Alma Mater Studiorum · Università di Bologna

SCUOLA DI SCIENZE

Dipartimento di Fisica e Astronomia

Corso di Laurea Magistrale in Astrofisica e Cosmologia

An high-precision strong lensing model of the galaxy cluster PSZ1 G311.65-18.48

Tesi di Laurea Magistrale

Presentata da:

Giada Venusta Pignataro

Relatore:

Chiar.mo Prof.

Massimo Meneghetti

Correlatore:

Dott.

Pietro Bergamini

Appello IV

Anno Accademico 2019/2020

Abstract

Strong gravitational lensing is one of the best performing methods to study the mass distribution in the central regions of galaxy clusters. The formation of highly amplified and distorted multiple images of background sources provides the constrains on the lens models. The progress witnessed recently on the lens modeling techniques was made possible by the several observational programs conducted with the Hubble Space Telescope (HST) and complemented by spectroscopic observations of a relatively large number of massive galaxy clusters. The effort made in these campaigns produced high-quality photometric and spectroscopic data that improve our ability in identifying the lensing observables and the development of new methods to combine lensing and galaxy kinematics measurements to constrain the cluster mass distribution on a wide range of scales.

In this work we use these novel techniques to conduct a strong lensing analysis of the cluster PSZ1 G311.65-18.48, with the aim of constructing a parametric lens model of the mass distribution in its core. We use HST observations conducted in several bands, complemented with VLT/MUSE spectroscopic observations. The MUSE observations provide redshift estimates for the lensed sources and help reducing the mis-identification of the multiple images. Spectroscopic data are also used to measure the inner velocity dispersion of a set of cluster member galaxies, information incorporated into the lens model to better parametrize the sub-halo cluster component. The multiple image and cluster member catalogues are the main ingredients required for the construction of a parametric lens model with the software `LENSTOOL`.

Throughout this thesis we describe the process that led to the construction of the cluster final model, which evolved through several intermediate steps where the model complexity increased progressively. The aim was to recover the mass distribution which minimizes the differences between observed and model-predicted multiple image positions. Our resulting reference model is the first lens model for this galaxy cluster ever published and reproduces the observed multiple images with very high accuracy. It can be used for several applications, including the characterization of the source producing the famous Sunburst Arc, or predicting the re-appearance of the possible transient source recently reported by Vanzella et al. (2020 [1]).

Sommario

Il lensing gravitazionale forte è attualmente uno dei migliori metodi utilizzati per studiare la distribuzione di massa nel centro degli ammassi di galassie. Infatti, si può sfruttare il fenomeno di formazione delle immagini multiple di sorgenti di background per porre dei constrains ai modelli di lente. Negli ultimi anni si è stati testimoni di un miglioramento nelle tecniche di modellistica, grazie alle campagne osservative condotte su grandi campioni di ammassi di galassie. Queste osservazioni hanno prodotto infatti dati spettroscopici e fotometrici di alta qualità, che hanno aiutato nell'identificazione delle immagini multiple. Questo ha inoltre permesso lo sviluppo di nuovi metodi per combinare il lensing a misure cinematiche sulle galassie appartenenti all'ammasso, per poterne determinare la distribuzione di massa su diverse scale.

In questo lavoro utilizziamo queste nuove tecniche per condurre un'analisi di lensing forte per l'ammasso di galassie PSZ1 G311.65-18.48, e costruire un modello di lente parametrico della distribuzione di massa nella zona centrale. I nostri test si basano principalmente su osservazioni in diverse bande del Telescopio Spaziale Hubble, combinati con dati spettroscopici dello strumento MUSE al VLT. Infatti le osservazioni MUSE ci permettono ottenere misure di redshift per le sorgenti lensate e quindi di migliorare l'identificazione delle immagini multiple. I dati spettroscopici sono utilizzati anche per compiere misure di dispersione di velocità interna delle galassie appartenenti all'ammasso: queste informazioni cinematiche vengono incorporate nel modello di lente, per caratterizzare al meglio la componente di sotto-aloni dell'ammasso di galassie. Il catalogo di immagini multiple e di galassie membro ottenuti, sono gli ingredienti di base per la costruzione di un modello parametrico con il software `LENSTOOL`.

Nel corso di questa tesi viene descritta la costruzione del modello finale per l'ammasso, risultato da una serie di step intermedi in cui è stato progressivamente aumentato il livello di complessità. Lo scopo è quello di ottenere una rappresentazione della distribuzione di massa che minimizzi le differenze tra le posizioni delle immagini multiple osservate e predette dal modello. Il modello di riferimento risultante, è il primo modello di lente pubblicato per questo ammasso di galassie, e riproduce le immagini osservate con alta precisione. Questo potrà essere utilizzato per diverse applicazioni, tra cui la caratterizzazione della sorgente che produce il famoso "Sunburst Arc" e la predizione di future apparizioni del possibile transiente recentemente scoperto da Vanzella et al., (2020 [1]).

Contents

List of Figures

List of Tables

Chapter 1

Cosmology

1.1 Cosmological Principles

In order to build a scientific theory of the Universe, in the beginning it was necessary the construction of simplified models based on some guiding principle. This is the so called Cosmological Principle, that states the homogeneity and isotropy of the Universe on sufficiently large scales (meaning the scales traced by the large-scale structure of the galaxies' distribution): homogeneity is the property of being identical in every point of space, while isotropy means appearing the same in every direction. In other words, in the Universe there are neither privileged positions nor directions if observing a sufficiently large portion of it; in fact the Universe is obviously not homogeneous, but there is quite good observational evidence that the Universe does have these properties, such as the near-isotropy of the Cosmic Microwave Background (CMB) radiation. Before the discovery of the CMB radiation that led to the theory of the expansion of the Universe, the cosmologists supported the Steady State Model, which extended the homogeneity and isotropy of space, to time as well: this was stated in the Perfect Cosmological Principle. The steady-state Universe was abandoned in the 1960s in favor of the Big Bang model.
Another fundamental base of the description of our Universe is Einstein's General Relativity theory: since on large scales the main interacting force is gravity, the physical description of the Universe must be based on how its geometry is shaped by gravity.

1.1.1 Geometry of space

The first step towards the development of a cosmological model is to establish a metric for our Universe, that allows to measure the interval between two space-time events:

$$ds^2 = c^2dt^2 - g_{ij}dx^i dy^j = c^2dt^2 - dl^2 \tag{1.1}$$

(with i,j=1,2,3) where the first term represents the time part, and the second term represent the space part, and g_{ij} is the metric tensor describing the space-time geometry. This interval ds is invariant under a change of coordinate system and the path of a light ray is given by $ds^2 = 0$. Focusing on the purely spatial term, in a homogeneous and isotropic Universe, this can assume three forms depending on geometry: flat Euclidean space, positively curved (sferic), and negatively curved (iperbolic). These can all be expressed in a general form using K, which is the Curvature Parameter and defines the constant curvature of the Universe:

- $K = 0$ a flat, infinite space;
- $K = +1$ a closed space, geometrically a 3D sphere;
- $K = -1$ an open space, geometrically a 3D hyperboloid.

With this K parameter and under the Cosmological Principle assumptions, the metric tensor assumes a simple form called the Friedmann-Robertson-Walker (FRW) metric with line element:

$$ds^2 = c^2dt^2 - a^2(t)\left[\frac{dr^2}{1-Kr^2} + r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right] \tag{1.2}$$

where $a(t)$ is the scale factor of the Universe (or expansion parameter), a function of time which has the dimensions of a length.

1.1.2 Redshift

It is useful to introduce a more directly observable parameter which is related to the scale factor $a(t)$: the redshift z. Consider luminous emitting source, since the source is moving with the expansion of the Universe, its spectrum experiences a cosmological redshift quantified by

$$z = \frac{\lambda_o - \lambda_e}{\lambda_e} \tag{1.3}$$

where λ_e is the wavelength of radiation emitted by the source at the time t_e, while λ_o is the wavelength of radiation observed at the origin of our coordinate system at a subsequent time t_o. In Weinberg (1972 [2]) it is demonstrated that the redshift is related to the scale factor as

$$1 + z = \frac{a(t_o)}{a(t_e)} \tag{1.4}$$

where $a(t_o)$ is the scale factor at the moment of the photon detection, and $a(t_e)$ is the scale factor at the moment of emission. So the redshift z is positive, meaning a shift of the spectrum to longer wavelengths, only if the Universe is expanding.

1.2 Distances and the Hubble Law

We define the point P_0 as the origin of a set of polar coordinates (r, θ, ϕ). We can assume $dt = d\phi = d\theta = 0$ and integrate the FRW metric in this coordinate system to determine the distance measured by a "chain" of observers in every point between P_0 and a generic point P at time t. This is defined as the

Proper distance

$$d_P = \int_0^r \frac{a(t)dr'}{\sqrt{1 - Kr^2}} = a(t)F(r) \tag{1.5}$$

with

$$F(r) = \begin{cases} \arcsin(r) & K = 1 \\ r & K = 0 \\ \text{arcsinh}(r) & K = -1. \end{cases} \tag{1.6}$$

The proper distance calculated ad the present time t_0 is called

Comoving distance

$$d_C := d_{\mathrm{P}}\,(t_0) = a_0 F(r) \tag{1.7}$$

where $a_0 = a(t_0)$. Since the proper distance d_P of a source can change in time as a consequence of the time dependence of the scale factor, the source at P has a radial velocity with respect to P_0:

$$v_{\mathrm{r}} = \dot{a}F(r) = \frac{\dot{a}}{a} d_{\mathrm{P}} = H(t) d_P. \tag{1.8}$$

This equation is the **Hubble law**, and the quantity $H(t)$ is improperly called the Hubble constant; in fact, $H(t)$ (Hubble parameter) it is not constant in time but at a fixed time assumes the same value in every point of the Universe.

From the Hubble law and the definition of the redshift z, in the limit where z is small, it can be shown that

$$z = H_0 d_P. \tag{1.9}$$

It is also useful to express the cosmic scale factor for times t close to t_0 expanding in a power series:

$$a(t) = a_0 \left[1 + H_0 (t - t_0) - \frac{1}{2} q_0 H_0^2 (t - t_0)^2 + \cdots\right] \tag{1.10}$$

where

$$q_0 = -\frac{\ddot{a}(t_0)\, a_0}{\dot{a}(t_0)^2} \tag{1.11}$$

is called the deceleration parameter, which is dimensionless.

We have seen that there is not a unique way to define a distance of an astronomical object in cosmology, and in addition to the proper and comoving distance, which are not directly measurable, we can define other kinds of distances that are, in principle, measurable:

the **Luminosity distance** of a source at a distance r at the time t with emitted power L and flux f, is defined as :

$$d_L \equiv \left(\frac{L}{4\pi f}\right)^{1/2}. \tag{1.12}$$

Due to the expansion of the Universe it is necessary to take into account time-dilation effect, a stretch of the spherical surface centred on the source, and a cosmological redshift on the photons; therefore we obtain

$$d_L = a(t_0) r (1 + z). \tag{1.13}$$

Useful to gravitational lensing purposes,

the **Angular Diameter distance** of a source in r at t with proper diameter D_P and $\Delta\theta$ the angle subtended by it, is defined as:

$$d_A := \frac{D_P}{\Delta\theta}. \tag{1.14}$$

So we can define the duality relation that follows:

$$d_A = d_L \frac{a^2(t)}{a^2(t_0)} = \frac{d_L}{(1 + z)^2}. \tag{1.15}$$

1.3 Friedmann Models

The fundamental basis to every cosmological model stands in the system of Einstein's Equations, which sum up how the space-time geometry, described by the metric tensor g_{ij}, is determined by its energy content (or equivalently, matter) described by the energy-momentym tensor T_{ij}:

$$R_{ij} - \frac{1}{2} g_{ij} R = \frac{8\pi G}{c^4} T_{ij} + \Lambda g_{ij} \tag{1.16}$$

where Λ is the cosmological constant, R_{ij} and R are the Ricci tensor and Ricci scalar, respectively. Einstein inserted the cosmological constant in his equations in order to find static solutions, since they believed the Universe was static. Today we still need the presence of a cosmological constant, which is used to formalize the vacuum energy. In the Big Bang Model assuming a FRW metric, the form of the energy-momentum tensor is that of a perfect fluid. We can define

$$\tilde{T}_{ij} := T_{ij} + \frac{\Lambda c^4}{8\pi G} g_{ij} = -\tilde{p} g_{ij} + \left(\tilde{p} + \tilde{\rho} c^2\right) u_i u_j \tag{1.17}$$

where u_i is the velocity four-vector, $\tilde{p}$ and $\tilde{\rho}$ are the effective pressure and density respectively, related to p and ρ by

$$\tilde{p} = p - \frac{\Lambda c^4}{8\pi G}, \quad \tilde{\rho} = \rho + \frac{\Lambda c^2}{8\pi G}\,. \tag{1.18}$$

If the metric is FRW, the time-time and space-space components of (1.16) yield:

$$\ddot{a} = -\frac{4\pi}{3} G \left(\tilde{\rho} + \frac{3\tilde{p}}{c^2}\right) a\,, \tag{1.19}$$

$$\dot{a}^2 + K c^2 = \frac{8\pi}{3} G \tilde{\rho} a^2 \tag{1.20}$$

known as the Friedmann Equations. Those are related by a third equation, obtainable when assuming the adiabatic expansion of the Universe:

$$\mathrm{d}\left(\rho c^2 a^3\right) = -p \mathrm{d} a^3\,. \tag{1.21}$$

1.3.1 Perfect fluid

The perfect fluid form for (1.17) is required by the Cosmological Principle and forms the basis for the Friedmann models. To solve those equations it is necessary to specify an equation of state for the fluid

in the form of $p = p(\rho)$. In general we can write

$$p = w\rho c^2. \tag{1.22}$$

The value of the parameter w varies depending on the component we are considering

- $w = 0$ represents dust, pressureless material, it describes any non-relativistic fluid or gas;
- $w = 1/3$ represents radiation or, in general, a fluid of non-degenerate ultrarelativistic particles;
- $w = -1$ represents the cosmological constant, the vacuum component.

Inserting (1.22) in the condition (1.21) we obtain a relation between ρ , the scale factor, and w:

$$\rho \propto a^{-3(1+w)}\,. \tag{1.23}$$

This results in the following equations for the different components of the Universe:

- $\rho_M = \rho_{0,M}\left(\frac{a_0}{a}\right)^3 = \rho_{0,M}(1+z)^3$ in a dust dominated Universe;
- $\rho_R = \rho_{0,R}\left(\frac{a_0}{a}\right)^4 = \rho_{0,R}(1+z)^4$ in a radiation dominated Universe;
- $\rho_\Lambda = \rho_{0,\Lambda}\left(\frac{a_0}{a}\right)^0 = \rho_{0,\Lambda}$ in a cosmological constant dominated Universe.

The values of $\rho_{0,w}$ are estimated today and as a consequence of the different evolution in time it is possible to divide the evolution of the Universe in different epochs, in which we can consider only one component to be dominant, and so disregard any contribution from the others. Following this argument the duration of every epoch can be found by imposing the equivalence of different components energy density:

- matter-radiation equivalence, $\rho_M = \rho_R$:
  $(1+z_{eq}) = \frac{\rho_{M,0}}{\rho_{R,0}} \sim 3\cdot 10^4 \Longrightarrow z_{eq}^{MR} \sim 3\cdot 10^4$
- matter-Λ equivalence, $\rho_M = \rho_\Lambda$:
  $(1+z_{eq}) = \left(\frac{\rho_{\Lambda,0}}{\rho_{M,0}}\right)^{1/3} \sim 1.7 \Longrightarrow z_{eq}^{M\Lambda} \sim 0.7$

1.3.2 The Big Bang singularity

Friedmann Equations solutions give the time evolution of $a(t), p(t)$ and $\rho(t)$, when the equation of state is known. They can be reformulated in terms of the Hubble parameter calculated today, H_0, so from

(1.20):

$$H_0^2\left(1-\frac{\Lambda c^2}{3H_0^2}-\frac{\rho_0}{\rho_{crit,0}}\right)=-\frac{Kc^2}{a_0^2}\,. \tag{1.24}$$

Where we defined the Universe *Critical Density* at t_0 as:

$$\rho_{\mathrm{crit},0}:=\frac{3H_0^2}{8\pi G} \tag{1.25}$$

and we can define $\rho_{0,\Lambda}:=\frac{\Lambda c^2}{8\pi G}$ so that the equation becomes:

$$H_0^2\left(1-\Omega_{0,\Lambda}-\Omega_0\right)=-\frac{Kc^2}{a_0^2} \tag{1.26}$$

where we defined the *Density Parameter* $\Omega=\Omega(t)$. For the w component $\Omega_w=\frac{\rho_w}{\rho_{crit}}$, which link the content of energy of the Universe to its curvature (or geometry) as:

- K=0 $\Longrightarrow \Omega_{\Lambda,0}+\Omega_0=\Omega_{tot,0}=1$ flat geometry;
- K=+1 $\Longrightarrow \Omega_{\Lambda,0}+\Omega_0=\Omega_{tot,0}>1$ spherical geometry;
- K=-1 $\Longrightarrow \Omega_{\Lambda,0}+\Omega_0=\Omega_{tot,0}<1$ hyperbolic geometry.

Combining (1.23) and (1.26) we obtain an expression for *H(t)* that reads:

$$H^2(t)=H_0^2\left(\frac{a_0}{a}\right)^2\left[1-\sum_w\Omega_{0,w}+\sum_w\Omega_{0,w}\left(\frac{a_0}{a}\right)^{1+3w}\right] \tag{1.27}$$

It can be shown that if $-1/3<w<1$ the Universe history has a point in time where *a(t)=0* and the density diverges, known as the *Big Bang singularity*. We can write the first Friedmann Equation as

$$\ddot{a}=-\frac{4\pi}{3}G\left(\rho+\frac{3p}{c^2}\right)a=-\frac{4\pi}{3}G\rho(1+3w)a \tag{1.28}$$

where we substituted Eq.(1.22). We can see that $\ddot{a}<0$, provided $(1+3w)>0$, since $\rho>0$. This, along with the observational evidence of the expansion of the Universe ($\dot{a}>0$), establishes that the graph of *a(t)* has negative concavity. It exists a moment in time (see Fig.(1.1)) where *a(t)=0* and we can label this moment $t=0$, since the density diverges and it is a singularity: the Big Bang. This also implies that the time between the singularity and the time t, must be always smaller than the characteristic expansion time, also known as the Hubble time, $\tau_H=1/H$.

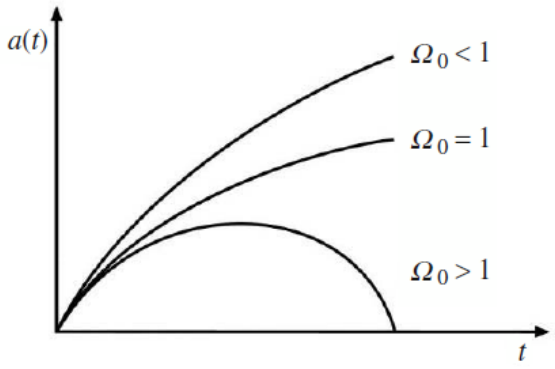

Fig. 1.1. Evolution of the scale factor in different geometry universes. Note the negative concavity of *a(t)* that shows the intersection with the time axis, instant called the Big Bang. Image from Coles, Lucchin (2002 [3]).

