create a *RAM-disk* where the input and output files are temporarily saved and deleted.

9.2.2 From measured σ_m^{gal} to `LensTool` fiducial σ_{LT}

As we showed in Section 3.7 the dPIE mass distribution is implemented in `LensTool` through a fiducial velocity dispersion σ_{LT} related to the 1D central velocity dispersion σ_0 (in Eq. 2.45) by: $\sigma_0 = \sqrt{3/2}\sigma_{LT}$ (see Eq. 3.30).

To convert the model predicted aperture-average velocity dispersions $\hat{\sigma}_m^{gal}$ and $\hat{\sigma}^{gal}$ to their fiducial `LensTool` values ($\hat{\sigma}^{ref}_{(m),LT}$), we relate them through

$$\hat{\sigma}^{gal}_{(m),LT} = \hat{\sigma}^{gal}_{(m)}/c_p(R) \ , \tag{9.13}$$

where R is the aperture radius chosen for the cluster member spectral extraction. Adopting orbital isotropy and a (spherical) galaxy surface brigtness profile proportional to the dPIE matter density, the projection coefficient $c_p(R)$ is given by the Eq. 2.63.

9.3 Functional Tests

Using `LensTool` we create a toy model of a galaxy cluster to test the ability of `BayesLens` in recovering the correct halos and sub-halos mass parameters and in predicting the correct multiple-image positions. `BayesLens` performances are also compared to the results obtained from `LensTool` alone.

Our simple mock galaxy cluster is based on the best-fit lens model developed in Caminha et al. (2017b) (hereafter in this chapter C17b) for the galaxy cluster MACSJ1206 at redshift $z = 0.44$. With respect to C17b, the cluster-scale component of our simulated cluster is made by a single halo, parametrized through a PIEMD mass profile. The center of the halo has an offset of $0.92''$ from the BCG reference position. The core radius of the halo has a value of $r^h_{cut} = 3.0''$, its ellipticity $e = (a^2 - b^2)/(a^2 + b^2)$ is fixed to a value of $e^h = 0.7$, with a position angle of $\theta^h = 19.14°$ counterclockwise from West. The central velocity dispersion has a value of $\sigma^h_{LT} = 1000\,\mathrm{km\,s^{-1}}$. The halo parameters, except the ellipticity and the position angle, are also given in Table 9.2. No shear nor foreground structure are present in our simulated cluster.

The clumpy sub-halo component of the mock cluster is composed by 138 cluster galaxies, selected from the C17b members catalogue to have $m_{F160W} < 22$. All these galaxies are parametrized as circular dPIE profiles whose r_{core}, r_{cut} and σ_{LT} values are determined as follows. For the r_{cut} we adopt the scaling relation in Eq. 3.29 with slope $\beta_{cut} = 0.66$, and a similar relation with $\beta_{core} = 0.5$ is used for the core radius r_{core}. The two normalisations, computed at the BCG luminosity L_0 ($m_{F160W} = 17.1851$), are $r^{ref}_{core} = 0.01''$ and $r^{ref}_{cut} = 5.0''$. We assign to the cluster members line-of-sight stellar velocity dispersions, averaged within a circular aperture, assuming a 20% Gaussian scatter around the scaling relation in Eq. 3.28 with $\alpha = 0.27$ and normalization equal to $350\,\mathrm{km\,s^{-1}}$at the BCG luminosity (see Table 9.1). To determine the `LensTool` fiducial velocity dispersions, we deproject the aperture-averaged velocity dispersions using Eq. 9.13 and assuming apertures of $R = 0.8''$. Note that the BCG measured velocity dispersion has a value of $391.87\,\mathrm{km\,s^{-1}}$, i.e. about $40\,\mathrm{km\,s^{-1}}$ larger than the normalization of the σ-mag scaling relation. This unrealistic mismatch (see Fig. 6.3) is useful to test the `BayesLens` capability to fit the correct scaling relations leaving however the BCG free to scatter (around the best-fit scaling relation) up to the right velocity dispersion.

Given the total mass distribution for the mock cluster, we use `LensTool` to ray-trace the position of 15 sources, randomly selected from the C17b catalogue,

	Scaling relations hyperparameter		
	$\hat{\alpha}$	$\hat{\sigma^{ref}}$ [km s^{-1}]	$\hat{r}^{ref}_{cut}$ [arcsec]
Mock	0.27	350.0	5.00
BL	$0.27^{+0.02}_{-0.02}$	$352.3^{+13.6}_{-14.2}$	$4.95^{+0.71}_{-0.57}$
LT	$0.22^{+0.09}_{-0.09}$	$441.6^{+28.0}_{-28.6}$	$3.26^{+0.72}_{-0.57}$
$\mathrm{LT_{kin}}$	0.27	$396.2^{+11.6}_{-12.5}$	$6.07^{+0.73}_{-0.59}$

Table 9.1: Scaling relations hyperparameter values. In the first row (Mock) the true input values of the mock galaxy cluster are reported. The values in the second, third and fourth rows are the 50^{th}, 16^{th} and 84^{th} percentiles computed from the marginalized posteriors of the BL, LT and $\mathrm{LT_{kin}}$ models respectively (see text).

	Halo parameters			
	$\hat{x}^h$ [arcsec]	$\hat{y}^h$ [arcsec]	$\hat{\sigma}^h_{LT}$ [km s^{-1}]	$\hat{r}^h_{core}$ [arcsec]
Mock	-0.85	0.32	1000.0	3.00
BL	$-0.81^{+0.12}_{-0.13}$	$0.34^{+0.08}_{-0.08}$	$1000.4^{+4.0}_{-4.2}$	$2.99^{+0.25}_{-0.25}$
LT	$-0.52^{+0.09}_{-0.09}$	$0.36^{+0.06}_{-0.07}$	$997.4^{+3.0}_{-3.3}$	$3.01^{+0.18}_{-0.17}$
$\mathrm{LT_{kin}}$	$-0.56^{+0.09}_{-0.09}$	$0.33^{+0.06}_{-0.06}$	$996.5^{+2.9}_{-3.0}$	$3.06^{+0.16}_{-0.16}$

Table 9.2: Same as Table 9.1 but for cluster-scale halo parameters.

	Galaxies velocity dispersions [km s^{-1}]			
	$\hat{\sigma}^{gal}_{m,BCG}$	$\hat{\sigma}^{gal}_{m,1}$	$\hat{\sigma}^{gal}_{2}$	$\hat{\sigma}^{gal}_{3}$
Mock	391.9 ± 14.2	255.0 ± 11.7	275.2	242.9
BL	$392.9^{+15.0}_{-15.9}$	$254.8^{+11.9}_{-12.7}$	$286.2^{+18.9}_{-20.9}$	$244.6^{+16.7}_{-15.3}$
LT	$441.6^{+28.0}_{-28.6}$	$257.0^{+14.8}_{-13.8}$	$359.7^{+20.7}_{-20.6}$	$304.3^{+17.0}_{-16.4}$
$\mathrm{LT_{kin}}$	$396.2^{+11.6}_{-12.5}$	$216.0^{+5.3}_{-5.8}$	$315.4^{+8.7}_{-9.4}$	$261.7^{+6.8}_{-7.4}$

Table 9.3: Same as Table 9.1 but for the four cluster member velocity dispersions of Fig. 9.4.

to their multiple images on the lens plane. The sources are within a redshift range of [1.01 - 6.06] and produce a total of 82 multiple images. 15 of these images are excluded from the final set, either because they are strongly de-magnified (more than 3 magnitudes fainter than the source luminosity), or because they are too close to a cluster member to be really observable. We also exclude the 10^{th} family constituted by a single lens image. Therefore, our final mock multiple images catalogue consists in 66 multiple images, shown in Fig. 9.1. The final simulated cluster, despite a purposely simple mass distribution, contains a number of sub-halos and multiple images comparable to those of most CLASH or HFF clusters.