1.3.3 Einstein-de Sitter Universe

The Einstein-de Sitter (EdS) Universe is a flat, one component model, so that $\Omega_{0,w} = 1$. With such model one can obtain explicit solutions for all the parameters, written as follow:

$$a(t) = a_0 \left(\frac{t}{t_0}\right)^{\frac{2}{3(1+w)}} , \tag{1.29a}$$

$$t = t_0(1+z)^{\frac{2}{3(1+w)}} , \tag{1.29b}$$

$$H(t) = \frac{2}{3(1+w)}\frac{1}{t} , \tag{1.29c}$$

$$q = \frac{1+3w}{2} , \tag{1.29d}$$

$$\rho = \frac{1}{6\pi G(1+w)^2 t^2} . \tag{1.29e}$$

These equations imply that an EdS Universe is characterized by a Hubble parameter which decreases as a function of time and a constant deceleration parameter, so in this Universe the expansion decelerates constantly with time. The expansion parameter grows indefinitely with time, so the speed of expansion is determined by the value of w. Until now the value of the density parameter Ω was considered at $t = t_0$, but Ω is a function of time as well so it is useful to have an expression to parametrize its evolution with

redshift; from its definition and (1.27) we can rewrite:

$$\Omega_w^{-1}(z) - 1 = \frac{\Omega_{0,w}^{-1} - 1}{(1+z)^{1+3w}} \tag{1.30}$$

For $z \to \infty$, Ω always tends to 1 and we can notice that if $\Omega_{0,tot}$ is greater or smaller than one, then it is greater or smaller than one at every redshift, so even if approaching the Big Bang it tends towards unity, it never crosses the limit. On the other hand if $\Omega_{0,tot} = 1$ then it stays always one: this means that the evolution of the Universe can not change its curvature.

1.3.4 Latest cosmological parameters results

The latest results from the Planck mission (Planck Collaboration et al., 2018 [4]), based on the measurements of the CMB fluctuations, are consistent with a flat Universe with the following cosmological parameters:

$$H_0 = (67.4 \pm 0.5)\ kms^{-1}Mpc^{-1}\,, \tag{1.31a}$$

$$\Omega_{tot} = \sum \Omega_i + \Omega_\Lambda = 1.011 \pm 0.006\,, \tag{1.31b}$$

$$\Omega_m = 0.315 \pm 0.007\,. \tag{1.31c}$$

Results of (1.31b) type justify the restriction to flat cosmologies in our models. Other methods have been used to measure H_0: from the calibration of the tip of the red-giant branch for Type Ia supernovae is found $H_0 = 69.8 \pm 0.8$ km s^{-1}Mpc^{-1} (Freedman et al., 2019 [5]), while the H0LiCOW collaboration based on the method of time delay on lensed quasars found $H_0 = 73.3^{+1.7}_{-1.8}$ km s^{-1}Mpc^{-1} (Wong et al., 2019 [6]). The lower value found by Planck is to impute to the strong degeneracy of H_0 with other parameters such as Ω_m. However this result is in good agreement with the measurements that exploit the baryon acoustic oscillations (BAOs) to calibrate the intrinsic magnitude of the SNe Ia, finding $H_0 = 67.8 \pm 1.3$ km s^{-1}Mpc^{-1} (Macaulay et al., 2019[7]). Anyway we can use the Planck result to calculate the critical density today, finding $\rho_{crit,0} \simeq 2 \times 10^{-29}$ g cm$^{-3}h^2$, where $h = 0.677 \pm 0.004$. In this work we adopt a ΛCDM cosmology with $\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$ and $H_0 = 70$ kms^{-1}Mpc^{-1}.

1.4 Structure formation

The currently accepted Big Bang model presents different problems, most of which can be explained by introducing an epoch of accelerated expansion in the early history of the Universe: the *inflation phase*. The inflation theory is in agreement with homogeneous, isotropic, flat, and magnetic monopole-free Universe models, and it is fundamental to introduce inhomogeneities in our theory. The idea is that quantum fluctuations on microscopic scales in density and temperature at the end of the inflationary phase can grow to create the large-scale structures we observe today.

1.4.1 Jeans instability

The origin of galaxies and galaxies clusters can be explained by Jeans gravitational instability theory, that showes how small density fluctuations in the primordial fluid can grow in time depending on the balance between the structure self-gravity and pressure, leading eventually to collapse into a gravitationally bound object. These density perturbations are defined as

$$\delta := \frac{\delta\rho}{\rho} = \frac{\rho - \bar{\rho}}{\bar{\rho}} \tag{1.32}$$

where $-1 < \delta < +\infty$ and it is dimensionless, and $\bar{\rho}$ is the mean background density. Density fluctuations are related to temperature fluctuations observable in the CMB spectrum and they have the same order of magnitude, which is measured to be $\frac{\delta T}{T} \sim 10^{-5}$ (e.g. Smooth et al., 1992 [8]). Today the density fluctuations amplitude is $\simeq 10^2$ so we know that they have grown between $z_{ls} \simeq 10^3$ (last-scattering redshift, the moment CMB radiation formed) and $z = 0$. In order to develop this theory it is necessary to introduce a length scale called the Cosmological Horizon radius: this is the radius of the region in casual connection with the observer, meaning that inside this radius it is important to consider every force counteracting gravity, while over this radius the only meaningful interaction is due to gravitational force. Its definition is

$$R_H(t) := a(t) \int_0^t \frac{c\mathrm{d}t'}{a\left(t'\right)}, \tag{1.33}$$

it is possible to demonstrate that R_H is a finite quantity that grows as the Universe expands.
Jeans theory (Jeans and Howard, 1902 [9]) starts by considering a static Universe and the idea is to study the evolution of a perturbation applied to a known solution of our problem. Working in a Newtonian approximation we need to look at the dynamics of the fluid that describes the Universe and the density fluctuations. Consider a perfect fluid characterized by the density ρ, the velocity v, the gravitational

potential ϕ, the pressure p, and the entropy S. The equations of hydrodynamics are:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \quad \text{Continuity equation,} \tag{1.34a}$$

$$\frac{\partial v}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\frac{1}{\rho}\nabla\rho - \nabla\phi \quad \text{Euler equation,} \tag{1.34b}$$

$$\nabla^2\phi = 4\pi G\rho \quad \text{Poisson equation,} \tag{1.34c}$$

$$\frac{\mathrm{d}S}{\mathrm{d}t} = 0 \quad \text{Entropy conservation.} \tag{1.34d}$$

One solution of this system is the "background" one, so that $\rho = p = \phi$ =const. and $\mathbf{v}$=0. Now we can apply to this solution a small perturbation, so that $\delta << 1$, and we can use the linear theory. The perturbed system of equations, considering only first-order terms, is then

$$\begin{cases} \frac{\partial \delta\rho}{\partial t} + \rho_b \nabla \cdot \delta\mathbf{v} = 0 \\ \frac{\partial \delta\mathbf{v}}{\partial t} = -\frac{c_0^2}{\rho_b}\nabla\delta\rho + \nabla\delta\phi \\ \nabla^2\delta\phi = 4\pi G\delta\rho\,. \end{cases} \tag{1.35}$$

Jeans' approach was to consider plane-wave solutions in the form of $f(\mathbf{r}, t) = f_k \exp(i\mathbf{k} \cdot \mathbf{r} + i\omega t)$ where $\mathbf{r}$ is the space vector, t the time coordinate, f_k the amplitude, ω is the frequency and $\mathbf{k} = 2\pi\hat{k}/\lambda$ is the wave vector. Working on Fourier space we can obtain the dispersion relation:

$$\omega^2 = k^2 c_s^2 - 4\pi G\rho_b \tag{1.36}$$

which has two different solution depending on ω^2 sign, so the critical value between real and imaginary solutions is $\omega^2 = 0$ that defines the Jeans scale as

$$\lambda_J = \frac{2\pi}{k_J} = c_s\sqrt{\frac{\pi}{G\rho_b}} \tag{1.37}$$

also expressed in terms of mass

$$M_J = \frac{4}{3}\pi\rho_b\lambda_J^3\,. \tag{1.38}$$

In order to have an indefinite growth of the perturbation, its scale-length must be larger than Jeans scale, or equivalently, its mass must exceed the Jeans mass.

1.4.2 Expanding Universe

To extend the Jeans criterion to an expanding Universe it is necessary to keep in mind the length-scales given by Jeans scale (λ_J) and the Horizon radius (R_H), and also define the fundamental time-scales given by the equivalence time (t_{eq}), corresponding to matter-radiation equivalence described in section (1.3.1), and the decoupling time, when the baryonic matter and radiation decouple ($z_{dec} \approx 1000$). We can describe our perturbation as a closed, spherical Universe ($\Omega > 1$) in a less dense, flat, EdS background Universe. Given its definition, we can "sync" the Hubble parameter of the two universes, in order to find an expression relating the density fluctuation to the scale factor:

$$\delta(t) \propto a^{-2}\rho_b^{-1}. \tag{1.39}$$

Thus, since the Universe is radiation and matter dominated before and after the equivalence, respectively, we need to distinguish between different growth regimes:
when $\lambda > R_H$ all perturbations follow the evolution of the main component of the Universe

$$t < t_{eq} \rightarrow \delta \propto a^2 \propto t \tag{1.40a}$$

$$t > t_{eq} \rightarrow \delta \propto a \propto t^{2/3}; \tag{1.40b}$$

when $\lambda_J \ll \lambda < R_H$ gravity effects are influenced by the microphysical processes. To study the evolution of perturbations on scales smaller than the Horizon radius in an expanding Universe, we need rewrite the system of the equation of hydrodynamics in terms of a velocity $\mathbf{u} = H\mathbf{r} + \mathbf{v}$. This expression links the velocity in physical coordinates $\mathbf{u}$, to the velocity in comoving coordinates $\mathbf{v}$. Here, the expansion of the Universe is parametrized by the Hubble law, and $\mathbf{v}$ is defined as the peculiar velocity of the perturbation, namely a way to quantify its velocity component within respect to the expansion. To solve the perturbed hydrodynamic equations we still look for a solution in wave form written as $f(\mathbf{x}, t) = f_k(t)\exp(i\mathbf{k}\cdot\mathbf{x})$ where now the amplitude of the wave has a time dependency as well. It can be demonstrated that this leads to a dispersion relation

$$\ddot{\delta}_k + 2\frac{\dot{a}}{a}\dot{\delta}_k + \left(k^2c_s^2 - 4\pi G\rho_b\right)\delta_k = 0 \tag{1.41}$$

which is an ordinary differential equation (ODE) of the second order describing how the density field varies as a function of time in an expanding Universe. Note that the second term in the equation

contains the Hubble parameter. Assuming the simple power law form $\delta_k \propto t^\alpha$ the ODE solutions are

$$t > t_{eq} \rightarrow \begin{cases} \delta_- \propto a^{-3/2} \propto t^{-1} \\ \delta_+ \propto a \propto t^{2/3} \end{cases} . \tag{1.42}$$

Of course, we are interested in the growing solution: in a matter dominated Universe the perturbation growth is linearly proportional to the expansion factor.
In a radiation dominated Universe, namely $t < t_{eq}$, the ODE (1.41) has a slightly different form:

$$\ddot{\delta}_k + 2\frac{\dot{a}}{a}\dot{\delta}_k + \left(k^2 c_s^2 - \frac{32}{3}\pi G \rho_b\right)\delta_k = 0 \tag{1.43}$$

and the solutions are again a growing and a damping one. The problem arises when comparing the Jeans scale and the Horizon radius in this epoch: it can be shown that the Jeans' scale is always larger than the Horizon Radius, meaning that in the radiation era it is not possible to find a solution for photons density perturbations at scales smaller than the Horizon radius; the waves propagate at such high speed that fluctuations are cancelled.

The solutions in Eq.(1.42) hold for every type of matter, but it must be considered that dark and baryonic matter are coupled and interact gravitationally. Dark matter (DM) decouples from radiation way earlier than baryons, and since that moment perturbations evolve as in Eq.(1.42); to find how dark matter perturbations evolve before the equivalence we need to solve a dispersion relation in the form

$$\ddot{\delta}_{k,DM} + 2\frac{\dot{a}}{a}\dot{\delta}_{k,DM} + k^2 c_s^2 \delta_{k,DM} - 4\pi G \rho_{b,DM}\delta_{k,DM} = 0 \tag{1.44}$$

whose growing solution inside the Horizon radius ($\lambda < R_H$) is

$$\delta_{k,DM} = 1 + \frac{3}{2}\frac{a}{a_{eq}} . \tag{1.45}$$

This solution implies that, since R_H grows with time, different scales can enter the horizon at different times, modifying their growth trend: indeed the so called *Mezaros effect* explains that dark matter perturbations can grow up to a factor 5/2 in the radiative era, namely $\frac{\delta(t_{eq})}{\delta(t_H)} \leq \frac{5}{2}$.

Until t_{eq} baryons are coupled with radiation because of frequent Thompson scattering, and their perturbations can not evolve. At t_{dec} baryons perturbation can finally start to grow and we would expect them to follow the solution δ_+ in Eq.(1.42), but their growth is not free since dark matter

perturbations already had time to create potential wells, where the baryons can fall. We find solutions to the dispersion relation of baryonic fluctuations for $\lambda_J \ll \lambda < R_H$ at $t > t_{dec}$, written as

$$\delta_{k,B} = \delta_{k,DM} \left(1 - \frac{a_{dec}}{a}\right) . \tag{1.46}$$

This expression parametrizes the **baryon catch-up** phenomenon: after decoupling baryonic fluctuations evolve faster to "catch-up" with the dark matter perturbations evolution. In other words, the growth of baryonic perturbations is accelerated by the presence of the dark matter potential wells. When $a \gg a_{dec}$ the size of the dark matter and baryon perturbation is equal, and the evolution continues following Eq.(1.42).

1.4.3 Dissipation

As seen the Jeans criterion can be expressed in terms of mass rather than length. It is useful to understand how the dark matter or the baryonic Jeans mass evolves with the cosmic time, to evaluate when perturbations can grow and how massive they must be to not be dissipated. It is important to distinguish between two types of dark matter: cold dark matter (CDM) and hot dark matter (HDM). We talk about CDM if the dark matter particles are not relativistic at the epoch of decoupling from the other components of the Universe; vice versa if the particles are still relativistic at the decoupling we talk about HDM. This is important because we need to determine the behaviour of the mean particle velocity in the different cosmological epochs, in order to use Eq.(1.38). It can be show that DM Jeans mass in both scenarios grows until equivalence and then starts to decline: this means that there is a maximum value for Jeans mass, and this is what sets apart the two scenarios. Simulations showed that the maximum value of the DM Jeans mass, reached at equivalence is

- $M_J(a_{eq}) \approx 10^5 M_\odot$ for CDM: all perturbations with mass greater than this value had the possibility to grow and collapse to create smaller structures first (globular cluster sizes), leading to the large-scale structures observed today by hierarchic aggregation (**bottom-up scenario**);
- $M_J(a_{eq}) \approx 10^{15} M_\odot$ for HDM: only perturbations with such masses can collapse and form structures, meaning that the first structures formed would be greater than today's galaxy clusters, and all other smaller object formed by fragmentation (**top-down scenario**).

Current observations and studies on galaxy clusters and globular clusters ages lead to support CDM scenarios for our Universe.

The baryonic Jeans scale is always as in Eq.(1.37) but in order to evaluate the particles motion we can

now use the sound speed, since it is collisional matter. Also, in this case the Jeans mass grows until reaching a maximum value at the time of decoupling, and then it decreases: since at decoupling baryons loose the radiation support, it can be shown that there is a sharp drop in normalization. The maximum value for baryons Jeans mass is $M_J(a_{dec}) \approx 10^{16} M_\odot$ but since the evolution of baryonic perturbation is led by dark matter potential wells, it is not relevant.

So far we discussed how perturbations can grow. However, there are also damping solutions that we did not consider yet. On scales $\lambda < \lambda_J$, dissipation is important for both dark and baryonic matter:

- since dark matter is collision-less it is not real dissipation. Rather, it is a free streaming effect: at decoupling DM particles propagate freely responding to the mean gravitational field of the Universe, damping the fluctuations. We can define the free-streaming scale as the space travelled by a DM particle in a time t, if the perturbation velocity is v :

$$\lambda_{FS}(t) = a(t) \int_0^t \frac{v\,(t')}{a\,(t')} \mathrm{d}t'. \tag{1.47}$$

  Thus the free-streaming mass is $M_{FS} \propto \lambda_{FS} \rho_{DM}$ and, if a perturbation at the time t has a mass smaller than the free-streaming mass, it will be cancelled. It can be shown that free-streaming mass grows as a function of time until equivalence, then it remains constant. Since the normalization of both the Jeans and the FS mass is the same, at t_{eq} all perturbations with $M < M_{FS}(a_{eq})$ are cancelled, because $M_J(a_{eq}) = M_{FS}(a_{eq})$ but then the Jeans mass starts to decrease;

- since baryonic matter is collisional we can talk about dissipation caused by the continuous stochastic process of particles collisions that dampens the wave. To quantify the total space travelled by the particles we can define the Silk scale and mass, which follow $\lambda_S \propto a^{5/2}$ and $M_S \propto a^{9/2}$ before equivalence, and $\lambda_S \propto a^{9/4}$ and $M_S \propto a^{15/4}$ after equivalence; Silk mass grows until decoupling, where it reaches its maximum $M_S(a_{dec}) \approx 10^{12} M_\odot$.

Since the perturbation amplitude distribution changes depending on the considered scale, not all scales grow at the same rate.

1.4.4 Non-linear theory

In today's Universe we observe matter over-densities $\delta \gg 1$ and this implies that there is a time when the growth of perturbations becomes non-linear. This happens when the perturbations grow enough to reach $\delta \approx 1$. In this regime on a non-linear description is needed. The growth of perturbations

can be derived analytically only in a very simple approximation, whereas in every other case numerical simulations are needed. Analytically the simplest way to study the non-linear evolution of perturbations is considering a spherical and symmetric perturbation, with null initial co-moving velocity. To do so the perturbation is described as a closed Universe in a background EdS Universe, and from that we can find a condition for the positive solution to collapse:

$$\delta_+ > \frac{3}{5}\frac{1-\Omega_i}{\Omega_i(1+z_i)}, \tag{1.48}$$

where $\Omega_i = \Omega(t_i)$ is the density parameter of the background Universe at t_i. There is a threshold to overcome within a time frame for the collapse to happen and this condition is always verified for flat and closed background universe. However for open background universes it can be more difficult.
Dealing with the perturbation as a closed Universe we ensure its collapse at the so-called *turn-around* point, where the scale factor is maximum (and the density is minimum) and then starts to decrease, as seen in Fig.(1.1). We want to know what is the ratio between the perturbation density and the background Universe density at the moment of the turn-around:

$$\chi\left(t_{\max}\right) := \frac{\rho_p\left(t_{\max}\right)}{\rho\left(t_{\max}\right)} \approx 5.6. \tag{1.49}$$

Thus, because of Eq.(1.32), we have that $\delta(t_{max}) = \chi - 1 \approx 4.6$, meaning that the perturbation is already non-linear at the turn-around time. After the turn-around time, the density increases together with temperature, so matter heats up and the increasing pressure counteracts gravity, leading to an equilibrium state. Simulations show that when the collapse stops the perturbation has radius R_e at $t_e \approx 2t_{max}$, and then oscillates around this radius until it settles and virialize at $t_{vir} \approx 3t_{max}$. Between the equilibrium and virial time the perturbation density is constant, while the background Universe is still expanding, meaning that the ratio in density varies between these two instants:

$$\frac{\rho_p\left(t_e\right)}{\rho\left(t_e\right)} \approx 2^2 8\chi \approx 180 \tag{1.50}$$

and

$$\frac{\rho_p\left(t_{vir}\right)}{\rho\left(t_{vir}\right)} \approx 3^2 8\chi \approx 400. \tag{1.51}$$

If we applied linear theory we would have found non-linear results, but 2 order of magnitude smaller,

$$\begin{aligned} \delta_+ (t_e) &\approx 1.686 \\ \delta_+ (t_{vir}) &\approx 2.20 \end{aligned} \tag{1.52}$$

stressing the importance of developing a non-linear theory. The problems in this approximation are that spherical collapse is not realistic and the non-linear theory results are strongly dependent on the cosmological model we choose to adopt.

1.4.5 Press-Schechter function

In order to test our models of cosmic structures evolution a connection with observables is needed. This step is possible defining a mass function $n(M)dM$ which gives the number of objects with mass between M and $M + dM$ at the time t per unit volume. Fixing a mass-to-light ratio (M/L), it is possible to switch between the mass function and the luminosity function, which can be measured more easily. The Press-Schechter mass function (Press & Schechter, 1974 [10]) was found starting from the following assumptions: (i) the density perturbations distribution is Gaussian, as stated by the inflationary theory; (ii) spherical collapse model with threshold value $\delta_c = 1.686$, as found in linear theory in (1.52); (iii) Jeans linear theory for the evolution of perturbations. Press and Schechter (1994), showed that the mass function is

$$n(M)\mathrm{d}M = \sqrt{\frac{2}{\pi}} \frac{\alpha \bar{\rho}_M}{M_*^2} \left(\frac{M}{M_*}\right)^{\alpha-2} \exp\left[-\left(\frac{M}{M_*}\right)^{2\alpha}\right] \mathrm{d}M, \tag{1.53}$$

where $n(M)$ is the number density mass of the halos with mass M, $\bar{\rho}_M$ is the mean density of the halos with mass M, α is the power law index and M_* is a characteristic mass. In this form, Eq.(1.53) is a power law function with an exponential cut off at M_*, which increases with time, meaning that as time passes more massive structure are forming in the Universe (Fig 1.2).

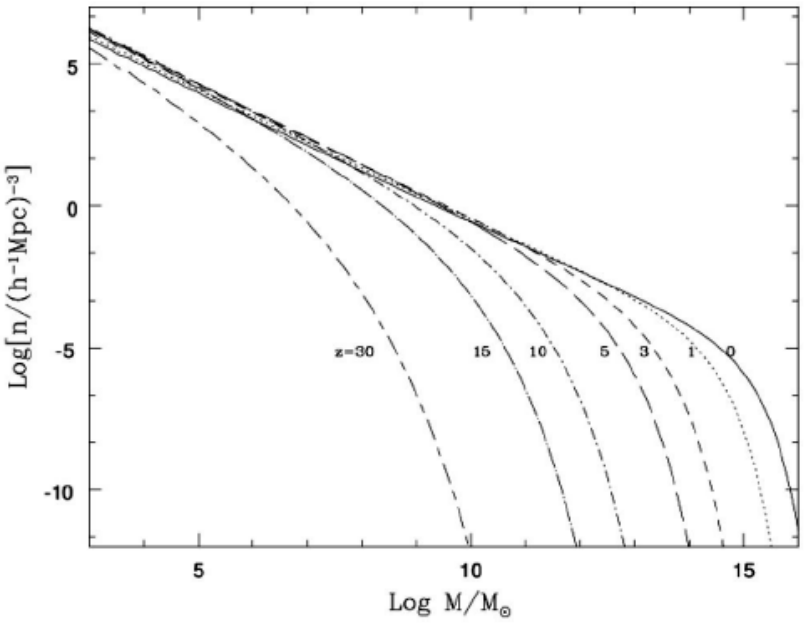

Fig. 1.2. Logarithmic Press-Schechter mass function, at different redshifts. At smaller redshifts the "knee" of the function is shifted towards larger masses.

Chapter 2

Gravitational Lensing

As seen in Chapter (1) Einstein's Theory of Relativity describes how mass, and hence gravity, shapes the geometry of space-time. A gravitational lens can occur when a sufficiently large mass concentration creates a gravitational field strong enough to distort and magnify the light coming from other sources behind it, along the same line of sight. In fact, light propagates on null geodesics as seen in (1.1.1) and these paths are affected by the small- and large-scale matter clumps in our Universe. Gravitational lensing is a powerful tool to study the distribution of matter in the cosmic structures and to observe the distant Universe.

2.1 Weak field approximation

Light deflection can be described by Fermat's principle, namely approaching the problem as light refraction and defining an effective refraction index, n. In most astrophysical scenarios lenses are "weak", meaning that their gravitational potential is small, $\Phi \ll c^2$, and are small within respect to the source-lens-observer system. Neglecting the expansion of the Universe, the line element of the local metric tensor can be written as a small perturbation of the Minkowski metric,

$$\mathrm{d}s^2 = g_{\mu v}\mathrm{d}x^{\mu}\mathrm{d}x^{v} = \left(1 + \frac{2\Phi}{c^2}\right)c^2\mathrm{d}t^2 - \left(1 - \frac{2\Phi}{c^2}\right)(\mathrm{d}\vec{x})^2. \tag{2.1}$$

From the null geodesic form, $ds^2 = 0$, the photon effective speed in the gravitational perturbation is smaller than in its absence:

$$c' = \frac{d\vec{x}}{dt} \approx c\left(1 + \frac{2\phi}{c^2}\right) \tag{2.2}$$

This equation leads to the definition of the effective refraction index,

$$n = \frac{c}{c'} = \frac{1}{1 + 2\phi/c^2} \approx 1 - \frac{2\phi}{c^2} \tag{2.3}$$

needed to apply the Fermat's principle. The total deflection angle of a photon is be the integral over the gradient of the potential perpendicular to the light path, along the proper light path. However we can adopt the *Born approximation* and integrate over the unperturbed light path:

$$\hat{\vec{\alpha}}(b) = \frac{2}{c^2} \int_{-\infty}^{+\infty} \vec{\nabla}_\perp \phi \mathrm{d}z \tag{2.4}$$

where b is the impact parameter of the photon with respect to the mass M. If the lens is a point mass, then the potential is $\Phi = -GM/r$, where G is the gravitational constant and $r = \sqrt{x^2 + y^2 + z^2} = \sqrt{b^2 + z^2}$. Thus the deflection angle becomes

$$|\hat{\vec{\alpha}}| = \frac{4GM}{c^2 b} = 2\frac{R_s}{b} \tag{2.5}$$

where $R_s = 2GM/c^2$ is the Schwarzchild radius of a mass M. Having multiple point masses the total deflection angle can be computed by superposition, given the linearity in M.

In a typical lensing system we can consider the source sphere, with radius D_S centered on the observer O, and the lens sphere, with radius D_L centered on the lens. In all cases of astrophysical interest deflection angles are small, therefore the spheres can be replaced by the corresponding tangent planes, the source plane and the lens plane. As shown in Fig. (2.1), the separation between the ray and the optical axis is described by the two-dimensional vector $\vec{\xi}$ on the lens plane, and by $\vec{\eta}$ on the source plane, and D_S, D_L and D_{LS} are angular diameter distances.

Since most of the deflection occurs within $\Delta z \sim b$ and the distances to the source and the observer are greater than this, we can apply the thin screen approximation and assume that the deflection occurs in the lens plane. Considering now a more realistic model where the distribution of matter is three-dimensional, this approximation allows to describe the matter distribution with its surface density,

$$\Sigma(\vec{\xi}) = \int \rho(\vec{\xi}, z) \mathrm{d}z \tag{2.6}$$

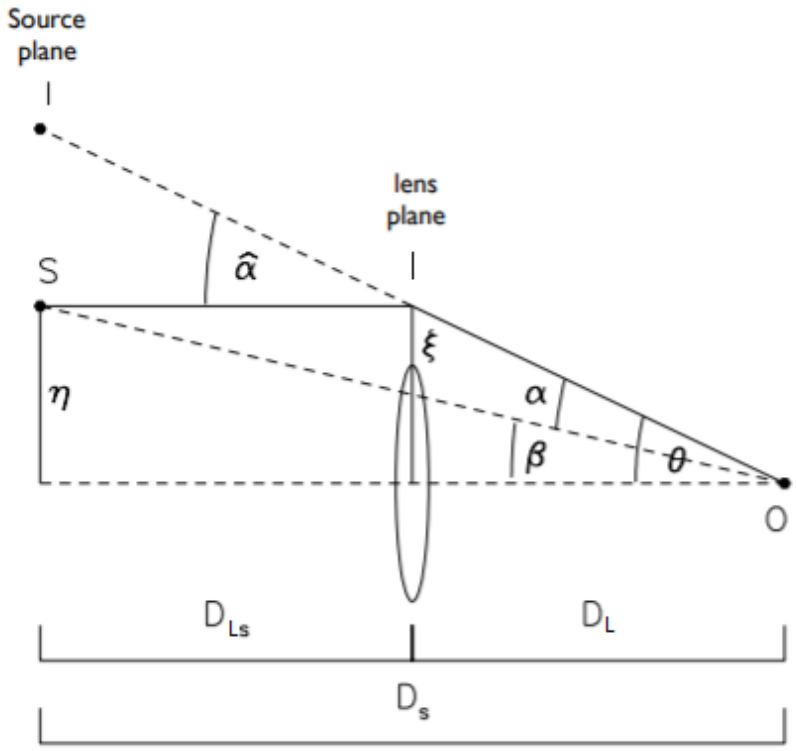

Fig. 2.1. Lensing system geometry. Figure from Narayan and Bartelmann, 1997 [11]

where ρ is the 3D density. Given the linearity in M of (2.5), the deflection angle is

$$\vec{\alpha}(\vec{\xi}) = \frac{4G}{c^2} \int \frac{\left(\vec{\xi} - \vec{\xi}'\right) \Sigma \left(\vec{\xi}'\right)}{\left|\vec{\xi} - \vec{\xi}'\right|^2} \mathrm{d}^2\xi' \tag{2.7}$$

2.2 The lens equation

In order to define the observable light path we need to derive a relation between the intrinsic source position ($\vec{\beta}$) and the apparent one ($\vec{\theta}$), also referred as the image position. From the geometry of Fig(2.1), if $\hat{\vec{\alpha}}, \vec{\theta}, \vec{\beta}$ are small, we can obtain the lens equation

$$\vec{\theta} D_{\mathrm{S}} = \vec{\beta} D_{\mathrm{S}} + \hat{\vec{\alpha}} D_{\mathrm{LS}}, \tag{2.8}$$

and, by defining the reduced deflection angle,

$$\vec{\alpha}(\vec{\theta}) \equiv \frac{D_{LS}}{D_S} \hat{\vec{\alpha}}(\vec{\theta}), \tag{2.9}$$

we obtain $\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta})$. We can now define a length scale ξ_0 on the lens plane, corresponding to a length scale η_0 on the source plane, such that $\vec{x} \equiv \frac{\vec{\xi}}{\xi_0}$ and $\vec{y} \equiv \frac{\vec{\eta}}{\eta_0}$; this allows the derivation of a dimensionless lens equation:

$$\vec{y} = \vec{x} - \vec{\alpha}(\vec{x}) \tag{2.10}$$

where

$$\vec{\alpha}(\vec{x}) = \frac{D_{\mathrm{L}} D_{\mathrm{LS}}}{\xi_0 D_{\mathrm{S}}} \vec{\alpha}\left(\xi_0 \vec{x}\right) \tag{2.11}$$

is the scaled deflection angle.

2.2.1 Lensing potential and convergence

The deflection depends on the projection of the 3D potential on the lens plane and on the lens-source system geometry. Hence with a proper rescaling we can define the *effective lensing potential* that characterizes the mass distribution:

$$\hat{\Psi}(\vec{\theta}) = \frac{D_{\mathrm{LS}}}{D_{\mathrm{L}} D_{\mathrm{S}}} \frac{2}{c^2} \int \Phi\left(D_{\mathrm{L}} \vec{\theta}, z\right) \mathrm{d}z. \tag{2.12}$$

It can be shown that the gradient of the lensing potential is the reduced deflection angle. In dimensionless form, this property can be written as

$$\vec{\nabla}_x \Psi(\vec{x}) = \vec{\alpha}(\vec{x}), \tag{2.13}$$

where $\Psi = \frac{D_L^2}{\xi_0^2} \hat{\Psi}$. In addition, the Laplacian of the lensing potential is twice the convergence κ, i.e.

$$\kappa(\vec{x}) = \frac{1}{2} \nabla_x^2 \Psi(\vec{x}), \tag{2.14}$$

where

$$\kappa(\vec{\theta}) \equiv \frac{\Sigma(\vec{\theta})}{\Sigma_{\mathrm{cr}}} \quad \text{with} \quad \Sigma_{\mathrm{cr}} = \frac{c^2}{4\pi G} \frac{D_{\mathrm{S}}}{D_{\mathrm{L}} D_{\mathrm{LS}}}. \tag{2.15}$$

As shown by Eq.(2.15), the convergence is the ratio between the surface density and the *critical surface density* Σ_{cr}. The latest is a function of the angular diameter distances of the system.

From their definitions, lensing quantities such as the lensing potential (2.12) and the critical surface density (2.15) strongly depends on distances. Fig.(2.2) shows how the factor $\frac{D_{\mathrm{S}}}{D_{\mathrm{L}} D_{\mathrm{LS}}}$ varies with the lens or source redshift. As it can be seen in Fig.(2.2a) to a larger source-lens distance corresponds a larger convergence, it is easier to exceed the Σ_{cr} threshold for a higher source redshifts. When this happens, the lens is called "strong"; for a given source redshift, varying the lens redshift, the maximum lensing

effect is produced by lenses at intermediate distaces (blue curve) between the source and the observer. The peaks of the curves in Fig.(2.2b) shift to higher redshifts if the source distance increases.

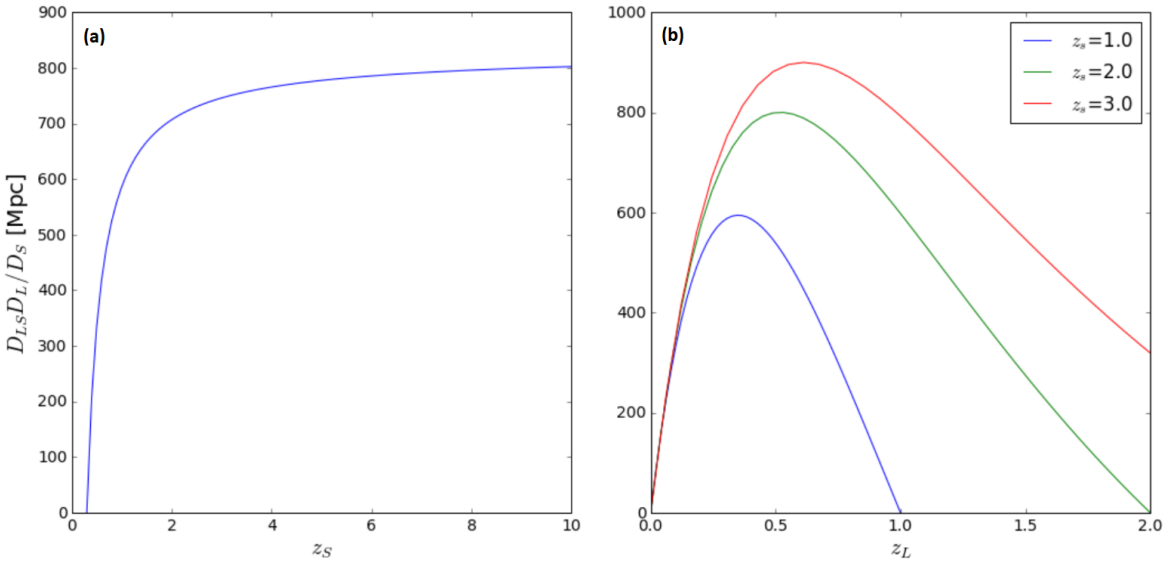

Fig. 2.2. How the distance factor varies with the source redshift (a), and how it varies with the lens redshift when the distance to the source increases (b). Figure from M. Meneghetti lectures (2020 [12]).

2.2.2 First order lens mapping

The lens equation (2.8) ideally allows us to determine the unperturbed source position $\vec{\beta}$ corrisponding to an observed image position $\vec{\theta}$, if the deflection angle $\vec{\alpha}$ given in Eq.(2.4) is known. For an extended source we should solve the lens equation for every point-source element in order to reconstruct its distorted image. We can work using a first order approximation if the source size is significantly smaller than the angular scale on which the deflection angle changes. By computing the distance between two points $\vec{\beta}$ and $\vec{\beta}' = \vec{\beta} + d\vec{\beta}$ on the source plane and mapping it on the lens plane we define a linear mapping between the source and the lens plane, described by the Jacobian matrix

$$A \equiv \frac{\partial \vec{\beta}}{\partial \vec{\theta}} = \left(\delta_{ij} - \frac{\partial \alpha_i(\vec{\theta})}{\partial \theta_j} \right) = \left(\delta_{ij} - \frac{\partial^2 \hat{\Psi}(\vec{\theta})}{\partial \theta_i \partial \theta_j} \right). \tag{2.16}$$

This metric is called the *lensing Jacobian*. It is a rank-two symmetric tensor that can be split into an isotropic and an anisotropic part

$$\begin{aligned} A_{iso,i,j} = \frac{1}{2} tr A \delta_{ij} &= \left[1 - \frac{1}{2} \left(\hat{\Psi}_{11} + \hat{\Psi}_{22} \right) \right] \delta_{ij} \\ &= \left(1 - \frac{1}{2} \nabla^2 \hat{\Psi} \right) \delta_{ij} = (1 - \kappa) \delta_{ij}, \end{aligned} \tag{2.17a}$$

$$\begin{aligned} A_{aniso_{i,j}} = A_{i,j} - \frac{1}{2} tr A \delta_{ij} &= \delta_{ij} - \hat{\Psi}_{ij} - \frac{1}{2}\left(1 - \hat{\Psi}_{11} + 1 - \hat{\Psi}_{22}\right)\delta_{ij} \\ &= -\hat{\Psi}_{ij} + \frac{1}{2}\left(\hat{\Psi}_{11} + \hat{\Psi}_{22}\right)\delta_{ij} \\ &= \begin{pmatrix} -\frac{1}{2}\left(\hat{\Psi}_{11} - \hat{\Psi}_{22}\right) & -\hat{\Psi}_{12} \\ -\hat{\Psi}_{12} & \frac{1}{2}\left(\hat{\Psi}_{11} - \hat{\Psi}_{22}\right). \end{pmatrix} \end{aligned} \tag{2.17b}$$

From (2.17b) we can define the shear tensor Γ, also often written in a form of a pseudo-vector, $\vec{\gamma} = (\gamma_1, \gamma_2)$, whose components are

$$\gamma_1 = \frac{1}{2}\left(\hat{\Psi}_{11} - \hat{\Psi}_{22}\right) \tag{2.18}$$

$$\gamma_2 = \hat{\Psi}_{12} = \hat{\Psi}_{21} \tag{2.19}$$

The shear tensor has eigenvalues $\pm\sqrt{\gamma_1^2 + \gamma_2^2} = \pm\gamma$, so we can write the shear tensor as

$$\begin{pmatrix} \gamma_1 & \gamma_2 \\ \gamma_2 & -\gamma_1 \end{pmatrix} = \gamma \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix} \tag{2.20}$$

where ϕ identifies the direction of the eigenvector corresponding to the positive eigenvalue.

From (2.17a), we use the property (2.14) to make the convergence explicit. Summarizing, we can rewrite the lensing Jacobian matrix as

$$\begin{aligned} A &= \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix} \\ &= (1 - \kappa)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \gamma \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix} \end{aligned} \tag{2.21}$$

from which we can see how the convergence determines an isotropic distortion of the image, so it's expandend/contracted in all directions by the same factor, while the shear determines an anisotropy by stretching the shape along a particular direction. As shown in Fig.(2.3), a circular source of radius r mapped on the lens plane will be an ellipse, with semi-major and -minor axes defined by

$$a = \frac{r}{1 - \kappa - \gamma} = \frac{r}{\lambda_t} \quad , \quad b = \frac{r}{1 - \kappa + \gamma} = \frac{r}{\lambda_r} \tag{2.22}$$

where λ_t and λ_r are the tangential and radial eigenvalues of the Jacobian matrix, respectively. The ellipticity is defined as

$$e = \frac{a - b}{a + b} = \frac{\gamma}{1 - k} \equiv g, \tag{2.23}$$

where we introduced the reduced shear g.