In our tests, we consider an error of $0.5''$ on the multiple-image positions. We also suppose that the stellar velocity dispersions of 58 luminous cluster galaxies are measured, within apertures of $R = 0.8''$. This corresponds to the number of velocity dispersions that we were able to measure in Chapter 6, from the MUSE datacube of

MACS 1206. To associate an error to these simulated measures we use the following empirical relation derived from the measurements in Chapter 6:

$$\frac{\delta\sigma_m^{gal}}{\sigma_m^{gal}} = 1.6 \times 10^{-3}\, m_{F160W}^3 - 8.53 \times 10^{-2}\, m_{F160W}^2 + 1.509\, m_{F160W} - 8.879 \quad (9.14)$$

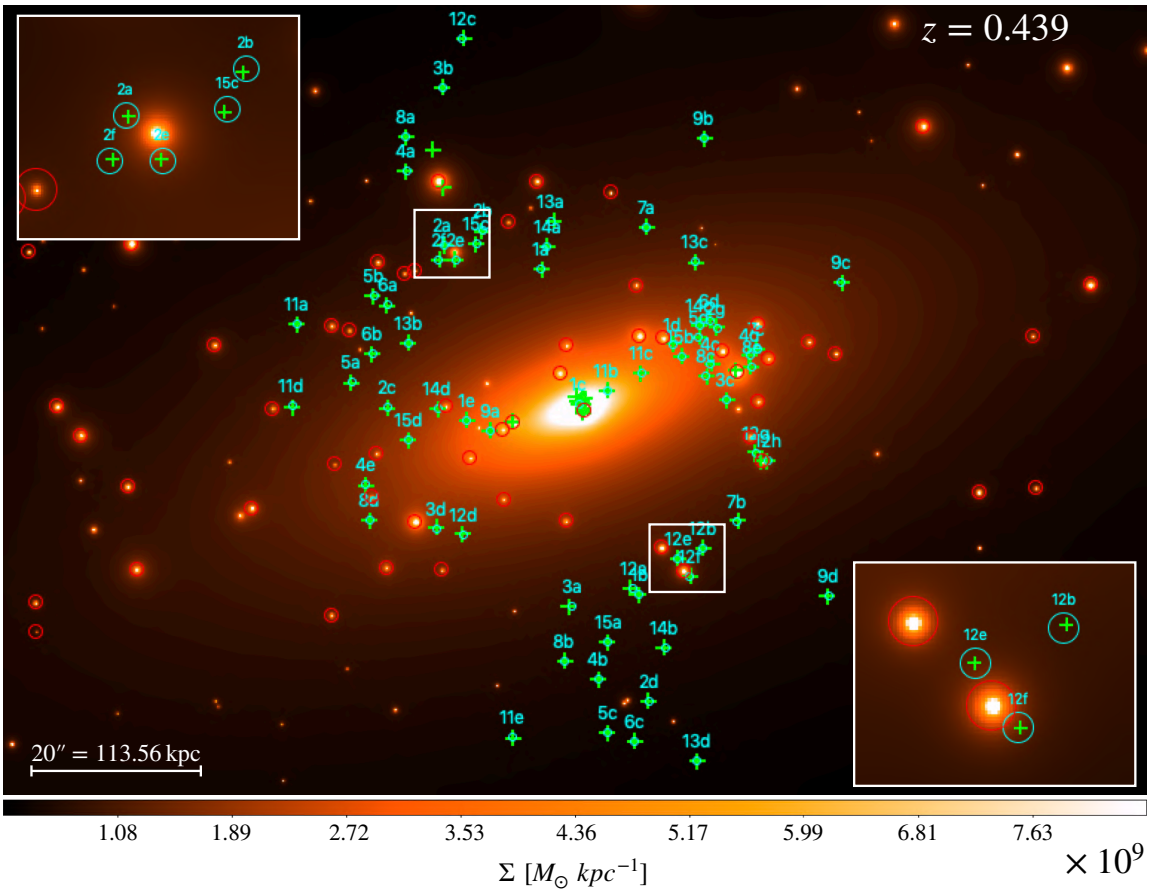


Figure 9.1: Mass density distribution, colour coded in $\mathrm{M}_\odot\,\mathrm{kpc}^{-1}$, of the central region of the mock cluster used for our functional tests, loosely based on the Caminha et al. (2017b) lens model for the cluster MACS J1206.2−0847 at redshift $z = 0.44$. Red circles mark the galaxies for which we have a 'measured' velocity dispersion. Our working hypothesis is that these velocity dispersions are measured within R=0.8″ apertures, displayed by the red circles. The small cyan circles are the mock observed multiple image positions. Their positions are known with an isotropic error of 0.5″. The green crosses mark the `BayesLens` predicted multiple images positions. The insets show two galaxy-scale strong lensing systems in the inner region of our mock cluster.

9.3.1 LensTool on the mock cluster

In this subsection, we describe the parameter ranges adopted in LensTool to probe the mass distribution of the mock cluster.

The cluster-scale component is parametrized as a single PIEMD profile. The x, y coordinates of its center can vary within flat priors, $3''$ wide, centered on the BCG position. For the core radius $\hat{r}^h_{core}$, ellipticity e^h, position angle $\hat{\theta}^h$ and fiducial velocity dispersion $\hat{\sigma}^h_{LT}$, we adopt uniform priors inside the following intervals respectively: $[1.0''$ - $7.0'']$; [0.0 - 0.9]; $[5.0°$ - $35.0°]$; and $[700.0$ - $1300.0]\,\mathrm{km\,s^{-1}}$.

The sub-halos component of the cluster is parametrized using the two scaling relations in Eqs. 3.28 and 3.29 normalized at the BCG luminosity, plus a similar relation for r_{core}. We try two alternatives for the priors on the scaling relation parameters. In a first model (hereafter identified as LT), we adopt $\hat{r}^{ref}_{core} = 0.01''$ and $\hat{\beta}_{core} = 0.5$, while $\hat{\alpha}$, $\hat{\sigma}^{ref}_{LT}$, $\hat{\beta}_{cut}$ and $\hat{r}^{ref}_{cut}$ are chosen within flat priors between [0.20 - 0.33], $[250.0$ - $450.0]\,\mathrm{km\,s^{-1}}$, [0.5 - 1], and $[1''$ - $11'']$ respectively. In a second model (hereafter identified as $\mathrm{LT_{kin}}$), we derive the scaling relation parameters optimizing, through a Bayesian approach, the posterior p_{sr} (in Eq. 9.2) using the 58 measured cluster members stellar velocity dispersions. The same procedure was adopted in Chapter 6 to develop the cluster lens models with kinematic priors. For the mock cluster we find a slope $\hat{\alpha} = 0.27$ and a Gaussian prior on the lens model reference sigma $\hat{\sigma}^{ref}_{LT}$ with mean $311\,\mathrm{km\,s^{-1}}$ and standard deviation $18\,\mathrm{km\,s^{-1}}$ (see Sec. 9.1.2). Exploiting Eq. 6.1 and assuming $\gamma = 0.2$, we derive also $\hat{\beta}_{cut} = 0.66$. This value of γ describes well the tilt of the canonical fundamental plane relation for elliptical galaxies. The other scaling relation parameters keep the same values abopted in LT model.

The LT and $\mathrm{LT_{kin}}$ optimizations are performed on the lens-plane using the positions of the 66 multiple images inside the simulated catalogue. The posterior probabilities are sampled using 10 walkers and the final MCMC chains are $\sim 2 \times 10^5$ and $\sim 8 \times 10^4$ long for LT and $\mathrm{LT_{kin}}$ respectively, burn-in excluded.

The predicted multiple images positions and the final total Δ_{rms} are determined from the best-fit cluster lens models.

9.3.2 BayesLens on the mock cluster

For the cluster-scale halo and for the sub-halo $\hat{\alpha}$, $\hat{\beta}_{core}$, $\hat{r}^{ref}_{core}$ and $\hat{r}^{ref}_{cut}$ parameters, we adopt the same priors, or fixed values, as in the LT model. As in $\mathrm{LT_{kin}}$, $\hat{\beta}_{cut}$ is derived from $\hat{\alpha}$ through Eq. 6.1 with $\gamma = 0.2$, however now $\hat{\alpha}$ is one of the free

parameters of the model.

Contrary to LT and $\mathrm{LT_{kin}}$, in our code the σ - mag scaling relation refers to measured aperture-averaged stellar velocity dispersions. For its normalization, we adopt a uniform prior between 250.0 and 450.0 $\mathrm{km\,s^{-1}}$. In the `BayesLens` model (hereafter BL), the 58 measured stellar velocity dispersions are used on one hand to determine the best σ - mag scaling relation parameters, and on the other to derive a Gaussian prior for each measured galaxy. Thanks to these priors, the lens model tends to prefer solutions with cluster members velocities that coincide with the measured values (see Eq. 9.6), unless the lensing data require them to deviate. The presence of a non-zero scatter also enables the models to fairly sample the whole parameter space, thereby avoiding the underestimation of systematics.

To derive the term p_{im} of the total posterior (Eq. 9.11), for each `emcee`-walker position, we performed a fast `LensTool` source-plane optimization of 300 MCMC chain-steps.

The total posterior p_{tot} is sampled using 500 walkers and from the final MCMC chain (composed by 2.5×10^6 steps) we remove a burn-in phase wider than the largest autocorrelation time of the optimized parameters. To determine the predicted multiple image positions (Fig. 9.1) and the final total Δ_{rms} we exploit `LensTool` trough the following procedure. Fixing all the sub-halos parameters to their best-fit BL model values, we perform an image plane `LensTool` maximum-likelihood optimization leaving all the cluster-scale halo parameters free to vary within their uniform priors. The multiple images positions and the predicted Δ_{rms} are obtained using this final best-fit `LensTool` model. This procedure is necessary because in this first version of `BayesLens` both $\hat{e}^h$ and $\hat{\theta}^h$ are optimize outside our code thought `LensTool` software-calls (see Sec. 9.2).

9.4 Results and discussion

While in LT, $\mathrm{LT_{kin}}$ and BL models we use the same parameterization and prior ranges for the cluster-scale halo component, the sub-halos are described in different ways. In LT the σ - mag and r_{cut} - mag scaling relation parameters are left free to vary within uniform priors and zero-scatter scaling relations are assumed. In $\mathrm{LT_{kin}}$ we used the 58 mock measured stellar velocity dispersions to fix the scaling relation slopes and to define a Gaussian prior on the σ - mag normalization. However, zero-scatter scaling relations are still assumed. Finally, in BL we use the cluster members stellar kinematics to derive informative Gaussian priors on the velocity

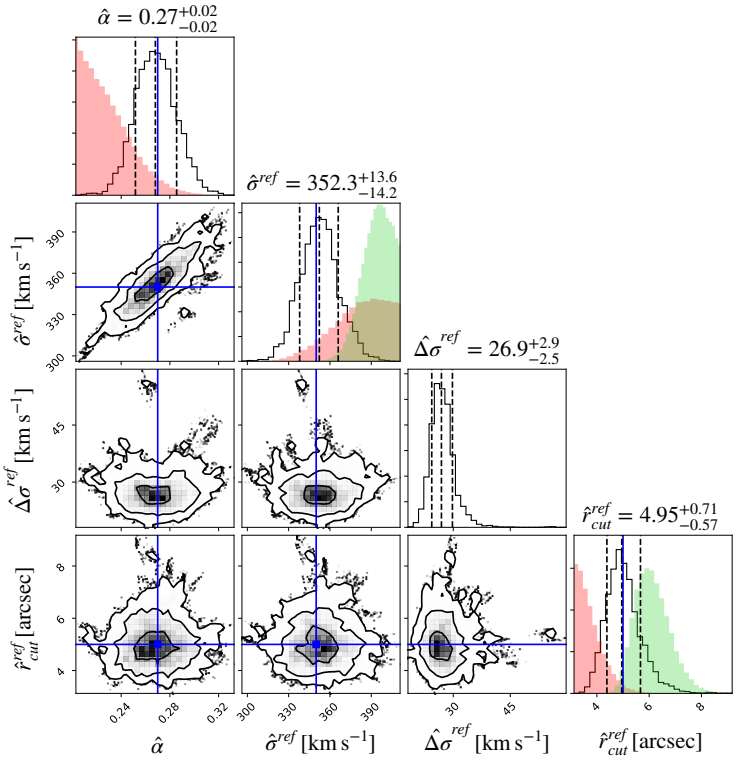


Figure 9.2: In black the BL (see text) marginalised posterior distributions on the hyperparameters of the σ - mag and r_{cut} - mag scaling relations are shown. In the plot we omit $\hat{\beta}_{cut}$ because its value is not optimized in `BayesLens` but directly derive from Eq. 6.1. The contours limit the 1σ, 2σ and 3σ regions. The black dashed vertical lines in the histograms correspond to the 16^{th}, 50^{th} and 84^{th} percentiles of the distributions (these values are reported in the titles). The blue solid lines are the true values of the parameters of the mock cluster. The red and green shaded areas are the LT and $\mathrm{LT_{kin}}$ results respectively.

dispersions of the measured galaxies. The unmeasured galaxies are left free to scatter around the best-fit scaling relation obtained fitting the 58 measured stellar velocity dispersions.

In Figs. 9.2 and 9.3 we show in black the marginalized `BayesLens` posterior distributions for the scaling relations and cluster-scale halo parameters, respectively. The true parameter values assumed in the mock cluster are marked with solid blue lines, while the LT and $\mathrm{LT_{kin}}$ posteriors are the red and green shaded distributions in the histograms. Similarly, Fig. 9.4 shows the marginalized posterior distributions for the stellar velocity dispersions of four cluster galaxies. $\hat{\sigma}^{gal}_{m,BCG}$ and $\hat{\sigma}^{gal}_{2}$ are the velocity dispersions of the BCG and of the brightest galaxy without a measured velocity dispersion respectively, while $\hat{\sigma}^{gal}_{m,1}$ and $\hat{\sigma}^{gal}_{3}$ correspond to the two galaxy scale strong lensing systems zoomed in Fig. 9.1.

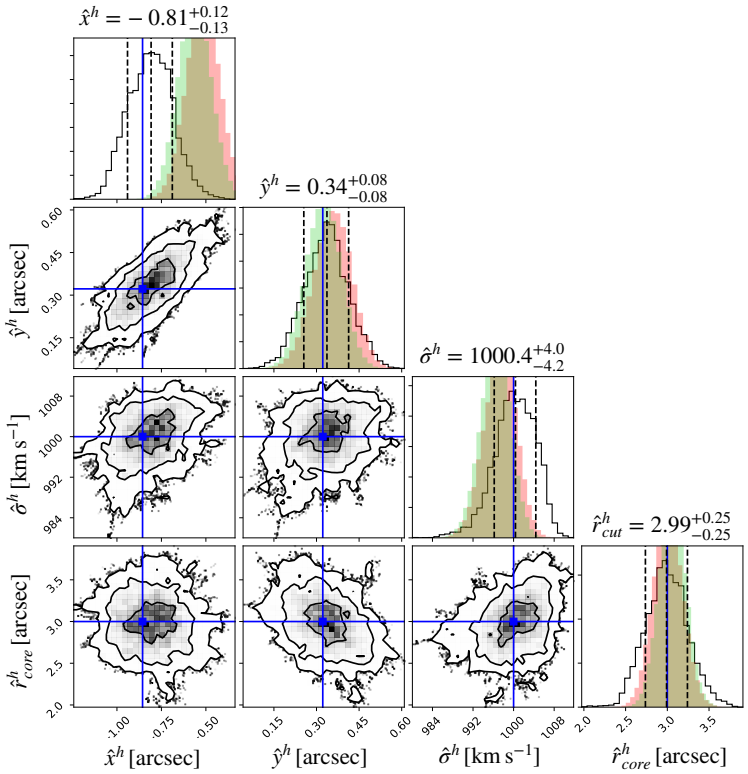


Figure 9.3: Same as Fig. 9.2 but for the cluster-scale halo parameters. We omit $\hat{e}^h$ and $\hat{\theta}^h$ distributions because they are separately optimized by `BayesLens` (see Sec. 9.2).

Thanks to the measured cluster members stellar velocity dispersions, `BayesLens` recovers, well within the 1σ uncertainties, all the of mock true parameters of the scaling relations (see Fig. 9.2). Conversely, the LT model, constrained by the multiple images positions alone, is not able to retrieve the correct slopes and normalizations: $\hat{\sigma}^{ref}$ is overestimated by ~ 90 km s^{-1}, while $\hat{r}^{ref}_{cut}$ and $\hat{\alpha}$ are underestimations of the true values (see Table 9.1). The too high σ-mag normalization is also responsible for the mismatch visible in Fig. 9.4 (and Table 9.3) between the $\hat{\sigma}^{gal}_2$ and $\hat{\sigma}^{gal}_3$ distributions and the mock true values. In the LT$_{\rm kin}$ model, the optimized scaling relation parameters $\hat{\sigma}^{ref}$ and $\hat{r}^{ref}_{cut}$ are better constrained than LT, but their values still differ, by more than 1σ, from the mock inputs (see Table 9.1). Indeed, in LT$_{\rm kin}$ the value of the normalization $\hat{\sigma}^{ref}$ is close to the BCG measured velocity dispersion (see Fig. 9.4 and Table 9.3). This is justified by the fact that the BCG mass influences the multiple images positions more than the other less massive galaxies and that in LT$_{\rm kin}$ we assume a zero-scatter scaling relation normalized to the BCG luminosity. This result does not question the robustness of the lens models in Chapters 6 and 7 since it is only due to the large and unrealistic mismatch (40 km s^{-1}) that we assume between the BCG and the reference velocity dispersion, σ^{ref}, in our mock