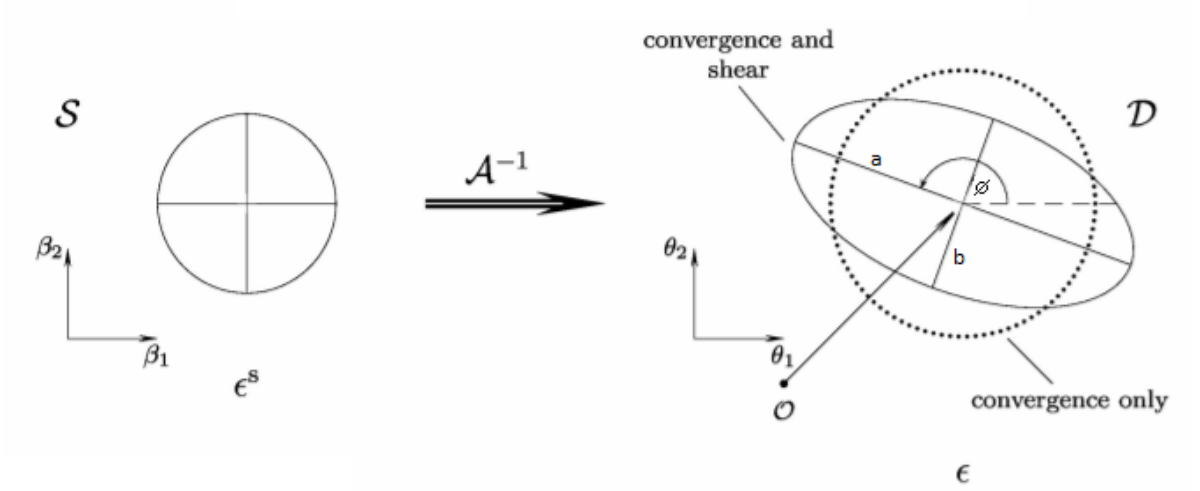

Fig. 2.3. Circular image distortion into an ellipse due to convergence and shear. Figure from S. Suyu lectures (2016 [13]) .

2.2.3 Magnification

Along with distortions, gravitational lensing introduces a magnification effect. If we are observing a lensed source, to compute the magnified flux within respect to the intrinsic one we only need to see how the source area changes, since the surface brightness I_ν is conserved (no photons are created or absorbed). This property is explained by Liouville theorem. Given the ellipse axes in Eq.(2.22), we want to compute how the area of the lensed image (I) compares to the source intrinsic area (S). We obtain:

$$I = \pi ab = \pi r^2 (detA)^{-1} = S(detA)^{-1} \rightarrow \mu \equiv \frac{1}{detA} = \frac{I}{S}, \tag{2.24}$$

where μ is the magnification. The link between the intrinsic and observed flux can be expressed as

$$F_\nu = \int_I I_\nu(\vec{\theta}) d^2\theta = \int_S I_\nu^S(\vec{\beta}(\vec{\theta}))\mu(\vec{\theta}) d^2\beta, \tag{2.25}$$

where the first integral is over the image plane, while the second over the source plane. From Eq.(2.24), we can define the radial ($\mu_r = 1/\lambda_r$) and tangential ($\mu_t = 1/\lambda_t$) magnification factors. Since the magnification is a function of $\vec{\theta}$, there exists a set of critical points where detA $= 0$, namely where $\lambda_r = 0$ and $\lambda_t = 0$. Along the lines defined by those conditions the magnification diverges. These lines are called *critical lines* and they are defined on the lens plane. If we map them onto the source plane using the lens equation we obtain the *caustics*. Since the Jacobian determinant vanishes along

these lines, the regions of the lens plane separated by the critical lines have values of the Jacobian determinant with opposite sign. This indicates that parities of the image formed on the two sides of the critical lines are inverted. An image near a critical line will be characterized by a strong tangential or radial deformation, as well as higher flux. Given a lens system, the number of images varies with the source position. And it can be shown that the number of images changes by two when the source crosses a caustic (P.Schneider, J. Ehlers, E.Falco, 1992 [14]): two images with opposite parity merge on the critical line. Working with massive and complex lenses such as galaxy clusters, critical and caustic lines can have extended and irregular shapes. A source can produce several multiple images. In addition, many sources can be strongly lensed simultaneously as seen in Fig.(2.4). The figure shows the that shows the model describing the mass distribution of the core of the galaxy cluster MACS J1206 (G.B. Caminha et al., 2017 [15]).

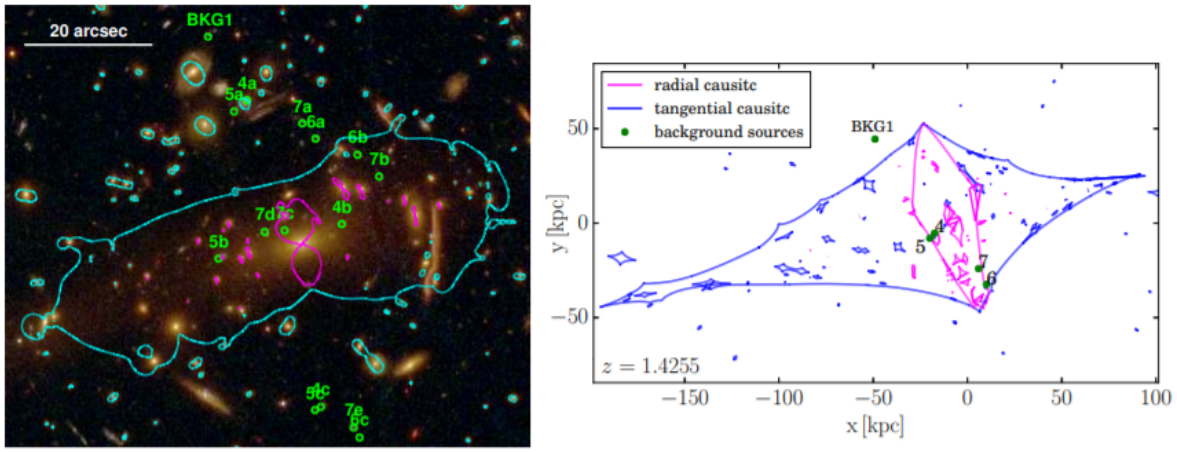

Fig. 2.4. MACS J1206 galaxy cluster lens model. On the left, tangential (cyan) and radial (magenta) critical lines on the image plane; on the right, tangential and radial caustic lines on the source plane, along with the reconstructed position of the background sources. Figure from G.B. Caminha et al. (2017 [15]).

2.2.4 Time-delay

Light passage through a gravitational potential also leads to a delay in the travel time between the source and the observer. Consequently, we see multiple images of the same source appearing on the lens plane at different times. This time-delay has two components. One is geometrical, due to different path taken by deflected light rays to reach the observer, and one is gravitational, and due to the different effective speed of light in presence of a gravitational potential. It can be demonstrated that the total

time-delay is described by a surface:

$$t(\vec{\theta}) = t_{geom} + t_{grav} = = \frac{1}{c} \underbrace{(1+z_L)\frac{D_{\rm L} D_{\rm S}}{D_{\rm LS}}}_{\text{time-delay distance}} \underbrace{\left[\frac{1}{2}(\vec{\theta}-\vec{\beta})^2 - \hat{\Psi}(\vec{\theta})\right]}_{\text{Fermat's potential}} = \frac{D_{\Delta r}}{c}\tau(\vec{\theta}), \tag{2.26}$$

where z_L is the lens redshift and the term $(1+z_L)$ is a factor introduced to account for the expansion of the Universe. The gradient of this surface gives the lens equation,

$$\nabla\left[\frac{1}{2}(\vec{\theta}-\vec{\beta})^2 - \hat{\Psi}(\vec{\theta})\right] = 0 \tag{2.27}$$

Thus the images form at the stationary points of the time-delay surface. The curvature of this surface is given by its Hessian matrix:

$$T = \frac{\partial^2 t(\vec{\theta})}{\partial\theta_i \partial\theta_j} \propto \left(\delta_{ij} - \hat{\Psi}_{ij}\right) = A. \tag{2.28}$$

This defines a link between the image magnification and the curvature of the time-delay surface, given the relation between the Hessian matrix and the lensing Jacobian. A flat time-delay surface at the image positions will result in a divergent magnification factor, vice versa a more curved surface will result in a smaller magnification. Different stationary points result in different types of images:

- Type I images: in the minima of the surface, both (2.28) eigenvalues are positive, so $detA > 0$ and $trA > 0 \longrightarrow$ positive magnification;
- Type II images: in the saddle points of the surface, (2.28) eigenvalues have opposite signs, so $detA > 0 \longrightarrow$ negative magnification (i.e. reversed image parity, not de-magnified);
- Type III images: in the maxima of the surface, both (2.28) eigenvalues are negative, so $detA > 0$ and $trA < 0 \longrightarrow$ positive magnification;

Fig.(2.5) shows an example of one-dimensional time-delay functions for a circularly simmetric lens for a source not directly behind it (Narayan and Bartelmann, 1997 [11]). The geometric time-delay is described by a parabola and the gravitational function will vary depending on the considered potential. In particular, for a source perfectly aligned with the centre of the lens ($\beta = 0$), we observe a configuration called the Einstein ring: three stationary points, two of which (the minima) merge onto a single ring. We can then introduce the Einstein radius, which is properly an angle, to define the radius of the Einstein

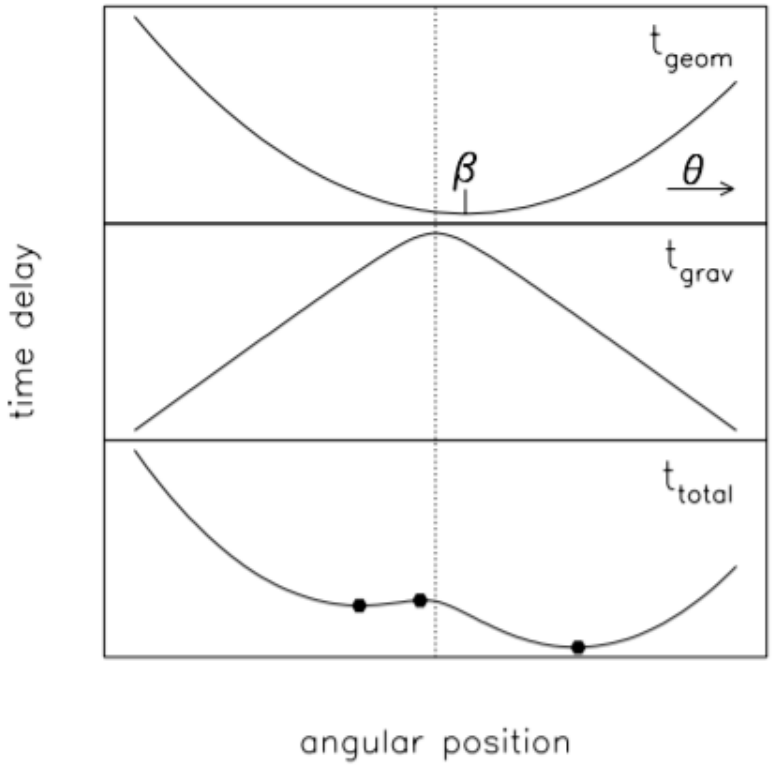

Fig. 2.5. Geometric, gravitational, and total time-delay one dimensional functions. Figure form Narayan and Bartelmann (1997 [11]).

ring as shown in Fig.(2.6):

$$\theta_E = \sqrt{\frac{4GM(\theta)}{c^2}\frac{D_{LS}}{D_s D_L}} \tag{2.29}$$

this can be used as reference scale to work in dimensionless form.

Note that the one-dimensional representation can be misleading, since the surface is three-dimensional: from the height difference of the stationary points of the surface is possible to determine the time-delay between the appearance of the images. For example, the explosion of a supernova in a galaxy lensed by MACS1149 galaxy cluster provided a opportunity to test model predictions of time-delay on short timescales compared to a human lifetime. (Treu et al., 2016 [16]);

Fig.(2.7) shows the time-delay surface contours in the area of the lensed galaxy, where we can see the multiple images of the Resfdal supernova. The time-delays between these images are of the order of a few days.

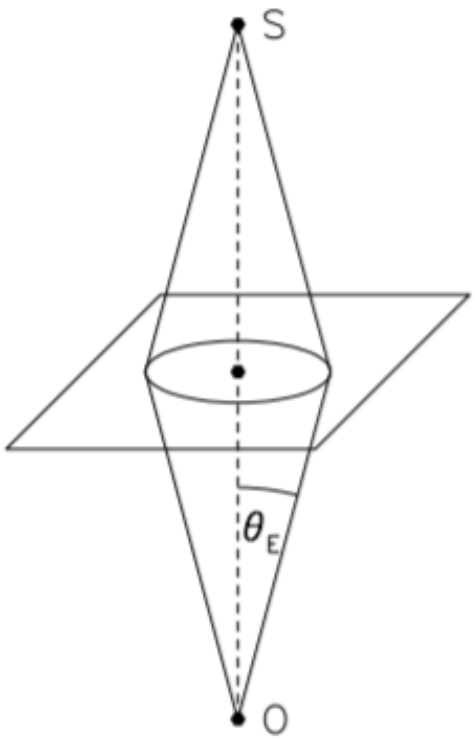

Fig. 2.6. A source S exactly behind the centre of an axially symmetric lens is mapped into the Einstein ring, with angular radius given by θ_E. Figure from Narayan and Bartelmann (1997 [11]).

2.3 Extended lens profiles

Massive galaxies and galaxy clusters are powerful gravitational lenses, capable of distorting and multiple-imaging several sources at the same time. Studying the mass distribution of these cosmic structures is important to understand how they form and evolve over cosmic time and to constrain the cosmological model (see e.g. Jullo et al. 2010 [17], Golse et al. 2002 [18]). Galaxies and clusters are characterized by complex extended mass distributions, whose properties are determined by their mass surface density. The surface density of an extended lens is characterized by several features, such as its mean density profile, the shape of iso-density contours, the smoothness or clumpiness of the lens. The lensing properties also depend on the environment surrounding the lens. This description of analytical lens models for extended lenses still use the thin screen approximation.

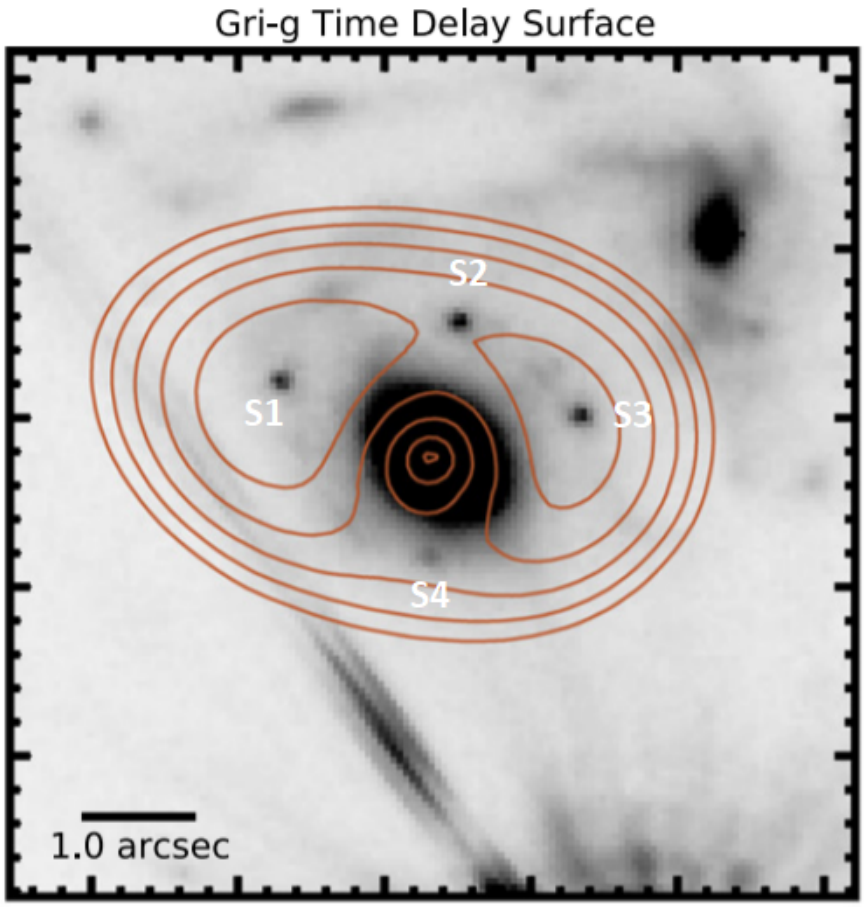

Fig. 2.7. Resfdal supernova in MACS1149. S1 is a minima, so it's taken as reference point and is the first multiple image appeared, while S2 is a saddle point and appeared with a time-delay of 5 days, S3 is another minima and with a time-delay of 11 days and S4 is a saddle point with a time-delay of 26 days. Figure from Treu et al., 2016 [16].

2.3.1 Axially symmetric profiles

For such lenses the properties at a given radius r are the same in every direction. For example, the lensing potential is $\hat{\Psi}(\vec{\theta}) = \hat{\Psi}(\theta)$. This symmetry allows a one-dimensional description of most equations. For a circularly symmetric lens, the deflection angle is radially directed and its amplitude only depends on the distance from the lens center, i.e. the deflection angle is parallel to $\vec{\theta}$ and its expression is the same of Eq.(2.4) for the point mass, but in this case $M = M(\theta)$ so the mass considered is the mass enclosed in a circle of radius θ, where the light ray passes. In dimensionless notation, introducing an arbitrary reference scale ξ_0, this can be written as

$$\alpha(x) \equiv \frac{m(x)}{x} \text{ with } m(x) = 2\int_0^x x'\kappa(x')\,\mathrm{d}x' \tag{2.30}$$

where $m(x)$ is the dimensionless mass, defined as $m(x) = \frac{M(\xi_0 x)}{\pi \xi_0^2 \Sigma_{cr}}$. Therefore the lens equation is

$$y = x - \frac{m(x)}{x}. \tag{2.31}$$

To derive an expression for the convergence κ we rewrite the Laplacian in Eq.(2.14) in polar coordinates. Then, from (2.15),

$$\kappa(x) = \frac{1}{2}\frac{m'(x)}{x} \tag{2.32}$$

and then the shear vector can be written as

$$\gamma(x) = |\kappa(x) - \bar{\kappa}(x)| \text{ with } \bar{\kappa}(x) = \frac{m(x)}{x^2} = 2\pi \frac{\int_0^x x'\kappa(x')\,\mathrm{d}x'}{\pi x^2} \tag{2.33}$$

where $\bar{\kappa}(x)$ is the mean convergence in a circle of radius x. Using these results it can be demonstrated (M. Meneghetti, 2017 [19]) that the determinant of the lensing Jacobian is

$$detA = \underbrace{\frac{y}{x}}_{\lambda_t}\underbrace{\frac{dy}{dx}}_{\lambda_r} = [1 - \bar{\kappa}(x)][1 + \bar{\kappa}(x) - 2\kappa(x)] \tag{2.34}$$

where λ_r and λ_t are, respectively, the radial and tangential eigenvalues of the Jacobian as defined in Sec.(2.2.2). Thus considering a circular source at the distance y from the center of the lens, its perimeter would be mapped onto the lens plane point by point, since the deflection angle is parallel to both x and y: as shown in Fig.(2.8) if the source dimension is δ this would be mapped into the mayor axis of the ellipse image, subtending the same angle. Therefore to find the tangential deformation we can simply compute the tangential eigenvalue of the Jacobian matrix, and to find the radial stretching we can compute the radial eigenvalue of the Jacobian matrix. Posing $detA$ in Eq.(2.34) equal to zero we can find the critical lines:

$$\alpha(x)/x = m(x)/x^2 = \bar{\kappa}(x) = 1 \tag{2.35a}$$

$$\alpha'(x) = m'(x)/x - m/x^2 = 2\kappa(x) - \bar{\kappa}(x) = 1. \tag{2.35b}$$

Eq.(2.35a) and Eq.(2.35b) give the tangential and radial critical lines, respectively. Both these conditions define circles on the lens plane.