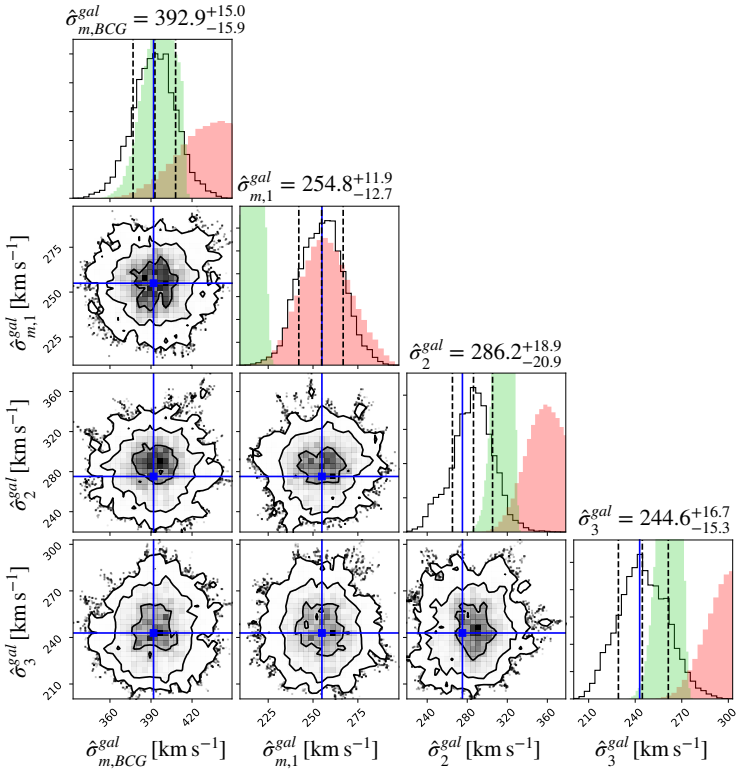


Figure 9.4: Same as Fig. 9.2 but for four cluster members velocity dispersions. To derive the `LensTool` distribution for a given cluster galaxy (red and green shaded areas) we use the scaling relations (Eqs. 3.28 and 3.29) and we project the fiducial `LensTool` velocity dispersion inside aperture of radius R=0.8″. $\hat{\sigma}^{gal}_{m,BCG}$ and $\hat{\sigma}^{gal}_{2}$ are the aperture-average velocity dispersion of the BCG (about 40 km s^{-1} larger than $\hat{\sigma}^{ref}$, see text) and of the brightest galaxy without a measured velocity dispersion respectively. $\hat{\sigma}^{gal}_{m,1}$ and $\hat{\sigma}^{gal}_{3}$ correspond to the two galaxy-scale strong-lensing systems identified in Fig. 9.1.

cluster. This mismatch is definitely not observed in the four clusters MACSJ1206, MACSJ0416, AS1063 and A2163 studied in Chapters 6 and 7 (see Fig. 6.3).

In BL we recover all the cluster-scale halo parameters, well within the 1σ uncertainties, while in LT and LT$_{\rm kin}$ the separation between the true and predicted center, $\hat{x}^h$, is significantly discrepant.

Even though the total mass of the mock cluster is redistributed in a slightly different way between the cluster halo and the member galaxies among different model classes, its value at large radii is accurately recovered by all models. This is visible in Fig. 9.5, which shows the cumulative total mass profiles of the mock cluster predicted by the LT, LT$_{\rm kin}$ and BL models. The LT$_{\rm kin}$ and BL models predict an almost identical profile for both the total mass of the cluster and the cluster member

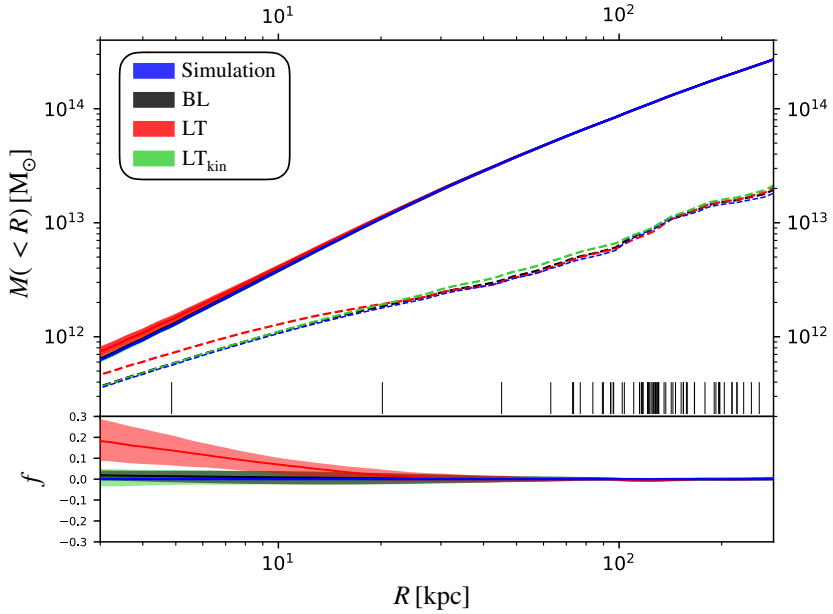


Figure 9.5: *Top*: Projected cumulative total mass profiles for the mock cluster as a function of the projected distance from the center of the BCG. The blue line corresponds to the true mass profile. The coloured areas are bounded by the 16^{th} and the 84^{th} percentiles of the distribution obtained using BayesLens (in black), `LensTool` (in red) and `LensTool` with the kinematic prior on the scaling relation parameters (in green). The coloured solid lines are the median values. The projected distances of the multiple images from the BCG center are shown as vertical black lines. *Bottom*: Relative variation of the total mass profiles with respect to the true mass profile of the mock cluster.

component. Differently, the LT model predicts a higher mass for the cluster member component in the inner region, which can be explained by an overestimated BCG velocity dispersion. However, at the distances where most of the multiple images form, yielding more constrains to the lens model, also the LT model predicts a correct value for the mass profiles.

To quantify the accuracy of the BL, LT and LT$_{\rm kin}$ models in reproducing the multiple image positions we show in Fig. 9.6 the x and y offsets between model-predicted and mock true positions. The better mass characterization obtained using `BayesLens`, with the additional information coming from the cluster member velocity dispersions, results in a lower root-mean-square separation between the mock observed and the model predicted multiple images positions on the lens plane (Δ_{rms}). In particular, for our simple mock cluster, the BL model yield a total $\Delta^{BL}_{rms} = 0.13''$. This value is ~ 3 times smaller than the LT$_{kin}$ and LT models, where a zero-scatter scaling relation is adopted.

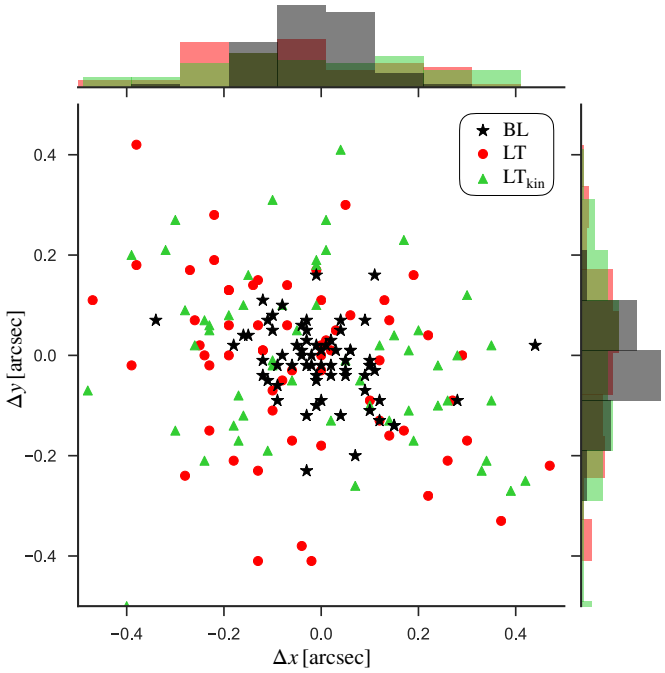


Figure 9.6: Offsets between the true 'observed' multiple image positions of sources lensed by our simple toy cluster model and their predicted positions as obtained by the BL (black stars), LT (red dots) and $\mathrm{LT_{kin}}$ (green triangles) lens models (see text). The histograms show the distribution of the offsets along the x and y directions.

Finally, in Fig. 9.7 we show the mock cluster members measured stellar velocity dispersions σ^{gal}_{true} against the values predicted by the three lens models. Both the LT and the LT_{kin} models show a large scatter on the measured velocity dispersions. Furthermore the LT model (red dots in the figure) systematically overestimate the cluster members velocity dispersions due to the too large σ - mag normalization highlighted before. The introduction, in `BayesLens`, of Gaussian priors centered on the measured cluster members velocity dispersions and the inclusion of the scatter around the best-fit scaling relation has the effect to appreciably reducing the separation between the model-predicted and the true velocity dispersions.

9.5 Conclusions

In this chapter, we have shown that an accurate description of cluster member galaxies is attainable, despite the large number of degrees of freedom, through fast and tractable hierarchical Bayesian inference. This extends beyond the state-of-the-art of current cluster lensing models. Indeed, unlike conventional models

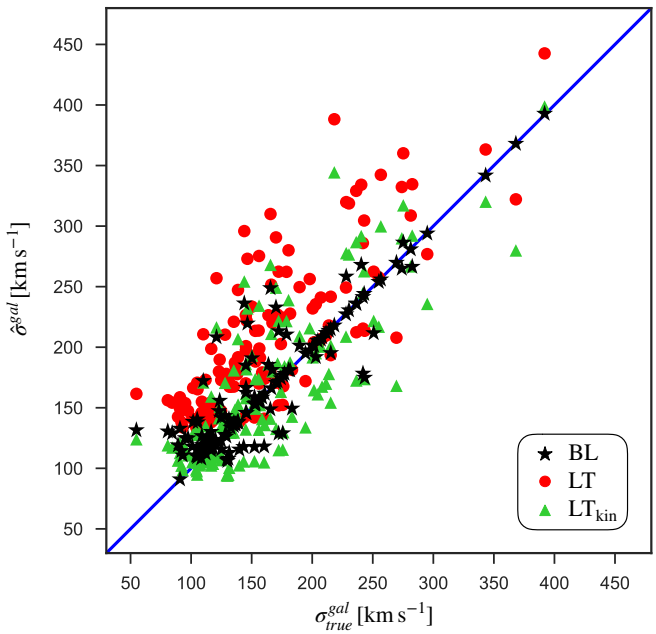


Figure 9.7: Velocity dispersions of cluster-member galaxies from the BL, LT and $\mathrm{LT_{kin}}$ lensing models (y-axis) versus their true values in the mock (x-axis). Symbols are as in Fig. 9.6. Zero-scatter models have either offset scaling relations or large r.m.s. image offsets. Our hierarchical model recovers the correct cluster member velocity dispersions of galaxies in our simple toy cluster model.

where galaxies are placed on a razor-thin scaling relation and some are 'freed' ad hoc, we populate a non-zero scatter scaling relation with hyperparameters that are inferred directly from the lensing constraints and (when available) stellar kinematic data.

The freedom in the parameters of each cluster-member galaxy, within the intrinsic scatter of their parent population, has multiple implications. First, the data themselves dictate which galaxy should be 'freed' and deviate significantly from a baseline scaling relation. Second, intrinsic scatter is a hyperparameter that is left free in the modeling inference, and this allows for a direct determination of galaxy-population properties.

An important feature of our inference is that the code is fully modular. This allows for calls to any chosen lensing code (besides `LensTool`, used here as the benchmark), as well as different prescriptions to relate the stellar velocity dispersions to the lensing parameters of cluster-member galaxies. Similarly, the currently used scaling relations may be easily replaced with fundamental-plane relations (with free hyperparameters), so as to study the evolution of the fundamental plane across

redshift and environment. Its modular structure also enables the use of additional constraints from e.g. flux ratios, extended-source reconstruction, time delays, etc.

So far, we tested or code only on a simple toy cluster model. In this case most of the cluster member velocity dispersions, and all the halo parameters and scaling relation hyperparameters are recovered within their 68% credibility ranges. Moreover the r.m.s. displacement of multiple images between observations and model predictions is ~ 3 times smaller than the zero-scatter model results.
We should emphasize that this is a performance test on mocks, and additional effects may play a role in real-life systems, such as faint substructure, deviations from simple geometry of the main DM halo(s) and cluster members, and additional contributions along the line of sight. While BL might eliminate part of the systematics in the lensing models, its precise impact has to be quantified on the real galaxy clusters.

Chapter 10

Conclusions and future perspectives

Strong gravitational lensing is one of the most powerful tools to probe the total mass distribution of galaxies and the cores of galaxy clusters. In recent years, remarkable progress has been made on the observational side, with dedicated HST imaging campaigns on sizeable samples of massive galaxy clusters, such as CLASH (Postman et al. 2012) and the Hubble Frontier Fields (Lotz et al. 2017) programs. Most importantly, dedicated spectroscopic programs, such as CLASH-VLT (Rosati et al. 2014), GLASS (Treu et al. 2015; Schmidt et al. 2014) and extensive VLT/MUSE observations have led to the development of high-precision strong lensing models, via the identifications of a large number of multiple images and cluster members. Spectroscopic observations have been the key to this transition to high-precision lens models by reducing significantly biases due to misidentifaction of multiple images via photometric techniques and distance-mass degeneracies in model parameters, which affected previous lens models largely based on imaging data sets.

These new models have provided robust total mass maps and mass density profiles in the inner regions of massive clusters, which now can be effectively compared with state-of-the-art cosmological simulations to test the foundations of the ΛCDM structure formation paradigm.

In this thesis, I have used new high-quality imaging and spectroscopic data sets, to further improve the accuracy and precision of cluster lens models by incorporating stellar kinematic information of a significant fraction of cluster galaxies. Specifically, I have focused my study on three massive galaxy clusters (MACS J1206.2−0847, MACS J0416.1−2403, and Abell S1063) at $z \sim 0.4$ with HST and deep MUSE observations for which high-precision models have recently been developed (Caminha

et al. 2017b,a, 2016; Bonamigo et al. 2018). These parametric models have taken advantage of the spectroscopic identification of 60-100 multiple images per cluster, as well as highly-complete and pure samples of cluster galaxies, to reproduce the position of multiple images with a rms accuracy of $\sim 0.5''$. In addition, the diffuse mass component in these clusters incorporates the gas mass distribution, traced by high-angular resolution *Chandra* observations, which dominates the baryonic mass budget. Despite the significant improvement of the these lens models, the parameters describing the sub-halo mass distribution (i.e. the clumpy mass component associated to cluster members), which is assumed to follow pre-determined scaling relations consistent with the cluster fundamental plane, may still suffer from possible degeneracies with other cluster mass parameters.

To reduce such degeneracies and avoid possible biases in the description of the sub-halo mass distribution in these three clusters, in this work I have measured the stellar velocity dispersion of 40-60 cluster members per cluster from high signal-to-noise MUSE spectra, and have incorporated this kinematic information in the sub-halo scaling relations of the lens models.

In detail, I summarize below the main results of my PhD thesis:

- From the MUSE datacubes of the three clusters MACS J1206.2−0847, MACS J0416.1−2403, Abell S1063 (a.k.a. RXC J2248.7−4431), we extracted about 70-120 cluster member spectra uncontaminated from nearby sources. Exploiting the public software `pPXF` (Cappellari & Emsellem 2004; Cappellari 2017), we measured stellar velocity dispersions for about 40-60 cluster galaxies covering a wide luminosity range of about 4-5 magnitudes, down to $m_{F160W} \sim 21.5$, i.e. ~ 2.5 mag below $L^\star$ (see Section 6.2). The robustness of `pPXF` measurements was tested through extensive spectra simulations in Section 5.4. In particular, the simulations were used to: verify the presence of systematic uncertainties in `pPXF` fitted parameters; quantify realistic statistical errors on measured velocity dispersions; establish a lower $\langle S/N \rangle$ of the spectra for reliable kinematics measurements; determine a lower limit for measured velocity dispersion due to MUSE instrumental resolution; determine the optimal wavelength range for `pPXF` fits.

- Extending the work by Elíasdóttir et al. (2007), we derived an expression for the projection coefficient, $c_p(R)$, which is needed to translate the measured velocity dispersions of cluster galaxies into central velocity dispersions, σ_0, setting the normalization of the cluster member mass profiles in the lens models. The projection coefficient is the result of a double numerical integration of a function that depends on the core (r_{core}) and truncation radius

(r_{cut}) of the galaxy mass profiles (dual pseudo-isothermal spheres) and on the aperture R chosen for the spectral extraction.

- Using a Bayesian approach, we determined the best-fit L-σ Faber-Jackson relation in each cluster. The posterior distributions of the normalization and slope of these relations were used as kinematic priors for the cluster member scaling relations in our lens models. Thus, we developed for the first time cluster lens models which incorporates kinematic priors on cluster member scaling relation parameters (see Section 6.3).

- The introduction of kinematic priors on cluster member scaling relations significantly reduced inner degeneracies of cluster lens models, by restricting the parameter space of the sub-halo component, making it consistent with galaxy kinematics. Thus, sub-halos masses and sizes are now robustly constrained by lens models (see Section 6.5), within the parametrization of sub-halos with truncated pseudo-isothermal spheres. On the other hand, the total mass projected density profiles are not altered significantly by the new models.

- Once re-normalized to the same absolute luminosity, the reference values of the cluster member scaling relations (one for the central velocity dispersion σ_0 and one for the truncation radius r_{cut}), independently derived for each one of the three clusters, were found in very good agreement. As a result, we statistically inferred the M_{tot}-σ_0 and r_{cut}-σ_0 relations, between the total mass, the truncation radius and the central velocity dispersion of cluster sub-halos (see Section 6.5).