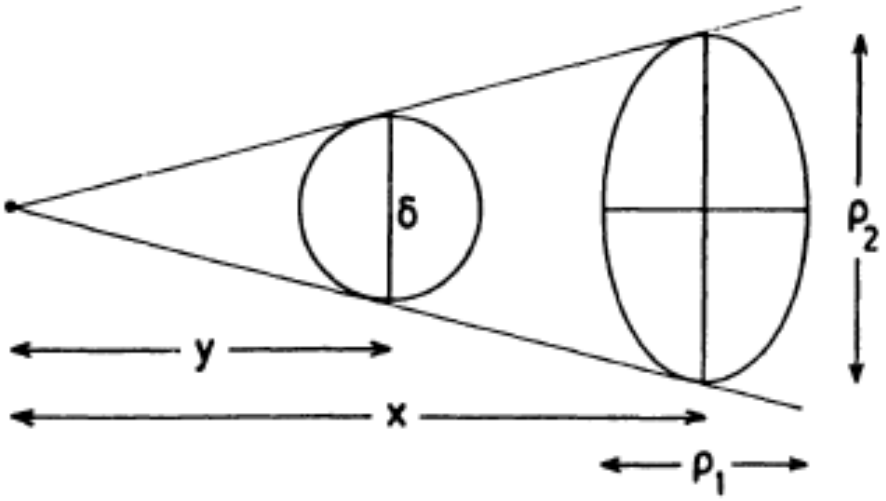

Fig. 2.8. Infinitesimal source of diameter δ mapped into an ellipse. Figure from P.Schneider, J. Ehlers, E.Falco (1992 [14]).

2.3.2 Power-law profiles

These profiles can be described by a convergence profile proportional to the n-th power of the radius x, such as

$$\begin{aligned} \kappa(x) = \frac{3-n}{2} x^{1-n} &\Rightarrow m(x) = x^{3-n} \\ &\Rightarrow \alpha(x) = x^{2-n} \\ &\Rightarrow \gamma(x) = \frac{n-1}{2} x^{1-n} \\ &\Rightarrow y = x - x^{2-n} \\ &\Rightarrow \det A = \underbrace{\left(1 - x^{1-n}\right)}_{\lambda_t} \underbrace{\left(1 - (2-n)x^{1-n}\right)}_{\lambda_r} \end{aligned} \tag{2.36}$$

where n is a parameter that can assume any real value, $n > 1$. In particular, we distinguish the following cases:

- **1<n<2**: mass and deflection angle profiles increase with x for all n, convergence and shear profiles decrease instead. The tangential critical line is a circles of radius $x = 1$ $\forall n$ and the corresponding caustic line is a point in $y_{crit} = 0$ $\forall n$; on the other hand, the radial critical line size depends on the value of n. Its size decreases as n increases, while the caustic line has the opposite trend. For $n = 2$, the radial critical line does not exist. For all axially simmetric lenses there are serveral ways to solve the lens equation, but the most useful is the so-called *image diagram*. Multiple image positions can be found at the intersection of the deflection angle profile with the lines $f(x) = x - y$

in order to find the multiple images positions x. As shown in Fig.(2.9) the image separation depends on the value of n: larger n gives a more curved deflection angle profile and thus more images with smaller separation and vice versa.

Lens profiles with $n = 1$ and $n = 2$ are the perfectly convergent lens and the Singular Isothermal Sphere model (see section (2.3.3)), respectively.

- **n>2**: the deflection angle profile decreases with x for all n and has a singularity in $x = 0$, namely these lenses always produce two images, one inside and one outside the Einstein radius, since the time-delay surface is not continuously deformable. The radial eigenvalue is never zero, giving an absolute radial magnification factor $\mu_r < 1$ always. Thus the images are always radially demagnified. The radial critical line does not exist. The case $n = 3$ corresponds to the point-mass lens.

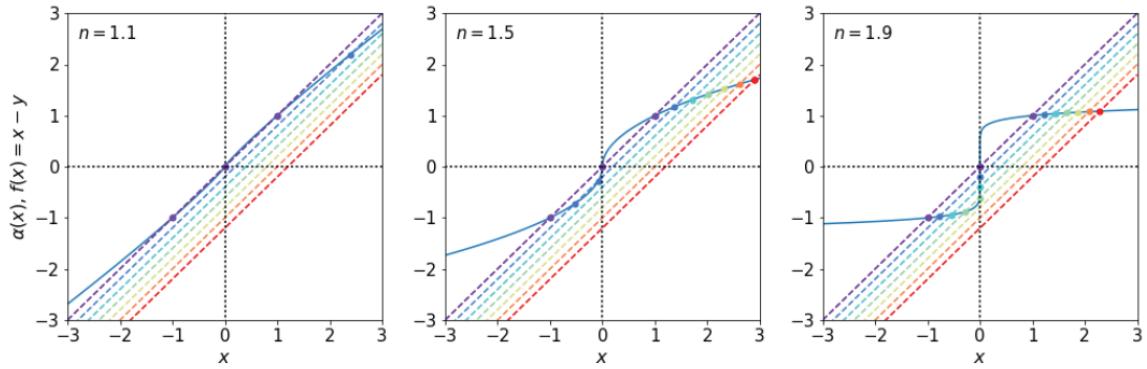

Fig. 2.9. Image diagrams for different values of n<2: the blue solid line is the deflection angle function, while the colored dashed lines are the $f(x) = x - y$ functions, with varying y value. Figure from M. Meneghetti, 2017 [19]

2.3.3 Singular Isothermal Sphere

A simple model to describe galaxy and cluster mass distributions is the Singular Isothermal Sphere (SIS), with surface mass density

$$\Sigma(\xi) = \frac{\sigma_v^2}{2G\xi} \tag{2.37}$$

where ξ the distance from the lens center. If this profile is described by assuming an ideal gas confined by a spherically symmetric potential, in thermal and hydrostatic equilibrium. The gas particle velocity dispersion is σ_v (i.e. stars in galaxy, galaxies in cluster). This profile is not physical, since it has a singularity at the center and the total mass is infinite, but the behaviour of (2.37) at $0 \ll \xi < \infty$

reproduces the observed flat rotation curves of spiral galaxies. By choosing $\xi_0 = 4\pi \left(\frac{\sigma_v}{c}\right)^2 \frac{D_L D_{LS}}{D_S}$ as length scale, we can write the dimensionless form of Eq.(2.37):

$$\Sigma(x) = \frac{\Sigma_{crit}}{2x} \Rightarrow \kappa(x) = \frac{1}{2x}. \tag{2.38}$$

Thus the SIS profile is a power-law lens profile with $n = 2$. From the convergence definition, we obtain the deflection angle and the lens equation:

$$\alpha(x) = \frac{x}{|x|} \tag{2.39a}$$

$$y = x - \frac{x}{|x|} \tag{2.39b}$$

The solutions of the lens equation depend on the source position y:

- when $0 < y < 1$ two solutions exist, on opposite sides of the lens center, one in $x_- = y - 1$ and one in $x_+ = y + 1$; their angular positions are $\boldsymbol{\theta}_\pm = \boldsymbol{\beta} \pm \boldsymbol{\theta_E}$, with θ_E from Eq.(2.29). Thus the image angular separation is always $\Delta(\theta_E) = 2\theta_E$.
- when $y > 1$ only one solution exists, i.e $x_+ = y + 1$.

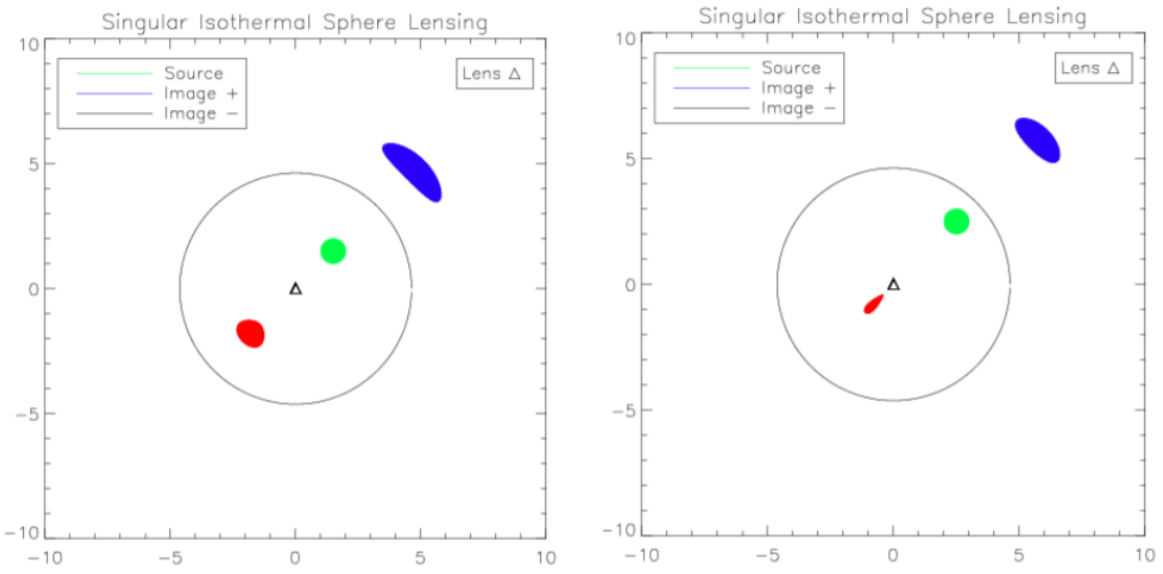

Fig. 2.10. Different alignments between a SIS lens (triangle) and an extended source (green dot). Figure from C. Jensen, 2012 [20].

Therefore the circle of radius $y_c = 1$, called the *cut*, has the same role of a radial caustic for power-law lens models with $n < 2$, separating the source plane in regions with different image multiplicities (but

it is not a caustic).

Computing the radial eigenvalue of the lensing Jacobian we find that it is $\lambda_r = 1$ always, meaning that the magnification is only tangential; thus the radial size of all images is always equal to the source's size, there is only a tangential stretch. The magnification for an image at x is given by

$$\mu(x) = \frac{y}{x}\underbrace{\frac{dy}{dx}}_{=1} = \frac{|x|}{|x|-1}. \tag{2.40}$$

Since for $y < 1$ the images form in x_- and x_+, substituting above

$$\underbrace{\mu_+(y)}_{\text{outside Einstein ring}} = 1 + \frac{1}{y} \quad ; \quad \underbrace{\mu_-(y)}_{\text{inside Einstein ring}} = 1 - \frac{1}{y} \tag{2.41}$$

from which we see that for y approaching unity the image in x_- fades until it disappears; moreover, if y increases μ_+ decreases and the image resembles more and more the source shape.

From Eq.(2.33), it follows that the shear and convergence profiles are equal

$$\gamma(x) = \frac{1}{2x} = \kappa(x) \tag{2.42}$$

with shear components:

$$\begin{aligned} \gamma &= \tfrac{1}{2}\tfrac{\cos 2\phi}{x} \\ \gamma &= \tfrac{1}{2}\tfrac{\sin 2\phi}{x} \end{aligned} \tag{2.43}$$

where ϕ is the polar angle.

Non-singular Isothermal Sphere To avoid the SIS profile central singularity an additional parameter ξ_c can be used to introduce a flat core in the surface density profile. The new modified surface density profile is then:

$$\Sigma(\xi) = \frac{\Sigma_0}{\sqrt{1+\xi^2/\xi_c^2}} \text{ with } \Sigma_0 = \frac{\sigma_v^2}{2G\xi_c}. \tag{2.44}$$

By introducing the same scale length ξ_0 defined in Sec.(2.3.3), and defining $x_c = \xi_c/\xi_0$, we can write the dimensionless convergence profile as

$$\kappa(x) = \frac{1}{2\sqrt{x^2+x_c^2}}. \tag{2.45}$$

The lens equation is

$$y = x - \sqrt{1 + \frac{x_c^2}{x^2}} - \frac{x_c}{x}, \tag{2.46}$$

which can be reduced to a third order polynomial. Thus the NIS lens can produce up to three images of a given source whose multiplicity depends on x_c. In particular it can be shown (M. Meneghetti, 2017 [19]) that the tangential and radial critical lines only exist for $x_c < 1/2$ (in θ_E units), and that the tangential caustic line is a point at $y_t = 0$ while the radius of the radial caustic line varies with x_c, as shown in Fig.(2.11). If $x_c > 1/2$ the convergence is never larger than unity, meaning the surface

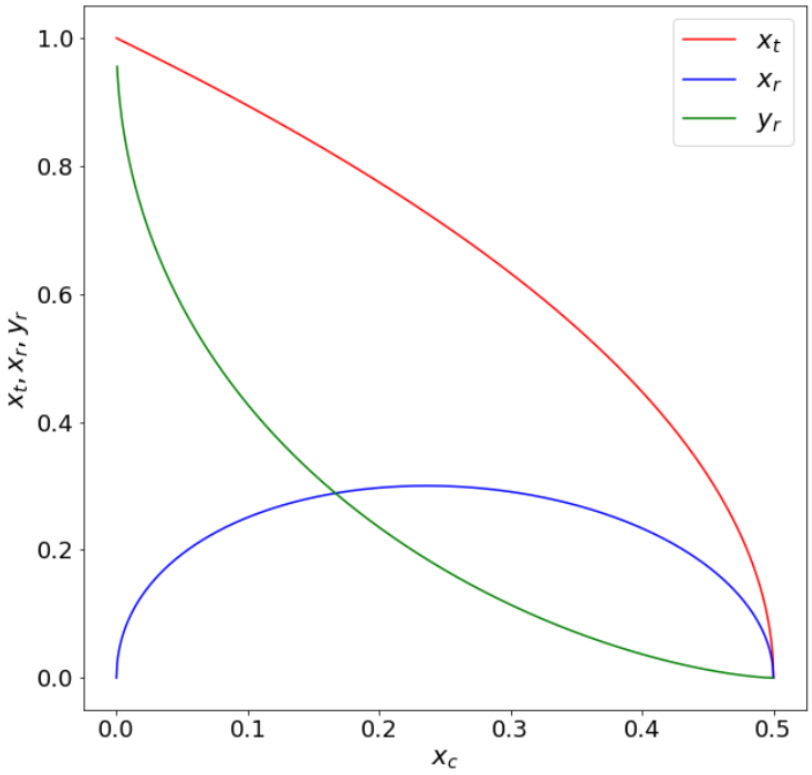

Fig. 2.11. NIS critical lines radius as function of x_c. Note that for $x_c = 0$, $y_r = 1$, i.e. we find the SIS cut. Figure from M. Meneghetti, 2017 [19].

density never exceeds the critical surface density $\Sigma_{\rm cr}$ and we are talking about a "weak" lens that can not produce multiple images because the critical lines are not produced.

2.3.4 Singular Isothermal Ellipsoid

Observations show that galaxies are better represented by elliptical mass distributions. Kormann, Schneider, and Bartelmann (1994 [21]) developed a lens model of the Singular Isothermal Ellipsoid (SIE), introducing the ellipticity in the SIS profile. In particular, they substitute the radial coordinate

ξ with $\sqrt{\xi_1^2 + f^2\xi_2^2}$, obtaining

$$\Sigma(\vec{\xi}) = \frac{\sigma_v^2}{2G}\frac{\sqrt{f}}{\sqrt{\xi_1^2 + f^2\xi_2^2}}. \tag{2.47}$$

The surface density in Eq.(2.47) is constant on ellipses with axis ratio $f = b/a$. By choosing the same reference scale length ξ_0 as for the SIS, we find the convergence

$$\kappa(\vec{x}) = \frac{\sqrt{f}}{2\sqrt{x_1^2 + f^2x_2^2}}. \tag{2.48}$$

The deflection angle can be computed from the gradient of the lensing potential in polar coordinates; in analogy with the SIS case, it can be shown that the shear is $\gamma = \kappa$ for the SIE as well. From the lensing Jacobian determinant, we find the two eigenvalues

$$\begin{aligned} \lambda_t(\vec{x}) &= 1 - 2\kappa(\vec{x}) \\ \lambda_r(\vec{x}) &= 1 \end{aligned} \tag{2.49}$$

from which we see that even the SIE does not produce a radial critical line, and the radial magnification is always unity; the tangential critical line is the ellipse of equation $\kappa(\vec{x}) = 1/2$. Due to the singularity at the center of the lens, even for the SIE profile we find a multiple image region not enclosed by the caustic, but by the cut. The number of multiple images that can be produced by this lens depends on how the caustic and cut lines are placed: as shown in Fig.(2.12), depending of the value of f, the cut (the ellipse) can enclose the tangential caustic (the astroid) or not, so since crossing the cut changes by 1 the image number, and crossing the caustic changes it by 2, the first three configurations can produce 1/2/4 images; in the last case instead, 3 multiple images can be produced since two cusps of the astroid cross the cut, but it can be demonstrated that this is possible only for $f < 0.3942$.

Non-singular Isothermal Ellipsoid To remove the singularity of the surface mass density, in the SIE model, we introduce a core radius ξ_c, so we can write the surface mass density and the convergence as

$$\Sigma(\vec{\xi}) = \frac{\sigma^2}{2G}\frac{\sqrt{f}}{\sqrt{\xi_1^2 + f^2\xi_2^2 + \xi_c^2}} \tag{2.50a}$$

$$\kappa(\vec{x}) = \frac{\sqrt{f}}{2\sqrt{x_1^2 + f^2x_2^2 + x_c^2}} \tag{2.50b}$$

Kormann, Schneider, and Bartelmann (1994) [21] discussed the critical lines and caustics topology and image multiplicity. In Fig.(2.13) we show that depending on the value of f and x_c, the lens can produce different multiplicity of a source behind it. The lens can have both radial and tangential critical lines

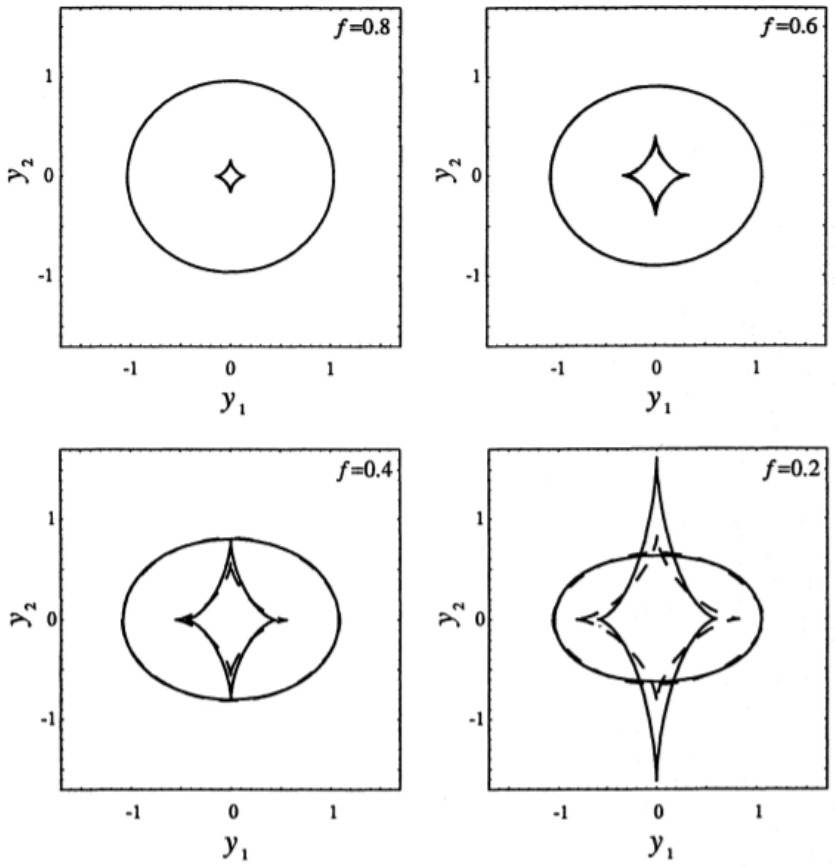

Fig. 2.12. Different f configurations of the SIE cut (ellipse) and tangential caustic (astroid). Figure from Kormann et al., 1994[21].

and caustics, producing up to 5 multiple images. For some other values, only the tangential critical line and caustic exist, so it can produce 1 or 3 multiple images. For some value, it can have no critical and caustic line at all.

2.3.5 The PIEMD profile

The description of this profile is fully provided in Elíasdóttir et al. (2017) [22], to which we refer in this work. The Pseudo Isothermal Elliptical Mass Distribution (PIEMD) profile has different variants and it a very useful tool in lens modeling; its main perks are analytic expressions of the lensing potential and first and second partial derivatives. It was first defined by Kassiola & Kovner (1993) [23] without a scale radius. The following expressions refer to a PIEMD profile with a core radius and a scale radius, also called the dual Pseudo Isothermal Elliptical mass distribution (dPIE). Thus, considering a spherical

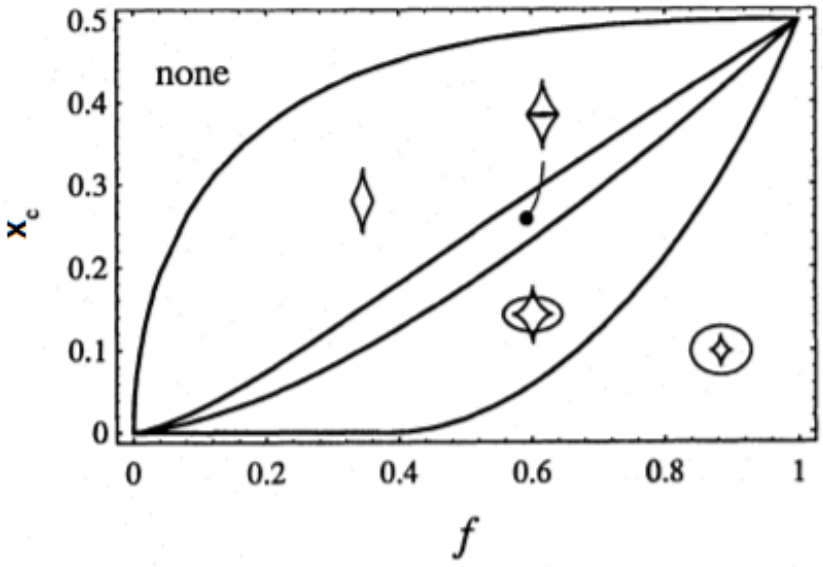

Fig. 2.13. NIE caustics topologies varying f and x_c. Figure from Kormann et al., 1994 [21]

system with scale radius s, the three-dimensional density distribution is

$$\rho(r) = \frac{\rho_0}{(1 + r^2/a^2)\,(1 + r^2/s^2)}; \quad s > a, \tag{2.51}$$

where a is the core radius and ρ_0 is the central density. The relation between the central density and the 1D-central velocity dispersion σ_0, is (Limousin et al., 2005 [24])

$$\rho_0 = \frac{\sigma_0^2}{2\pi G}\frac{a+s}{a^2 s}. \tag{2.52}$$

For $r < a$, the density follows $\rho \approx \rho_0 / \left(1 + r^2/a^2\right)$. For $a \leq r \leq s$, the density is isothermal. Finally, for $r > a$, the density falls off as $\rho \sim r^{-4}$. The surface mass density is written as

$$\begin{aligned} \Sigma(\xi) =& 2\int_{\xi}^{\infty} \frac{\rho(r) r}{\sqrt{r^2 - \xi^2}} \mathrm{d}r \\ =& \Sigma_0 \frac{as}{s-a}\left(\frac{1}{\sqrt{a^2+\xi^2}} - \frac{1}{\sqrt{s^2+\xi^2}}\right) \end{aligned} \tag{2.53}$$

where ξ is the two-dimensional radius and Σ_0 is

$$\Sigma_0 = \pi\rho_0 \frac{as}{s+a}. \tag{2.54}$$

The convergence κ can be computed from (2.15) and the shear is

$$\begin{aligned}\gamma(\xi) = \frac{\Sigma_0}{\Sigma_{\rm crit}}\frac{as}{s-a}\Bigg[2\left(\frac{1}{a+\sqrt{a^2+\xi^2}}-\frac{1}{s+\sqrt{s^2+\xi^2}}\right)\\ +\left(\frac{1}{\sqrt{a^2+\xi^2}}-\frac{1}{\sqrt{s^2+\xi^2}}\right)\Bigg]\end{aligned}. \tag{2.55}$$

Integrating Eq.(2.51), we can compute the total 3D mass enclosed by the physical radius r, as

$$\begin{aligned}M_{3D}(r) =&4\pi\int_0^r \rho(\bar{r})\tilde{r}^2{\rm d}\bar{r}\\ =&4\pi\rho_0\frac{a^2s^2}{s^2-a^2}\left[s\arctan\left(\frac{r}{s}\right)-a\arctan\left(\frac{r}{a}\right)\right]\end{aligned}. \tag{2.56}$$

We can now define the mass enclosed in the projected radius R, in order to obtain an expression for the projected lensing potential, thus

$$\begin{aligned}M_{2D}(\xi) =&2\pi\int_0^\xi \Sigma(\bar{\xi})\bar{\xi}{\rm d}\bar{\xi}\\ =&2\pi\Sigma_0\frac{as}{s-a}\left(\sqrt{a^2+\xi^2}-a-\sqrt{s^2+\xi^2}+s\right)\end{aligned}. \tag{2.57}$$

Note that this profile has a finite total mass, given by $M_{tot} = 2\pi\Sigma_0 as$. Setting ξ_L as a limiting radius, we can integrate the 2D mass profile to obtain the projected lensing potential as

$$\begin{aligned}\Psi(\xi) =&2G\int_\xi^{\xi_L}\frac{M_{2D}(\tilde{\xi})}{\xi}{\rm d}\bar{\xi}\\ =&4\pi G\Sigma_0\frac{as}{s-a}\left(\sqrt{s^2+\xi^2}-\sqrt{a^2+\xi^2}+a\ln\left(a+\sqrt{a^2+\xi^2}\right),\right.\\ &\left.-s\ln\left(s+\sqrt{s^2+\xi^2}\right)\right)+\text{ constant}\end{aligned} \tag{2.58}$$

from which we can finally compute the deflection angle:

$$\begin{aligned}\alpha(\xi) =&-\frac{2}{c^2}\frac{D_{LS}}{D_S}\frac{{\rm d}\Psi}{{\rm d}\xi}\\ =&\frac{8\pi G}{c^2}\frac{D_{LS}}{D_S}\Sigma_0\frac{as}{s-a}\left(\frac{\xi/a}{1+\sqrt{1^2+(\xi/a)^2}}-\frac{\xi/s}{1+\sqrt{1^2+(\xi/s)^2}}\right)\end{aligned} \tag{2.59}$$

LENSTOOL As we will further discuss in Chapter (4.2), the dPIE is implemented in the LENSTOOL software, that will be used in this work to construct a lens model. Hereafter, we will refer to cluster-scale halos as PIEMD profiles (which is a dPIE in the limit $r_{cut} \longrightarrow \infty$), and to galaxy-scale halos as dPIEs. The profile is defined by eight parameters: the redshift (z), the central position (RA,DEC), the ellipticity (e) and orientation (θ), the scale radius (r_{cut}), the core radius (r_{core}) and a central velocity

dispersion (σ_0). We note that the LENSTOOL parametrization of the dPIE profile requires a fiducial velocity dispersion σ_{LT}, related to the σ_0 as $\sigma_0 = \sqrt{3/2}\sigma_{LT}$. In particular, the projected velocity dispersion can be obtained using (2.56), thus

$$\sigma_P^2(\xi) = \frac{2G}{\Sigma(\xi)} \int_\xi^\infty \frac{M_{3D}(r)\rho(r)}{r^2} \sqrt{r^2 - \xi^2} \mathrm{d}r. \tag{2.60}$$

2.3.6 The Navarro-Frenk-White profile

Navarro, Frenk and White (1995) [25] derived from N-body numerical simulations a simple universal model to fit the density profiles of dark matter halos in a CDM scenario. The NFW profile is given by

$$\rho(r) = \frac{\rho_s}{(r/r_s)\,(1 + r/r_s)^2} \tag{2.61}$$

where r_s is the scale radius, and ρ_s is defined from $\rho(r = r_s) = \frac{\rho_s}{4}$. The profile is isothermal near r_s, but is significantly shallower than r^{-2} near the center, following $\rho(r) \propto r^{-1}$, and is steeper at large radii, changing to $\rho(r) \propto r^{-3}$. The total mass of a DM halo can be expressed in terms of a virial overdensity as

$$M_{vir} = \frac{4\pi}{3} r_{vir}^3 \frac{\Delta_{vir}}{\Omega_m(z)} \Omega_0 \rho_c \tag{2.62}$$

where Δ_{vir} is the overdensity within the virial radius r_{vir} and depends on cosmology and redshift, and $\Omega_m(z)$ as defined in section (1.3.2). In NFW framework to parametrized the DM halos the mass M_{200} is used, which is the mass enclosed by the virial radius r_{200}, namely the radius of a sphere of mean overdensity of 200 with respect the critical density (1.25): thus the covered mass range in NFW work is $\sim 3 \times 10^{11} < M_{200}/M_\odot <\sim 3 \times 10^{15}$. Another useful definition is the virial concentration $c_{vir} = r_{vir}/r_s$. From the definition of r_{200}, the concentration $c \equiv r_{200}/r_s$ can be linked to the characteristic density as

$$\rho_s = \frac{200}{3} \rho_c \frac{c^3}{(\ln(1 + c) - c/(1 + c))} \tag{2.63}$$

NFW found from the numerical simulations that the concentration, or equivalently the overdensity, is a function of the halo mass by reflecting the different formation redshift of different mass halos.

The description of different aspects of lensing by halos with NFW density profiles is derived in Bartelmann (1996 [26]) : the projected density profile is

$$\Sigma(x) = \frac{2\rho_s r_s}{x^2 - 1} f(x), \tag{2.64}$$

with

$$f(x) = \begin{cases} 1 - \frac{2}{\sqrt{x^2-1}} \arctan \sqrt{\frac{x-1}{x+1}} & (x > 1) \\ 1 - \frac{2}{\sqrt{1-x^2}} \operatorname{arctanh} \sqrt{\frac{1-x}{1+x}} & (x < 1) \\ 0 & (x = 1) \end{cases} . \tag{2.65}$$

So the lensing potential is given by

$$\Psi(x) = 4\kappa_{\rm s} g(x), \tag{2.66}$$

with $\kappa_s \equiv \rho_s r_s \Sigma_{\rm cr}^{-1}$ and

$$g(x) = \frac{1}{2} \ln^2 \frac{x}{2} + \begin{cases} 2 \arctan^2 \sqrt{\frac{x-1}{x+1}} (x > 1) \\ -2 \operatorname{arctanh}^2 \sqrt{\frac{1-x}{1+x}} (x < 1) \\ 0 (x = 1) \end{cases} . \tag{2.67}$$

Thus the deflection angle is

$$\alpha(x) = \frac{4\kappa_{\rm s}}{x} h(x) \tag{2.68}$$

with

$$h(x) = \ln \frac{x}{2} + \begin{cases} \frac{2}{\sqrt{x^2-1}} \arctan \sqrt{\frac{x-1}{x+1}} (x > 1) \\ \frac{2}{\sqrt{1-x^2}} \operatorname{arctanh} \sqrt{\frac{1-x}{1+x}} (x < 1) \\ 1 (x = 1) \end{cases} \tag{2.69}$$

Integrating the convergence we can obtain the dimensionless mass, as

$$\kappa(x) = 2\kappa_s \frac{f(x)}{x^2 - 1} \longrightarrow m(x) = 4k_s h(x). \tag{2.70}$$

The logarithmic slope of the surface density in Eq.(2.64) is significantly smaller than the SIS one at the center, so the magnification is larger. On the contrary, the profile is significantly steeper than isothermal at large radii and the decrease away from the critical lines is shallower.

2.3.7 External shear

When describing a lens placed in a particularly dense environment it is often needed to account for every other possible gravitational lens surrounding it. To parametrize this environment effect, ot often

used an external shear field, defined by means of a potential Ψ_γ that satisfies

$$\begin{aligned}\gamma_1 &= \tfrac{1}{2}\left(\Psi_{11} - \Psi_{22}\right) = \text{ const.}\\ \gamma_2 &= \Psi_{12} = \text{ const.}\\ \kappa &= \tfrac{1}{2}\left(\Psi_{11} + \Psi_{22}\right) = \text{ const.}\end{aligned} \tag{2.71}$$

so that Ψ_{11} and Ψ_{22} are constant. In this case, the potential is quadratic, i.e.

$$\Psi_\gamma = Cx_1^2 + Cx_2^2 + Dx_1x_2 + E \tag{2.72}$$

so the conditions in Eq.(2.71) can be expressed as

$$\begin{aligned}\gamma_1 &= \tfrac{1}{2}\left(\Psi_{11} - \Psi_{22}\right) = C - C'\\ \gamma_2 &= \Psi_{12} = D\\ \kappa &= \tfrac{1}{2}\left(\Psi_{11} + \Psi_{22}\right) = C + C'\end{aligned} \tag{2.73}$$

We distinguish two cases:

- Placing the lens on a matter sheet: there are no privileged directions, so the shear is zero, $\gamma_1 = \gamma_2 = 0$. Neglecting irrelevant constants the external convergence potential is

$$\Psi_\kappa = \frac{\kappa}{2}x^2 \tag{2.74}$$

  thus the deflection angle and the lens equation can be expressed as

$$\vec{\alpha} = \vec{\nabla}\Psi_\kappa = \kappa\vec{x}, \tag{2.75a}$$

$$\vec{y} = \vec{x} - \vec{\alpha} = \vec{x}(1-\kappa). \tag{2.75b}$$

  thus if $\kappa = 1$, every position $\vec{x}$ on the lens plane is mapped onto the same point $y = 0$, having a perfectly convergent lens.

- If the perturbation does not contribute to the convergence ($\kappa = 0$), then $C = -C' = \gamma_1/2$. The external shear potential is

$$\Psi_\gamma = \frac{\gamma_1}{2}\left(x_1^2 - x_2^2\right) + \gamma_2 x_1 x_2, \tag{2.76}$$

or in polar coordinates

$$\Psi_\gamma = \frac{\gamma}{2} x^2 \cos 2\left(\phi - \phi_\gamma\right), \tag{2.77}$$

where ϕ_γ is the angle defining the direction of the external shear.

Chapter 3

Galaxy Clusters

Galaxy clusters are bound systems of galaxies whose existence and observations currently helps our understanding and measurements of the main cosmological parameters, on top of providing a probe for the largely accepted standard cold dark matter scenario. As already explained in Sec.(1.4), the existence of these large-scale structures is the evidence of our not perfectly homogeneous Universe: the gravitational instability of primordial density perturbations leads to the formation of small structures that then merge forming galaxy clusters. Gravitational lensing is one of the most powerful methods to constrain the mass distribution of clusters, which also helps understanding the interaction between dark and baryonic matter. In this Chapter we describe the main properties of galaxy clusters, focusing on their strong lensing properties.

3.1 Main properties

To distinguish a galaxy clusters from other bound structures such as groups of galaxies we can rely on a richness criterion defined in Abell (1958 [27]) : if m_3 is the magnitude of the third brightest cluster member (within $\simeq 1.5h^{-1}$ Mpc from the cluster center) then a cluster must contain at least 30 galaxies brighter than $m_3 + 2^m$. Observed clusters show a population of 100-1000 galaxies, mostly Early Type Galaxies (ETGs), with a median line-of-sight velocity dispersion of tipically ~ 750km/s; from the velocity dispersion we find a total mass of $M_{tot} \sim 5 \times 10^{14} h^{-1} M_{\odot}$, but the stellar mass is only a small fraction of the total. Another important component is the hot gas, which is in virial equilibrium in the cluster's potential well. This intracluster medium (ICM) extends as far as the galaxy concentration and is extremely rarefied, with small density and high temperature: the temperature range is tipically $2 < T < 14$ keV, and the number density is $n \sim 0.1 - 0.01\text{cm}^{-3}$ in the center, and $n \sim 10^{-4}\text{cm}^{-3}$ in

Property	Value for clusters
Richness	30-300 galaxies
Radius	$(1-2)h^{-1}$Mpc
Velocity dispersion (los)	400-1400 km/s
Median velocity dispersion (los)	750 km/s
Mass ($1.5h^{-1}$Mpc)	$10^{14}-10^{15}h^{-1}M_{\odot}$
Luminosity (B band)	$6\times10^{11}-6\times10^{12}h^{-2}L_{\odot}$
$\langle\mathcal{M}/L_{\rm B}\rangle$	$\sim 300h\mathcal{M}_{\odot}/L_{\odot}$
T_X	2-14 keV
L_X	$\left(10^{42}-10^{45}\right)h^{-2}$ergs^{-1}
Mass fraction in galaxies	5%

Tab. 3.1. Main properties of rich galaxy clusters, for comparison with groups see Bahcall (1996 [28]).

the outskirts. Through the bremsstrahlung radiation from this hot plasma we can detect and measure the X-ray emissivity along the line of sight, and compute the total X-ray luminosity. The typical X-ray luminosity is of the order of $L_X \sim 10^{44}$ erg/s. These properties are listed in Bahcall (1996 [28]), along with others summarized in Tab.(3.1).

3.1.1 Cluster members

The formation and evolution of galaxies in clusters is still not fully understood since cluster member galaxies show many differences with respect to field galaxies. Mostly differences are attributed to the dense environment that can affect the galaxy interstellar medium (ISM) through ram pressure stripping (Gunn and Gott, 1972 [29]), evaporation (Cowie and Songalia, 1977 [30]), merging and other dynamical processes (Gorkom 2003 [31], Nulsen 1982 [32]). For instance, the Brightest Central Galaxies (BCGs) in clusters are classified as cD galaxies and are only found in cluster cores. Thus in clusters different morphologies are found at different densities: Dressler (1980 [33]) found a relationship between density and morphological type, for which ETGs and S0 are mostly found at high densities while spirals and irregulars exists in regions where gas density is lower; moreover the proportion of ETGs and late-type galaxies (LTGs) is different, being ETGs predominant in clusters while in the field we find $\mathrm{E}=10\%, \mathrm{S0}=10\%$, and $\mathrm{Sp}+\mathrm{Irr}=80\%$ (Sandage & Tamman 1979 [34]).

It is possible to highlight this population segregation in cluster in terms of star-formation (SF): high SF is detected in blue galaxies with a young stellar population, while low to none is found in red, quiescent galaxies. The dependence on environment of the color–magnitude relation of galaxies is described in Hogg et al. (2004 [35]). In Fig.(3.1) we show how in high density environments (e.g. cluster cores) red colors, hence ETGs, dominate the population. Galaxies are an important component of the cluster system that we need to consider in our mass models and that can have a non negligible role in creating

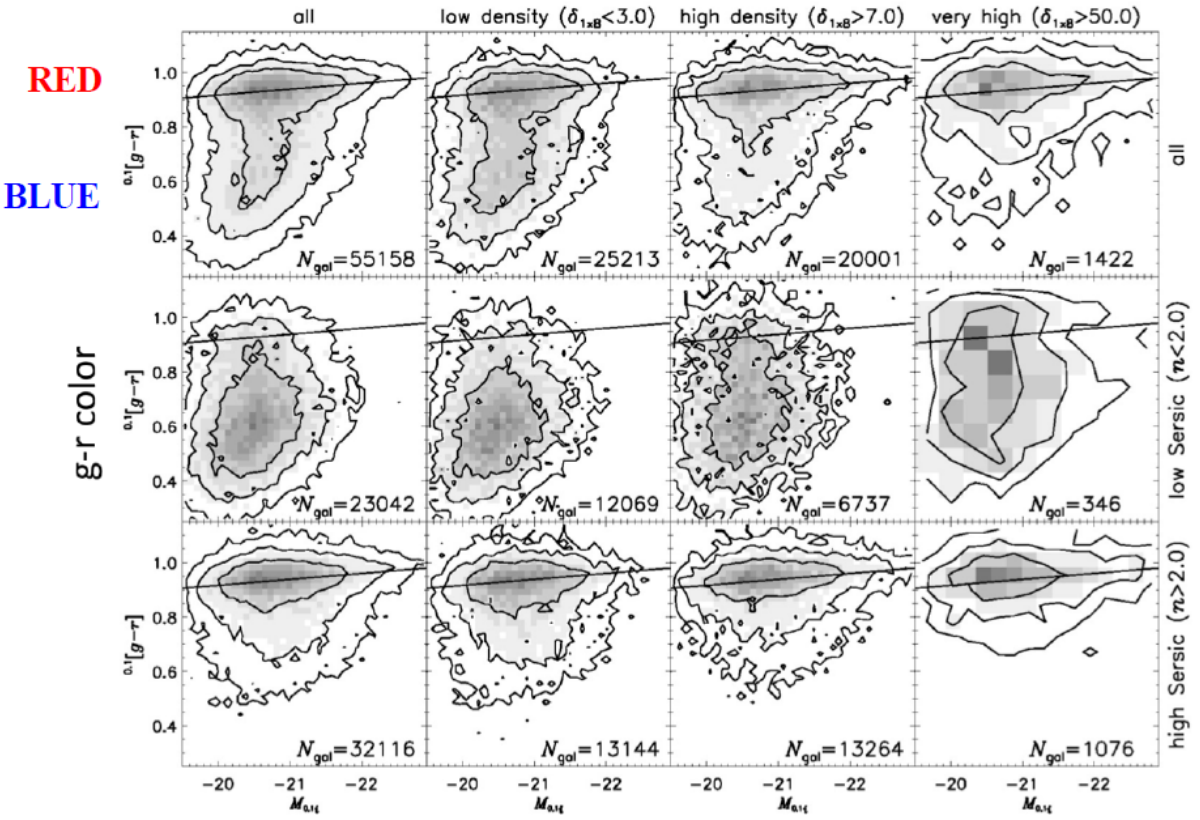

Fig. 3.1. Distribution in absolute i-band magnitude and [g-r] color for a galaxies sample. In each panel the grey scale monotonically represents the abundace of sample galaxies. The first row shows how the entire sample distribute at increasing density, we can see the separation in the red sequence (ETGs, low SF), and in the blue cloud (LTGs, high SF). At higher densities, hence clusters core, mostly the red sequence is present.