- Extending previous findings by Grillo et al. (2015) and Bonamigo et al. (2018), we derived fully consistent velocity dispersion (or circular velocity) and mass functions for the three clusters. This is due to the robust characterization of the sub-halo masses offered by our cluster lens models with kinematic priors (see Section 6.5.1).

- In Chapter 8, we compared the results of our cluster lens models with kinematic priors with the predictions of high-resolution N-body and hydrodynamical cosmological simulations. The sub-halo mass functions of observed and simulated cluster were interestingly found in good agreement. However, a discrepancy in the compactness and radial distribution of cluster substructures emerged, which we interpreted as the reason for the mismatch in the circular velocity function between observations and simulations, first reported in Grillo et al. (2015). In particular, observed sub-halos appear more compact (higher

circular velocity and smaller size) and more concentrated toward the cluster centers than sub-halos in hydro-simulations for a given mass. These differences are at the basis of a one-order-of-magnitude mismatch between the number of galaxy scale strong lensing events observed in real and simulated clusters (Meneghetti et al., submitted).
We still do not understand whether the origin of the lack of compact sub-structures in inner regions of simulated clusters has to be ascribed to numerical or resolution effects in the simulations, or some physical aspect missing in the CDM framework. The development of new higher resolution cosmological simulation and a detailed study of mass profiles of a large number of real cluster galaxies will be fundamental to discriminate among alternative solutions.

- A promising technique to study in more detail the mass distribution of single cluster members is to exploit galaxy-galaxy strong lensing systems (GGSLS). In such systems, single member galaxies act as small scale gravitational lenses embedded in the deep cluster-scale potential. In this case, cluster galaxies can produce several strongly magnified (up to tens times) multiple images on a scale of few kpc around their critical lines. In Chapter 7, we developed high-precision lens models (rms accuracy between $0.03''$ up to $0.15''$) for three GSSLSs belonging to MACS J1206.2−0847 and Abell 2163. Moreover, we used the measured velocity dispersion radial profiles of two lens galaxies and a single aperture velocity dispersion in another system to further reduce inner degeneracies among their mass free-parameters, particularly their truncation radii.

- In Chapter 9, we presented a preliminary version of a python code used to incorporate the measured cluster member stellar kinematics into common (external) lensing codes, such as `LensTool`. Moreover, the code exploits a fast and tractable hierarchical Bayesian inference formalism to include the scatter of cluster galaxies around the best-fit Faber-Jackson cluster member scaling relation in lens models. So far, we have not applied our python code to real galaxy clusters, however the first tests on a simple mock cluster appear very promising.

In the near future, we plan to develop lens+kinematic models for a larger sample of galaxy clusters. We mention below a number of interesting developments and further refinements of the methodology developed in this thesis.

1. It will important to include the observed scatter in the sub-halo scaling relations, which is still assumed zero in current modeling, by directly using the

measured velocity dispersions of cluster galaxies. This is now possible with the technique developed in Chapter 9.

2. One can verify the consistency of sub-halo mass (and velocity dispersion) functions between galaxy clusters at different redshifts. In particular, we can study the evolution of sub-halo mass functions over cosmic time and study their dependence from the mass of the host cluster. In Chapter 8, we found a remarkably agreement between the sub-halo mass functions of observed and simulated clusters, however a significant mismatch was discovered between the velocity dispersion functions. A larger sample of cluster lens models with a well characterized sub-halo population, as well as higher-resolution cosmological simulations, may help to shed more light on the reasons at the basis of this tension.

3. One can test the universality of the M_{tot}-σ_0 and r_{cut}-σ_0 relations in Eqs. 6.9 and 6.8, linking the three parameters that characterize the mass distribution of cluster galaxies (i.e. the central velocity dispersion, σ_0, the scale radius, r_{cut}, and the total mass, M_{tot}). Studying the evolution of these statistical relations over redshift, or as a function of the host cluster mass, we can obtain information about the strength of the physical processes that shape the sub-halo mass profiles (e.g. tidal striping, gas cooling, AGN and stellar feedback, etc.).

4. By extending the sample of galaxy-scale strong-lensing systems in clusters, more detailed structural information on DM halo of cluster galaxies can be obtained. In particular, one can study the radial dependence of galaxy halo parameters. For example, Fig. 8.8 shows (colored dots) that hydro-simulations predict a radial dependence of the sub-halo truncation radii. Moreover, the measurement of DM fractions and DM fraction profiles, after an accurate estimate of the stellar mass, can provide further important tests for cosmological simulations.

We should then mentioned two very important applications of high-precision lens models, such as those developed in this thesis.

Firstly, the use of lensing clusters as powerful cosmic gravitational telescopes relies entirely on the ability to robustly measure magnification maps, particularly in the high-magnification regime. This is now a mature science, and recent studies using new high-precision lens models have shown how one can discover and characterize the intrinsic physical properties (e.g. star-formation rates, stellar masses, sizes) of a new population of low luminosity, low mass ($\gtrsim 10^6\,M_\odot$) "proto-galaxies" at high-redshift (e.g. Vanzella et al. 2017; Karman et al. 2017).

Secondly, the discovery of the multiply lensed SN Refsdal at $z = 1,49$ (Kelly et al. 2015, 2016) in MACS J1149.5+2223 ($z = 0.54$), in combination with high-precision strong lensing modeling of the cluster mass distribution (Treu et al. 2016, again using HST imaging and deep spectroscopy) has opened the way to measure the Hubble constant exploiting the measured time delay between lensed SN images (Grillo et al. 2018). This is a technique which has proven very successful when applied to galaxy-scale systems in the field (Suyu et al. 2017; Birrer & Treu 2019). The combination of strong lensing and cluster galaxies' kinematics, that we have studied in this work, may have an important application in lens models of clusters with multiply lensed variable sources, such as SNe or QSO, to be discovered in the future.

We anticipate that the methodology we have started developing in this thesis, and its future refinements, will play a relevant role in the next decade, when a very large number of strong lensing systems will be discovered in large area surveys, such as LSST (Large Synoptic Survey Telescope) and Euclid. In addition, these techniques will find relevant applications in spectro-photometric studies of lensing clusters with facilities such as the James Webb Space Telescope (JWST) and the European-Extremely Large Telescope (E-ELT).

References

Agnello, A., Evans, N. W., & Romanowsky, A. J. 2014, MNRAS, 442, 3284

An, J. H. & Evans, N. W. 2011, MNRAS, 413, 1744

Auger, M. W., Treu, T., Bolton, A. S., et al. 2010, ApJ, 724, 511

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

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

Barnabè, M., Czoske, O., Koopmans, L. V. E., et al. 2009, MNRAS, 399, 21

Bartelmann, M. & Schneider, P. 2001, Phys. Rep., 340, 291

Beckwith, S. V. W., Stiavelli, M., Koekemoer, A. M., et al. 2006, AJ, 132, 1729

Bender, R., Burstein, D., & Faber, S. M. 1992, ApJ, 399, 462

Benítez, N. 2011, BPZ: Bayesian Photometric Redshift Code

Bergamini, P., Rosati, P., Mercurio, A., et al. 2019, arXiv e-prints, arXiv:1905.13236

Binney, J. & Mamon, G. A. 1982, MNRAS, 200, 361

Birrer, S. & Treu, T. 2019, Monthly Notices of the Royal Astronomical Society, 2172

Biviano, A., Rosati, P., Balestra, I., et al. 2013, A&A, 558, A1

Blumenthal, G. R., Faber, S. M., Primack, J. R., & Rees, M. J. 1984, Nature, 311, 517

Blumenthal, G. R., Pagels, H., & Primack, J. R. 1982, Nature, 299, 37

Bolton, A. S., Burles, S., Koopmans, L. V. E., Treu, T., & Moustakas, L. A. 2006, ApJ, 638, 703

Bolton, A. S., Treu, T., Koopmans, L. V. E., et al. 2008, ApJ, 684, 248

Bonafede, A., Dolag, K., Stasyszyn, F., Murante, G., & Borgani, S. 2011, MNRAS, 418, 2234

Bonamigo, M., Grillo, C., Ettori, S., et al. 2018, ApJ, 864, 98

Bond, J. R., Szalay, A. S., & Turner, M. S. 1982, Physical Review Letters, 48, 1636

Borgani, S. & Kravtsov, A. 2011, Advanced Science Letters, 4, 204

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

Brainerd, T. G., Blandford, R. D., & Smail, I. 1996, ApJ, 466, 623

Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503

Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682

Caminha, G. B., Grillo, C., Rosati, P., et al. 2016, A&A, 587, A80

Caminha, G. B., Grillo, C., Rosati, P., et al. 2017a, A&A, 600, A90

Caminha, G. B., Grillo, C., Rosati, P., et al. 2017b, A&A, 607, A93

Caminha, G. B., Rosati, P., Grillo, C., et al. 2019, arXiv e-prints, arXiv:1903.05103