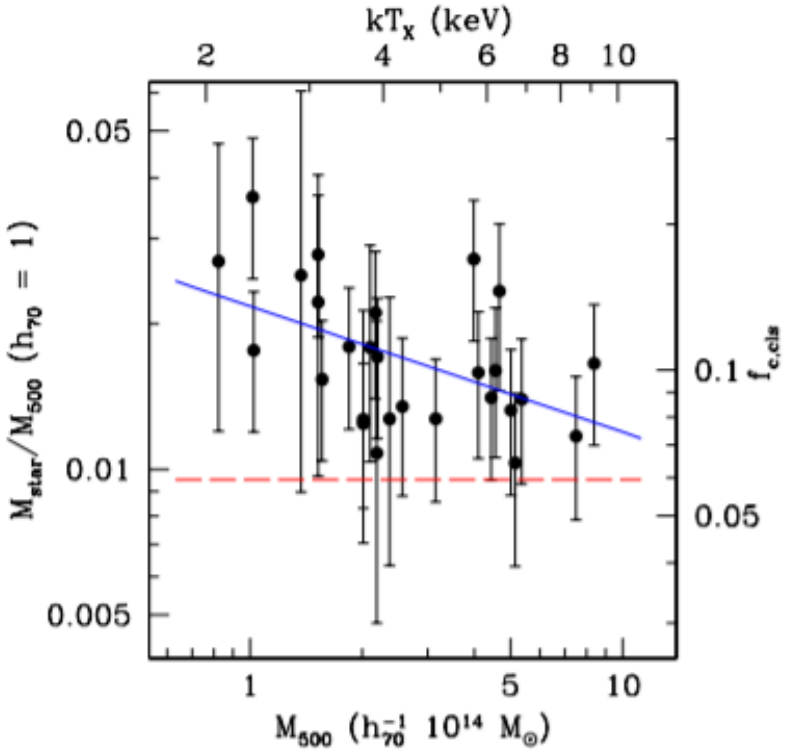

Fig. 3.2. Stellar mass fraction as a function of the cluster mass. The cluster mass is defined as in Eq.(2.62) with $\Delta_{vir} = 500$. The blue solid line marks the best fit relation for a 27 clusters sample. Figure from Lin et al., 2003 [37].

perturbations in the lensing signal. Nevertheless, the stellar component in clusters is not the dominant one, not even among baryonic matter: in fact, the cosmic value for $\rho_{baryons}/\rho_{tot}$ is ~ 0.17 (Spergel et al., 2003 [36]), but stars in cluster make up only $\rho_{stars}/\rho_{tot} \sim 0.02$ (Lin et al., 2003 [37]). In Fig.(3.2) we can see a plot that relates the fraction of halo mass that has been turned into stars in a 27 clusters sample with data from the 2MASS survey (Lin et al., 2003 [37]) and the total halo M_{500} mass: the considered mass is defined as in Eq.(2.62) at r_{500}, the radius of a sphere of mean overdensity of 500 with respect the critical density. Note the decreasing trend, namely the most massive clusters have a smaller stellar mass fraction than the least massive ones, in a range of $0.01 - 0.05$. Lin et al. (2003 [37]) also showed that massive galaxy clusters have a baryon fraction approaching the cosmic baryon fraction. Since stars in clusters only make up for a 0.02, baryons must be present in a different component. Considering that clusters potential wells are extremely deep, it is reasonable to assume that baryons can not escape. Thus most baryons must be in form of gas, with a fraction of $\rho_{gas}/\rho_{tot} \sim 0.15$.

3.1.2 Gas

Due to the high temperature (T$\sim 10^8$K) of the gas, galaxy clusters are luminous X-ray sources and this emission rises primarly from thermal bremsstrahlung from the ICM. Sazarin (1988 [38]) reviewed the physical properties of hot diffuse plasma: the particles are assumed to have a Maxwell-Boltzmann distribution at the temperature T; the particles mean free paths are generally much shorter than the length scales of clusters ($\sim$ 1 Mpc) so the ICM can be considered a collisional fluid. At the characteristic low densities of the ICM, excitation and de-excitation processes are described in the framework of collisional ionization equilibrium (CIE) (Smith and Hughes, 2010 [39]). The gas is generally thought to be in hydrostatic equilibrium in the potential well of the cluster, so under the assumption that the ICM is locally homogeneous and the cluster is spherically symmetric, the hydrostatic equilibrium equation can be written as (Gitti et al., 2012 [40]):

$$\frac{1}{\rho}\frac{dp}{dr} = -\frac{d\phi}{dr} = -\frac{GM(r)}{r^2}, \tag{3.1}$$

where $p = \rho kT/\mu m_p$ is the gas pressure, ρ is the gas density, ϕ is the gravitational potential of the cluster and *M(r)* is the total cluster mass enclosed in the radius r from the cluster center. Neglecting the gas self-gravity, its distribution can be fully determined by the potential $\phi(r)$ and the temperature $T(r)$ radial profiles. The galaxy cluster's total the gravitational mass M_{tot} is:

$$M_{\rm tot}(< r) = -\frac{kTr}{G\mu m_p}\left[\frac{d\ln\rho}{d\ln r} + \frac{d\ln T}{d\ln r}\right]. \tag{3.2}$$

Thus, measuring the density and the temperature profiles, $\rho(r)$ and $T(r)$, it is possible to estimate the gravitational mass of the cluster (see e.g. Gitti et al., 2007 [41], Voigt and Fabian, 2006 [42]).

An important and extensively used model for the density profile is the $\beta-$*model*, derived by Cavaliere and Fusco-Femiano (1976 [44]). They assume that both ICM and galaxies are in equilibrium in the cluster potential well and that the galaxy distribution is described by a King profile (King, 1962 [45]); moreover both gas and galaxies are considered "isothermal", namely the gas temperature is constant and the galaxies' velocity dispersion is isotropic. Under these assumptions the density profile can be written as

$$\rho(r) = \rho_0\left[1 + \left(\frac{r}{r_{\rm core}}\right)^2\right]^{-(3/2)\beta}, \tag{3.3}$$

where the parameter β is

$$\beta = \frac{\sigma_r^2}{kT/\mu m_p}. \tag{3.4}$$

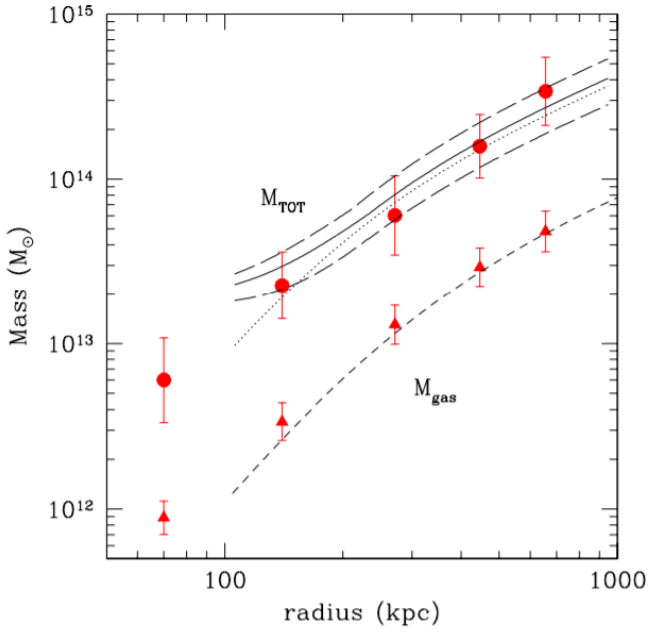

Fig. 3.3. Mass profiles of MS0735+7421 (Gitti et al., 2007 [43]). Red circles: total mass profile derived from deprojected gas temperature and density profiles. Black solid line: total mass calculated with Eq.(3.5) by estimating the temperature gradient from the deprojected temperature profile. The corresponding to the 90% confidence limits are given by the long dashed lines. Black dotted line: total mass calculated with Eq.(3.5) by assuming a constant temperature of 4.79 keV. Red triangles: gas mass profile derived from deprojected analysis. Black dashed line : gas mass profile derived from $\beta-$model fit.

In the last equation σ_r is the line-of-sight velocity dispersion. The great advantage of this model is that we can recover an analytical expression for the total mass:

$$M_{\rm tot}(<r) = \frac{kr^2}{G\mu m_p}\left[\frac{3\beta r T}{r^2+r_c^2} - \frac{dT}{dr}\right]. \tag{3.5}$$

Since the gas density increases in the central regions and the bremsstrahlung emission depends on the square of the gas density, the X-ray emissivity increase in those regions causing the gas to cool and flow towards the center: the cooling flows (see e.g. Ettori and Brighenti, 2008 [46]). In cool-core clusters the central regions are not well described by a single $\beta-$model, but a second $\beta-$model is needed (Fig. 3.4).

3.1.3 Dark Matter

Since the first studies of radial velocity in rich clusters, it became clear that the cluster visible and dynamical masses do not match (Zwicky, 1933 [48]). The discrepancy was reduced by X-ray observa-

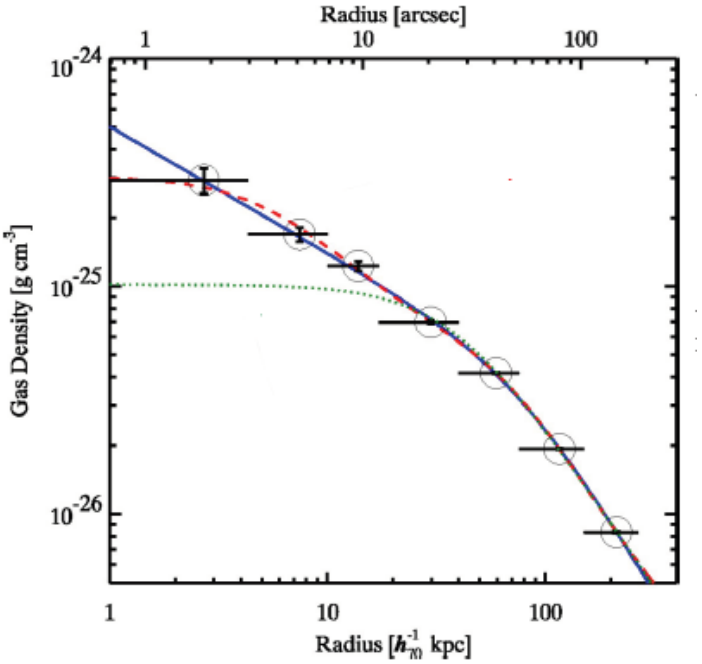

Fig. 3.4. *Chandra* radial gas density profile of Abell 2029, a cool-core cluster (Lewis et al., 2003 [47]): the large circles are centered on the data point, the dotted green curve is a single $\beta-$model, the dashed red curve is the double $\beta-$model.

tions that accounted for the gaseous mass, but it is not resolved without introducing the dark matter component dominating the overall mass (Sanders, 2003 [49]). Extensive reviews on galaxy clusters as cosmological probes can be found in e.g. Voigt (2005 [50]) and Allen et al. (2012 [51]); these works highlights that, as already discussed in Sec.(2.3.6), numerical simulations of cluster formation yield density profiles for the DM halos that are shallower than isothermal in the central regions and steeper than isothermal at large radii (see e.g. Gao et al. 2012 [52]; Newman et al. 2013a [53], 2013b[54]) can characterize an halo in terms of its virial concentration parameter c_{vir}, defined as the ratio between the virial radius and the scale radius (see Sec.(2.3.6)). The scale radius corresponds to the radius where the NFW profile slope is isothermal. Thus this radius marks the transition of the NFW profile from one trend to the other. Typical concentration parameters for simulated clusters are in the range $c \sim 4-10$ (Jing, 2000 [55]), and the higher concentration parameters are found in lower-mass objects, formed earlier in time when the overall density of the Universe was larger (Navarro et al., 1997 [56]). However recent N-body simulations showed that the density profile of clusters can differ from the NFW, see e.g. Meneghetti et al. (2014) [57] who found that a large fraction of simulated halos are better fitted by a generalized NFW (gNFW [1]), namely a steeper/shallower NFW profile modified by an additiona

[1] $\rho(r) = \frac{\rho_s}{\left(\frac{r}{r_s}\right)^{\beta}\left(1+\frac{r}{r_s}\right)^{3-\beta}}$

parameter β, and Einasto profiles (Einasto, 1965 [58]).

Cluster member galaxies are surrounded by halos of dark matter, meaning that it is important to quantify the distribution around the single galaxies given that these can serve as gravitational lenses inside the main cluster lens (see section (3.2.3)); as shown in Fig.(3.5) Hubble Space Telescope (HST) observed the lensing geometry of the Sloan Lens ACS Survey (SLACS) sample of 98 elliptical galaxies that strongly lens blue star-forming galaxies at intermediate redshift (Bolton et al., 2008 [59]), showing how the DM distribution is clearly spatially more extended than the luminous stellar component.

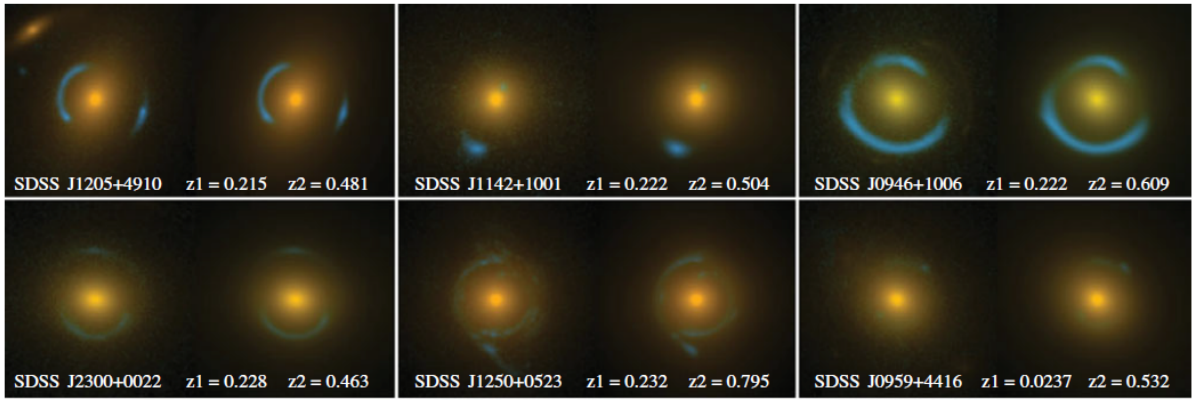

Fig. 3.5. HST images of some of the many elliptical lenses of the SLACS survey, that show the blue background galaxy deformed by the foreground red elliptical lens. Figure from Bolton et al., 2008 [59].

3.2 Galaxy clusters as lenses

3.2.1 Strong lensing

The appearance of multiple and highly distorted images, such as gravitational arcs, are common manifestation of strong lensing. In order to produce multiple images, and thus critical and caustic lines, a lens must be compact and massive enough to satisfy $\Sigma > \Sigma_{crit}$ somewhere. This condition is more easily satisfied in the center of galaxies, groups and galaxy clusters, where shear and convergence are close to unity. In order to find strong lenses nowadays the ideas are to either search for lensing features around the most probable lenses or search for highly deformed sources and then find their lenses; some examples of surveys with this purpose on galaxy scales are the CLASS (Myers et al., 2003 [60], Browne et al., 2002 [61]), HATLAS and SLACS.

Gravitational lensing surveys with galaxy clusters are mostly dedicated to study the matter distribution and evolution of cosmic structures, as well as to explore the distant and faint Universe. In the past decade several multi-cycle programs of the Hubble Space Telescope (HST) have produced detailed

observations of the center regions of many strong lensing clusters:

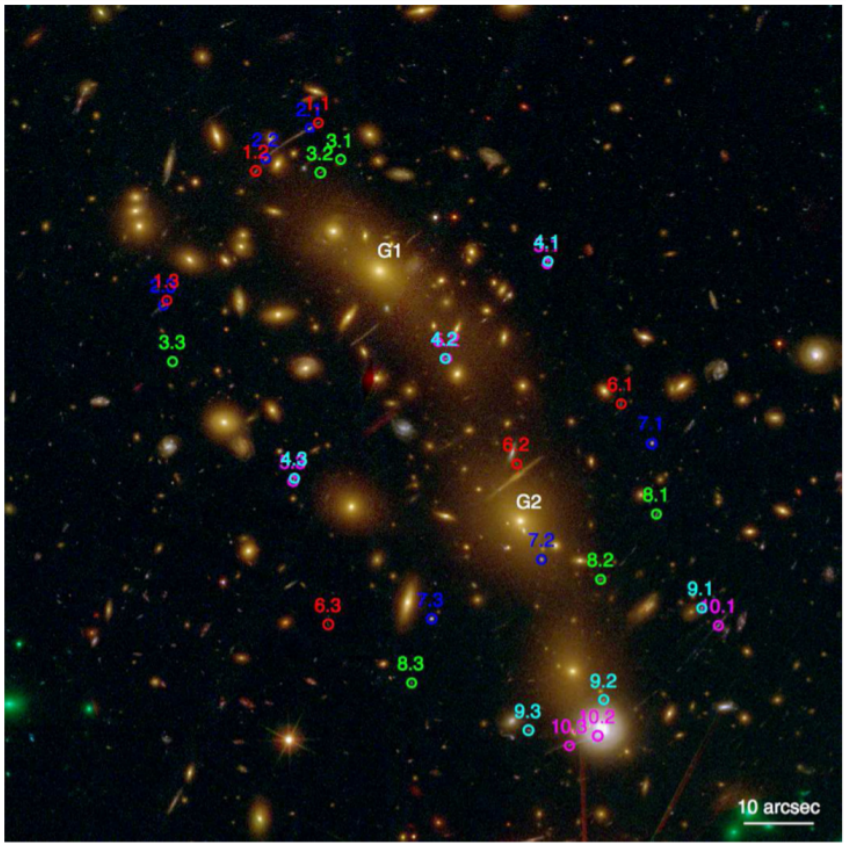

Fig. 3.6. HST image of cluster MACS J0416.1-2403, with labeled the multiple images families. Figure from Grillo et al., 2015 [62].