Cappellari, M. 2017, MNRAS, 466, 798

Cappellari, M. & Emsellem, E. 2004, Publications of the Astronomical Society of the Pacific, 116, 138

Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011, MNRAS, 413, 813

Carroll, S. M. 1997, arXiv e-prints, gr

Cerny, C., Sharon, K., Andrade-Santos, F., et al. 2018, ApJ, 859, 159

Chabrier, G. 2003, PASP, 115, 763

Chua, K. T. E., Pillepich, A., Rodriguez-Gomez, V., et al. 2017, MNRAS, 472, 4343

Coe, D., Salmon, B., Bradac, M., et al. 2019, arXiv e-prints [[arXiv]1903.02002]

Coe, D., Zitrin, A., Carrasco, M., et al. 2013, ApJ, 762, 32

Connor, T., Donahue, M., Kelson, D. D., et al. 2017, ApJ, 848, 37

Czoske, O., Barnabè, M., Koopmans, L. V. E., Treu, T., & Bolton, A. S. 2008, MNRAS, 384, 987

De Lucia, G., Kauffmann, G., Springel, V., et al. 2004, MNRAS, 348, 333

Despali, G. & Vegetti, S. 2017, MNRAS, 469, 1997

Diemand, J. & Moore, B. 2011, Advanced Science Letters, 4, 297

Djorgovski, S. & Davis, M. 1987, ApJ, 313, 59

Dressler, A. 1980, ApJ, 236, 351

Dressler, A. 1984, Annual Review of Astronomy and Astrophysics, 22, 185

Dressler, A., Lynden-Bell, D., Burstein, D., et al. 1987, ApJ, 313, 42

Eichner, T., Seitz, S., Suyu, S. H., et al. 2013, ApJ, 774, 124

Elbaz, D., Arnaud, M., & Boehringer, H. 1995, A&A, 293, 337

Elíasdóttir, Á., Limousin, M., Richard, J., et al. 2007, ArXiv e-prints [[arXiv]0710.5636]

Ettori, S., Donnarumma, A., Pointecouteau, E., et al. 2013, Space Sci. Rev., 177, 119

Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019, ApJ, 875, L1

Faber, S. M., Dressler, A., Davies, R. L., et al. 1987, in Nearly Normal Galaxies. From the Planck Time to the Present, ed. S. M. Faber, 175–183

Faber, S. M. & Jackson, R. E. 1976, ApJ, 204, 668

Fabjan, D., Borgani, S., Tornatore, L., et al. 2010, MNRAS, 401, 1670

Falco, E. E., Gorenstein, M. V., & Shapiro, I. I. 1985, ApJ, 289, L1

Fernández Lorenzo, M., Cepa, J., Bongiovanni, A., et al. 2011, A&A, 526, A72

Focardi, P. & Malavasi, N. 2012, ApJ, 756, 117

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, Publications of the Astronomical Society of the Pacific, 125, 306

Gavazzi, R., Treu, T., Rhodes, J. D., et al. 2007, ApJ, 667, 176

Gehrels, N. 1986, ApJ, 303, 336

Genel, S., Vogelsberger, M., Springel, V., et al. 2014, MNRAS, 445, 175

Giocoli, C., Meneghetti, M., Bartelmann, M., Moscardini, L., & Boldrin, M. 2012a, MNRAS, 421, 3343

Giocoli, C., Meneghetti, M., Ettori, S., & Moscardini, L. 2012b, MNRAS, 426, 1558

Giocoli, C., Tormen, G., Sheth, R. K., & van den Bosch, F. C. 2010a, MNRAS, 404, 502

Giocoli, C., Tormen, G., Sheth, R. K., & van den Bosch, F. C. 2010b, MNRAS, 404, 502

Girardi, M., Mercurio, A., Balestra, I., et al. 2015, A&A, 579, A4

Gobat, R., Rosati, P., Strazzullo, V., et al. 2008, A&A, 488, 853

Gómez, P. L., Valkonen, L. E., Romer, A. K., et al. 2012, AJ, 144, 79

Goodman, J. & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, Vol. 5, No. 1, p. 65-80, 2010, 5, 65

Grillo, C., Gobat, R., Presotto, V., et al. 2014, ApJ, 786, 11

Grillo, C., Lombardi, M., & Bertin, G. 2008, A&A, 477, 397

Grillo, C., Rosati, P., Suyu, S. H., et al. 2018, ApJ, 860, 94

Grillo, C., Suyu, S. H., Rosati, P., et al. 2015, ApJ, 800, 38

Gruen, D., Brimioulle, F., Seitz, S., et al. 2013, MNRAS, 432, 1455

Haardt, F. & Madau, P. 2001, in Clusters of Galaxies and the High Redshift Universe Observed in X-rays, ed. D. M. Neumann & J. T. V. Tran, 64

Halkola, A., Seitz, S., & Pannella, M. 2007, ApJ, 656, 739

Hashimoto, T., Laporte, N., Mawatari, K., et al. 2018, Nature, 557, 392

Herbig, T. & Birkinshaw, M. 1994, in Bulletin of the American Astronomical Society, Vol. 26, American Astronomical Society Meeting Abstracts, 1403

Hernquist, L. 1990, ApJ, 356, 359

Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19

Hoekstra, H., Bartelmann, M., Dahle, H., et al. 2013, Space Sci. Rev., 177, 75

Hogg, D. W., Blanton, M. R., Brinchmann, J., et al. 2004, ApJ, 601, L29

Holden, B. P., van der Wel, A., Franx, M., et al. 2005, ApJ, 620, L83

Hubble, E. P. 1926, ApJ, 64

Hubble, E. P. 1936, Realm of the Nebulae (Dover Publications)

Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236

Jaffe, W. 1983, MNRAS, 202, 995

Jauzac, M., Richard, J., Jullo, E., et al. 2015, MNRAS, 452, 1437

Jiang, L., Wu, J., Bian, F., et al. 2018, Nature Astronomy, 2, 962

Jørgensen, I., Chiboucas, K., Flint, K., et al. 2006, ApJ, 639, L9

Jorgensen, I., Franx, M., & Kjaergaard, P. 1996, MNRAS, 280, 167

Jullo, E. & Kneib, J.-P. 2009, MNRAS, 395, 1319

Jullo, E., Kneib, J.-P., Limousin, M., et al. 2007, New Journal of Physics, 9, 447

Karman, W., Caputi, K. I., Caminha, G. B., et al. 2017, A&A, 599, A28

Karman, W., Caputi, K. I., Grillo, C., et al. 2015, A&A, 574, A11

Kassiola, A. & Kovner, I. 1993, ApJ, 417, 450

Kawamata, R., Oguri, M., Ishigaki, M., Shimasaku, K., & Ouchi, M. 2016, ApJ, 819, 114

Keeton, C. R. & Madau, P. 2001, ApJ, 549, L25

Kelly, P. L., Brammer, G., Selsing, J., et al. 2016, ApJ, 831, 205

Kelly, P. L., Rodney, S. A., Treu, T., et al. 2015, Science, 347, 1123

Killedar, M., Borgani, S., Meneghetti, M., et al. 2012, MNRAS, 427, 533

Kneib, J.-P., Ellis, R. S., Smail, I., Couch, W. J., & Sharples, R. M. 1996, ApJ, 471, 643

Kochanek, C. S. 1993, ApJ, 419, 12

Kochanek, C. S. 1994, ApJ, 436, 56

Koekemoer, A. M., Avila, R. J., Hammer, D., et al. 2014, in American Astronomical Society Meeting Abstracts, Vol. 223, American Astronomical Society Meeting Abstracts #223, 254.02

Koopmans, L. V. E., Bolton, A., Treu, T., et al. 2009, ApJ, 703, L51

Kormendy, J. & Bender, R. 2013, ApJ, 769, L5

Kroupa, P. 2001, MNRAS, 322, 231

Lagattuta, D. J., Richard, J., Bauer, F. E., et al. 2019, MNRAS, 485, 3738

Lagattuta, D. J., Richard, J., Clément, B., et al. 2017, MNRAS, 469, 3946

Lewis, A. & Bridle, S. 2002, Phys. Rev. D, 66, 103511

Limousin, M., Kneib, J. P., Bardeau, S., et al. 2007, A&A, 461, 881

Limousin, M., Kneib, J.-P., & Natarajan, P. 2005, MNRAS, 356, 309

Limousin, M., Richard, J., Jullo, E., et al. 2016, A&A, 588, A99

Limousin, M., Sommer-Larsen, J., Natarajan, P., & Milvang-Jensen, B. 2009, ApJ, 696, 1771

Longair, M. S. 2008, Galaxy Formation (Springer Nature)

Lotz, J., Mountain, M., Grogin, N. A., et al. 2014, in American Astronomical Society Meeting Abstracts, Vol. 223, American Astronomical Society Meeting Abstracts #223, 254.01

Lotz, J. M., Koekemoer, A., Coe, D., et al. 2017, ApJ, 837, 97

Mancini, C., Daddi, E., Juneau, S., et al. 2019, MNRAS, 2068

Markevitch, M., Mushotzky, R., Inoue, H., et al. 1996, ApJ, 456, 437

Maurogordato, S., Cappi, A., Ferrari, C., et al. 2008, A&A, 481, 593

Mead, J. M. G., King, L. J., Sijacki, D., et al. 2010, MNRAS, 406, 434

Melchior, P., Suchyta, E., Huff, E., et al. 2015, MNRAS, 449, 2219

Meneghetti, M. 2019, Introduction to Gravitational Lensing (www.ita.uni-heidelberg.de)

Meneghetti, M., Natarajan, P., Coe, D., et al. 2017, MNRAS, 472, 3177

Meneghetti, M., Rasia, E., Vega, J., et al. 2014, ApJ, 797, 34

Merten, J., Meneghetti, M., Postman, M., et al. 2015, ApJ, 806, 4

Metcalf, R. B. & Madau, P. 2001, ApJ, 563, 9

Metcalf, R. B. & Petkova, M. 2014, MNRAS, 445, 1942

Meylan, G., Jetzer, P., North, P., et al., eds. 2006, Gravitational Lensing: Strong, Weak and Micro

Mo, H., van den Bosch, F. C., & White, S. 2010, Galaxy Formation and Evolution (Cambridge University Press)

Monna, A., Seitz, S., Balestra, I., et al. 2017, MNRAS, 466, 4094

Monna, A., Seitz, S., Greisel, N., et al. 2014, MNRAS, 438, 1417

Monna, A., Seitz, S., Zitrin, A., et al. 2015, MNRAS, 447, 1224

Munari, E., Grillo, C., De Lucia, G., et al. 2016, ApJ, 827, L5

Narayan, R. & Bartelmann, M. 1996, arXiv e-prints, astro

Natarajan, P., Chadayammuri, U., Jauzac, M., et al. 2017, MNRAS, 468, 1962

Natarajan, P. & Kneib, J.-P. 1997, MNRAS, 287, 833

Natarajan, P., Kneib, J.-P., & Smail, I. 2002, ApJ, 580, L11

Natarajan, P., Kneib, J.-P., Smail, I., et al. 2009, ApJ, 693, 970

Newman, A. B., Treu, T., Ellis, R. S., & Sand , D. J. 2011, ApJ, 728, L39

Newman, A. B., Treu, T., Ellis, R. S., et al. 2009, ApJ, 706, 1078

Niemiec, A., Jullo, E., Giocoli, C., Limousin, M., & Jauzac, M. 2019, MNRAS, 1263

Nipoti, C., Giocoli, C., & Despali, G. 2018, MNRAS, 476, 705

Oguri, M. 2019, glafic: gravitational lens adaptive-mesh fitting code (www.slac.stanford.edu)

Okabe, N., Bourdin, H., Mazzotta, P., & Maurogordato, S. 2011, ApJ, 741, 116

Peebles, P. J. E. 1982, ApJ, 263, L1

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2010, AJ, 139, 2097

Petkova, M., Metcalf, R. B., & Giocoli, C. 2014, MNRAS, 445, 1954

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A27

Planelles, S., Borgani, S., Fabjan, D., et al. 2014, MNRAS, 438, 195

Postman, M., Coe, D., Benítez, N., et al. 2012, ApJS, 199, 25

Puchwein, E., Bartelmann, M., Dolag, K., & Meneghetti, M. 2005, A&A, 442, 405

Rafelski, M., Teplitz, H. I., Gardner, J. P., et al. 2015, AJ, 150, 31

Rasia, E., Borgani, S., Murante, G., et al. 2015, ApJ, 813, L17

Renzini, A. 2006, ARA&A, 44, 141

Richard, J., Jauzac, M., Limousin, M., et al. 2014, MNRAS, 444, 268

Rosati, P. 2018, Nature Astronomy, 2, 944

Rosati, P., Balestra, I., Grillo, C., et al. 2014, The Messenger, 158, 48

Salpeter, E. E. 1955, ApJ, 121, 161

Sánchez Almeida, J., Terlevich, R., Terlevich, E., Cid Fernandes, R., & Morales-Luis, A. B. 2012, ApJ, 756, 163

Sánchez-Blázquez, P., Peletier, R. F., Jiménez-Vicente, J., et al. 2006, MNRAS, 371, 703

Schmidt, K. B., Treu, T., Brammer, G. B., et al. 2014, ApJ, 782, L36

Schneider, P. 2006, Extragalactic Astronomy and Cosmology (Springer Nature)

Schneider, P., King, L., & Erben, T. 2000, A&A, 353, 41

Sheth, R. K. & Tormen, G. 2002, MNRAS, 329, 61

Sijacki, D., Springel, V., Di Matteo, T., & Hernquist, L. 2007, MNRAS, 380, 877

Sijacki, D., Vogelsberger, M., Genel, S., et al. 2015, MNRAS, 452, 575

Soucail, G. 2012, A&A, 540, A61

Springel, V. 2005, MNRAS, 364, 1105

Springel, V. & Hernquist, L. 2003, MNRAS, 339, 289

Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, MNRAS, 328, 726

Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001b, MNRAS, 328, 726

Steinborn, L. K., Dolag, K., Hirschmann, M., Prieto, M. A., & Remus, R.-S. 2015, MNRAS, 448, 1504

Stock, D., Meyer, S., Sarli, E., et al. 2015, A&A, 584, A63

Suyu, S. H., Bonvin, V., Courbin, F., et al. 2017, MNRAS, 468, 2590

Suyu, S. H. & Halkola, A. 2010, A&A, 524, A94

Tormen, G., Bouchet, F. R., & White, S. D. M. 1997, MNRAS, 286, 865

Treu, T. 1998, in ESA Special Publication, Vol. 429, LIA Colloq. 34: The Next Generation Space Telescope: Science Drivers and Technological Challenges, ed. B. Kaldeich-Schürmann, 255

Treu, T., Brammer, G., Diego, J. M., et al. 2016, ApJ, 817, 60

Treu, T., Ellis, R. S., Kneib, J.-P., et al. 2003, ApJ, 591, 53

Treu, T., Koopmans, L. V., Bolton, A. S., Burles, S., & Moustakas, L. A. 2006, ApJ, 640, 662

Treu, T. & Koopmans, L. V. E. 2004, ApJ, 611, 739

Treu, T., Schmidt, K. B., Brammer, G. B., et al. 2015, ApJ, 812, 114

Treu, T., Stiavelli, M., Bertin, G., Casertano, S., & Møller, P. 2001, MNRAS, 326, 237

Umetsu, K. 2010, arXiv e-prints, arXiv:1002.3952

Umetsu, K. 2013, ApJ, 769, 13

Umetsu, K. & Broadhurst, T. 2008, ApJ, 684, 177

Umetsu, K., Broadhurst, T., Zitrin, A., Medezinski, E., & Hsu, L.-Y. 2011, ApJ, 729, 127

Umetsu, K., Medezinski, E., Nonino, M., et al. 2014, ApJ, 795, 163

Umetsu, K., Tada, M., & Futamase, T. 1999, Progress of Theoretical Physics Supplement, 133, 53

Valdes, F., Gupta, R., Rose, J. A., Singh, H. P., & Bell, D. J. 2004, ApJS, 152, 251

van de Voort, F., Schaye, J., Booth, C. M., & Dalla Vecchia, C. 2011, MNRAS, 415, 2782

van der Marel, R. P. 1994, MNRAS, 270, 271

Vanzella, E., Castellano, M., Meneghetti, M., et al. 2017, ApJ, 842, 47

Vogelsberger, M., Genel, S., Springel, V., et al. 2014a, Nature, 509, 177

Vogelsberger, M., Genel, S., Springel, V., et al. 2014b, MNRAS, 444, 1518

Warnick, K., Knebe, A., & Power, C. 2008, MNRAS, 385, 1859

Wiersma, R. P. C., Schaye, J., & Smith, B. D. 2009, MNRAS, 393, 99

Yin Cui, Yongzhou Xiang, Kun Rong, Feris, R., & Cao, L. 2014, in IEEE Winter Conference on Applications of Computer Vision, 213–219

Zheng, W., Postman, M., Zitrin, A., et al. 2012, Nature, 489, 406

Zitrin, A., Rosati, P., Nonino, M., et al. 2012, ApJ, 749, 97

Zwicky, F. 1937, ApJ, 86, 217