- Cluster Lensing and Supernova Survey with Hubble (CLASH): they mapped the distribution of dark matter in a sample of 25 rich galaxy clusters using strong lensing and studied the evolutionary and structural characteristics of the cluster member galaxies (Postman et al., 2012 [63]) ;
- Hubble Frontier-Fields (HFF): focused on 6 strong lensing clusters with HST 840 orbits, thus each object was observed with a depth comparable to the Hubble Ultra Deep Field (Ferguson H., 2000b [64]) down to to 29th magnitude in 7 bands. The goal was to combine the power of HST with the high magnification produced by clusters detect and study the distant galaxy populations (Lotz et al., 2014 [65], Koekemoer et al., 2014 [66]).
- Reionization Lensing Cluster Survey (RELICS): survey designed to deliver a large sample of bright high-redshift galaxies from the first billion years after the Big Bang, through HST and *Spitzer* observations of 41 galaxy clusters (Coe et al., 2019 [67]).

- Beyond Ultra-deep Frontier Fields and Legacy Observations (BUFFALO): will expand existing coverage of the Hubble Frontier Fields (HFF) to simultaneously find high-redshift galaxies and study the properties of dark matter and galaxy assembly (Steinhardt et al., 2020 [68]).

The datasets recovered from these programs are extensively used to produce gravitational lensing models of galaxy clusters. The observational constraints exploited to produce the mass models are mainly three: *astrometric* constraints, namely the positions of the multiple images from sources at different redshifts, *photometric* constraints, based on the measured flux of the images, and *time* constraints, from the time-delay measurements. The matter distribution can be recovered from the observed strong lensing constraints with *lens inversion* methods, that can be divided in two main classes: those following the free-form approach, where the observables are mapped onto the lens plane divided into individual cells and then linked to the lens potential or deflection field (see e.g. Merten et al. 2010 [69], Diego et al., 2005 [70], Lam et al., 2019 [71]); those following the parametric approach, where each cluster component is described by a mass component with an anylitic density profile positioned following the luminous matter and characterized by a set of parameters that need to be optimized to achieve the best reproduction of the observed positions and magnification of multiple images (see e.g. Natarajan & Kneib, 1997 [72], Sharon et al., 2012 [73]). Given the complexity of cluster lenses, the number of free parameters can be very large, hence the need to reduce them by exploiting scaling relations that can link the lensing properties of galaxy-scale mass components to easily accessible observables: for example, the galaxy luminosity can be related to the velocity dispersion by the Faber-Jackson (Faber & Jackson, 1976 [74]) scaling relation, reducing the free parameters to the normalization and reference luminosity only. A comprehensive comparison between lens modeling techniques is discussed in Meneghetti et al. (2017 [75]), based on HFF data and using synthetic lenses. These methods are further discussed in Sec.(4.1).

3.2.2 Weak lensing

In the weak lensing regime we can expect a slight distortion of background galaxies, but no multiple images or large shape deformation. Weak lensing measurements allow to trace the mass distribution of galaxy clusters at large radii from the center and the measurement of the cluster mass up to the virial radius.

In the weak lensing regime the amplitude of the deflection angle is small. In addition, the deflection angle varies on scales much larger that the size of the galaxies so that κ and γ are small and nearly constant. The lens equation can be linearised and the lens mapping is described by the lensing Jacobian. In this case a circular source is mapped into an ellipse with semi-axes linked to the local convergence

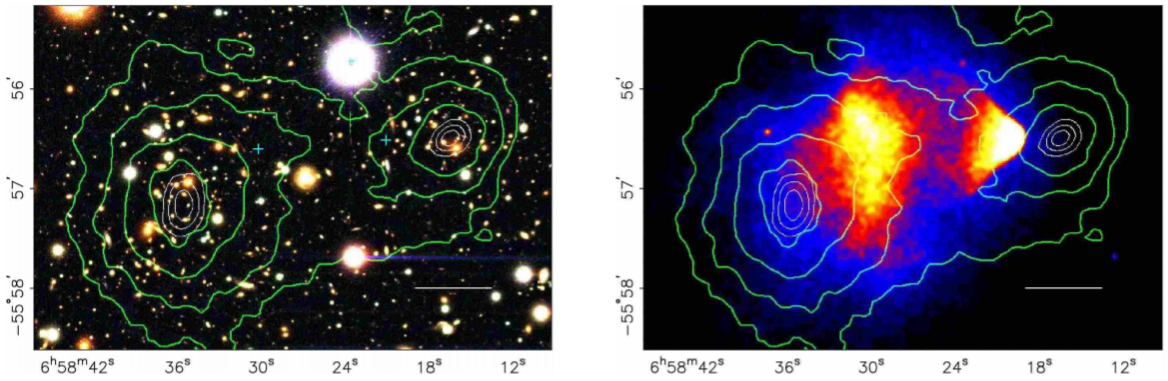

Fig. 3.7. Color image from the Magellan (left panel), and 500 ks Chandra image (right panel) of the Bullet Cluster. The white bar indicates 200 kpc at the distance of the cluster. The green contours are the weak lensing κ reconstruction, with the outer contour level at $\kappa = 0.16$ and increasing steps of 0.7. The white contours represents the 68.3%, 95.5%, 99.7% confidence leves on the position of the κ peaks. Figure from Clowe et al., 2006 [76].

and shear values. The main issue is that sources are not circular and not even perfectly elliptical, but rather irregular. To define the ellipticity of background galaxies we can use their surface brightness distribution $I(\vec{\theta})$, where $\vec{\theta}$ is the position on the lens plane. The image surface brightness centroid is

$$\vec{\theta} = \frac{\int d^2\theta q_I [I(\vec{\theta})]\vec{\theta}}{\int d^2\theta q_I [I(\vec{\theta})]}, \tag{3.6}$$

where q_I is a weight function that defines the iso-photal limit within which the surface brightness is measured. Then we define the tensor the second order brightness moments,

$$Q_{ij} = \frac{\int d^2\theta q_i [I(\vec{\theta})] \left(\theta_i - \bar{\theta}_i\right) \left(\theta_j - \bar{\theta}_j\right)}{\int d^2\theta q_i [I(\vec{\theta})]}, \quad i,j \in \{1,2\}, \tag{3.7}$$

whose trace gives the size of the image. From the element of this tensor, Q_{ij}, one can define the complex[2] ellipticity:

$$\epsilon = \epsilon_1 + i\epsilon_2 = \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22} + 2\left(Q_{11}Q_{22} - Q_{12}\right)^{1/2}} \tag{3.8}$$

Since galaxies are not circular it is necessary to separate the ellipticity introduced by lensing from the intrinsic ellipticity of the source. Hence we define the same tensor $Q^{(s)}_{ij}$ for the unlensed source in terms of the source surface brightness at $\vec{\beta}$ on the source plane. It is possible to show that

$$Q^{(s)} = \mathcal{A}Q\mathcal{A}^T = \mathcal{A}Q\mathcal{A} \tag{3.9}$$

[2] A complex number can be expressed as $a + ib$, where a and b are real numbers, and i represents the imaginary unit that satisfies $i^2 = -1$.

where $\mathcal{A}$ is the lensing Jacobian given in Eq.(2.16). Since the Jacobian can be written in terms of the complex *reduced shear* g, then we can find the link between the intrinsic and observed ellipticity:

$$\epsilon^{(\mathrm{s})} = \begin{cases} \frac{\epsilon - g}{1 - g^* \epsilon} & \text{if } |g| \leq 1 \\ \frac{1 - g\epsilon^*}{\epsilon^* - a^*} & \text{if } |g| > 1 \end{cases} \tag{3.10}$$

where the notation * denotes the complex conjugation[3]. Being the ellipticity a complex quantity, it is characterized by a module and a phase, and the phase is responsible for the orientation of the ellipse we use to describe the image. By assuming that the phases of the complex ellipticity, or equivalently the intrinsic orientations of the galaxies, are random then the expectation value of the intrinsic ellipticity is:

$$\langle \epsilon^{(s)} \rangle = 0 \longrightarrow \langle \epsilon \rangle = \begin{cases} g & \text{if } |g| \leq 1 \\ 1/g^* & \text{if } |g| > 1 \end{cases} \tag{3.11}$$

so in the weak lensing regime the mean value for the lensed galaxies in a region where κ **and** γ are constant gives the reduced shear. The noise on the shear value in a given region of the sky is determined by the intrinsic ellipticity dispersion

$$\sigma_\epsilon = \sqrt{\left\langle \epsilon^{(\mathrm{s})} \epsilon^{(\mathrm{s})^*} \right\rangle} \tag{3.12}$$

and can be reduced by averaging over many galaxy images. For a given number density of galaxies, if the sampled area is large we can average over more galaxies, at the expenses of spacial resolution. Alternatively, we can improve the measurements of the lensing signal by increasing the source number density with deeper observations. Other sources of noise are seeing or smearing effects, the instrument PSF can introduce anisotropy, and the pixels on the CCD yield a non continuous surface brightness profile.

The weak lensing allows to reconstruct the mass distribution in outskirts of galaxy clusters. Several algorithms exist to perform such reconstruction. For historical reasons we mention here the Kaiser & Squires (1993 [77]) free-form inversion method that yields an expression linking the projected mass density, or equivalently κ to the reduced shear:

$$\kappa(\vec{\theta}) = \frac{1}{\pi} \int \mathrm{d}^2\theta' \left[D_1 \left(\vec{\theta} - \vec{\theta}' \right) g_1 (1 - \kappa) + D_2 \left(\vec{\theta} - \vec{\theta}' \right) g_2 (1 - \kappa) \right] \tag{3.13}$$

where D is a filter function and this equation can be iteratively solved from $\kappa = 0$.

[3] The complex conjugate of the complex number $c = a + ib$ is given by $c^* = a - ib$.

3.2.3 Galaxy-galaxy lensing

A statistical technique similar to the weak lensing one can be used to measure the extent of dark matter halos around individual galaxies. In the weak lensing framework the dark matter halo can introduce small distortion to the shapes of background galaxies (Parker et al., 2007 [78]); we refer to this phenomenon as galaxy-galaxy lensing, firstly detected in Brainerd et al. (1996 [79]). On the other hand, if the galaxies are cluster members, their halos can introduce additional strong lensing effects on the background sources, already lensed by the overall mass of the cluster: these galaxy-galaxy strong lensing (GGSL) events manifest mostly through small-scale features, around the critical lines of the substructures. The Hubble Space Telescope cameras and resolution have allowed the identification of multiply imaged systems in galaxy clusters that can be traced back to these small-scale strong lensing events and thus offer the best way to probe the distribution and characteristics of substructures in a ΛCDM scenario. On this matter in Natarajan et al., 2017 [80] they compared the sub-halo mass function

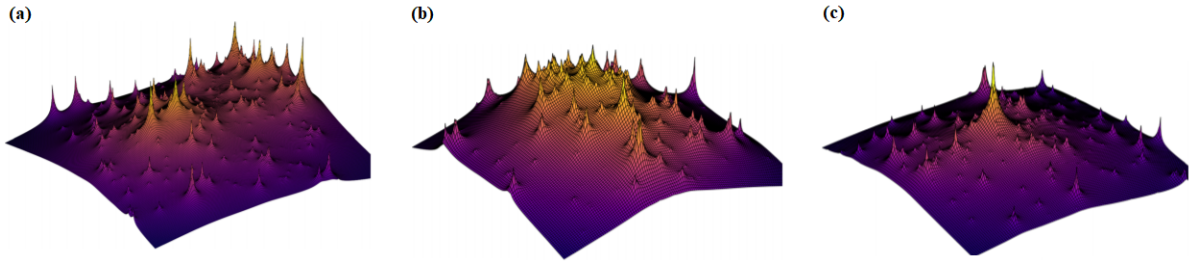

Fig. 3.8. The 3D visualization of the substructure distribution derived from HFF dataset for three clusters: (a) Abell 2744, (b) MACSJ 0416, (c) MACSJ 11449. Figure from Natarajan et al., 2017 [80].

(SHMF), obtained from the substructure distribution inferred from lensing constraints, to theoretical predictions and numerical simulations for three HFF clusters: while the simulated and observed SHMFs are in excellent agreement, the radial distribution of sub-halos is quite different, probably due to the dynamical properties and merging morphologies of the real clusters not well represented in simulations. A more recent study (Meneghetti et al., 2020 [81]) on GGSL events conducted on 11 galaxy clusters highlighted a discrepancy between the observed number of events in these clusters and the predicted one from CDM simulations; in Fig. (3.9) the observed and simulated probability of GGSL differs by more than one order of magnitude. This excess of observed events could point at incorrect assumptions on the role of dark matter in the standard cosmological model and its interaction with baryons.

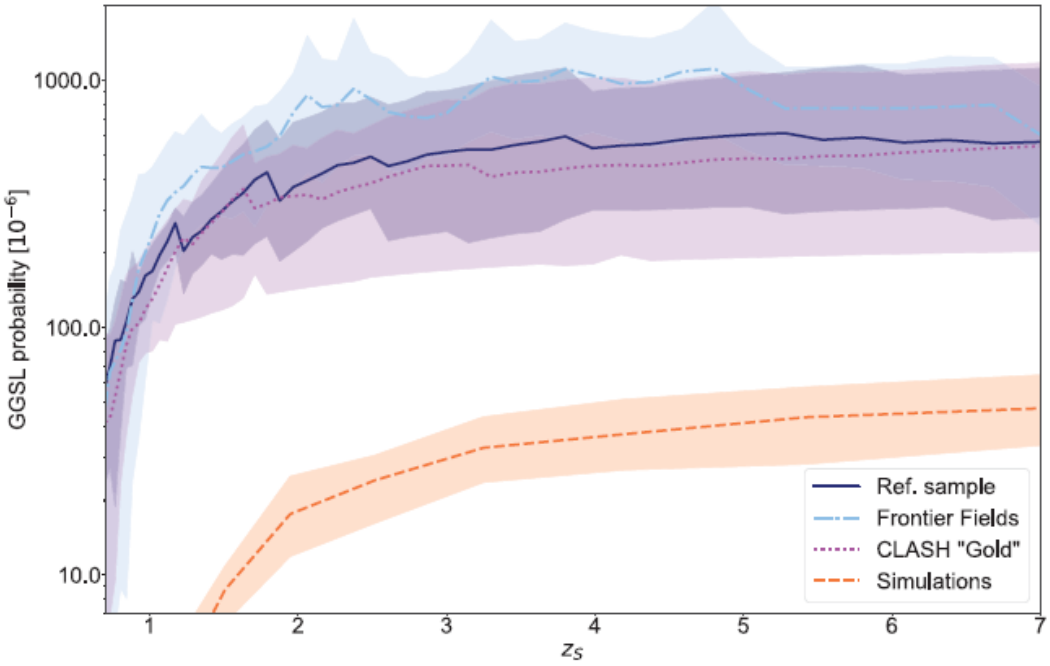

Fig. 3.9. Comparison between the mean GGSL probability for the reference sample (blue solid line), for the HFF (light blue dotted line), the CLASH (violet dotted line) samples, and simulations (orange dotted line). The difference is about one order of magnitude.

Chapter 4

Lens Modeling

4.1 Methods

Multiple images can be used to constrain the galaxy cluster mass distribution that produced them. As already briefly discussed in Sec.(3.2.1) lens modeling approaches can be distinguished in two main categories, the parametric ones and the free-form ones. Tipically, in the strong lensing regime, the parametric approach is the most used. The mass distribution is described by a set of mass clumps on different scales representing galaxies and DM halos, whose mass profiles are defined by a number of parameters (for example, the SIS mass distribution is described by three parameters, the center position *(x,y)* and the velocity dispersion). The recent high-quality observations with HST of some clusters revealed many families of multiple images usable to build detailed mass models. Furthermore this collection of positional constraints can be improved adding the information from morphology, color and redshift to map the multiple images to the same background galaxy (Broadhurst et al., 2005 [82]). The redshift-confirmed multiple image samples now available are the basis for new, more accurate cluster models.

The advantage of parametric models is that the mass distribution of a cluster can be described analytically using only few parameters. Sometimes this description of the lenses can be simplistic. In addition, it relies on strong assumptions, such that the light traces the mass. This encouraged the development of non-parametric (i.e. free-form) methods, where the description of the cluster mass distribution is described by arbitrary basis functions and relies on few or no assumptions on the DM distribution. The usual approach to these techniques is to map the mass or the potential on a grid of smaller quantity elements, the "pixels" (Diego et al. 2005a [70], Bradač et al. 2005 [83]), that can be

regular or irregular (Coe et al. 2008 [84], 2010 [85]) and then interpolate these quantities at the position of the strong or weak lensing constraints in order to minimise the χ^2 that tests how well the model matches the observations. The main strength of free-forms is the flexibility in describing more complex systems, such as merging clusters where the dark matter is not necessarily following the light. The large degree of flexibility can also be a weakness, because these methods are often prone to overfitting, without a suitable regularization (Kneib & Natarajan, 2011 [86]). Moreover, the non-parametric models often do not include the galaxy scale substructures, which are non negligible to the overall multiple image configuration.

To obtain the best from both techniques it is convenient to develop a so-called "hybrid" approach, where free-form is complimented with a mass component associated to the cluster members. In Meneghetti et al. (2017 [75]), these different methods are compared in the reconstruction of two simulated clusters: the results of this project highlighted how the parametric methods are generally more accurate in reconstructing the convergence map and in determining the mass around cluster members. Instead, the free-forms usually can reach a smaller spatial resolution. Moreover, free-forms radial profiles of the convergence and the enclosed mass measurements are as much accurate as the parametric ones.

4.2 LENSTOOL

In this work we will rely on the parametric software LENSTOOL (Kneib et al. 1996 [87], Jullo et al. 2007 [88], Jullo & Kneib 2009 [89]) to perform lens modeling. The software method is thus based on the decomposition of the examined lens mass into different mass component, each described by a set of parameters. Then the software explores the parameter space around the optimized results and produces a catalogue of expected multiple images with associated uncertainties. The current version exploits a Bayesian Markov Chain Monte Carlo (MCMC) approach.

In order to constrain the mass in a complex lens system we need to use multiple images with measured spectroscopic redshifts. Hence we need to define a likelihood function $\mathcal{L}$ for the observed positions $\mathbf{x^{obs}}$ of the images and parameters $\mathbf{p}$ of the model (Jullo et al. 2007 [88]):

$$\mathcal{L} = \Pr(\mathbf{x^{obs}} \mid \mathbf{p}) = \prod_{i=1}^{N} \frac{1}{\prod_{j=1}^{n_i} \sigma_{ij}\sqrt{2\pi}} \exp^{-\frac{\chi_i^2(\mathbf{p})}{2}} \tag{4.1}$$

where N is the number of families of multiple images, n_i is the number of multiple images from the i-th family, and σ_{ij} is the error on the position of the j-th image. The model is optimized by minimizing the $\chi^2(\mathbf{p})$ that defines the displacement between the observed and predicted ($\mathbf{x^{pred}}$) position of the images

on the lens plane:

$$\chi^2(\mathbf{p}) := \sum_{i=1}^{N} \sum_{j=1}^{n_i} \left(\frac{\left\| \mathbf{x}_{i,j}^{obs} - \mathbf{x}_{i,j}^{pred}(\mathbf{p}) \right\|}{\sigma_{i,j}} \right)^2 \tag{4.2}$$

We can write the number of degrees-of-freedom (DoF) of the model as the difference between the number of constraints (namely the positions of the total number of observed images N_{tot}) and the number of free-parameters N_{par} (Bergamini et al., 2020 [90]):

$$\text{DoF} \ = N_{\text{con}} - N_{\text{par}} = (2 \times N_{tot} - 2 \times N) - N_{\text{par}} \tag{4.3}$$

where the positions (in the (x, y) directions) of the background sources $(2 \times N)$ of the N families, being additional free parameters, are subtracted from the number of image positions. Optimizing on the lens plane means starting from the observed image positions, assuming a model to trace these images to a position on the source plane and then solve the lens equation from the source position to obtain the predicted images positions. Another way is optimizing on the source plane, which is definitely faster, since it only requires to find the set of parameters that trace the multiple images from the lens plane on a sufficiently small region on the source plane. This would require to minimize a source plane $\chi^2_S(\mathbf{p})$, analogous to (4.2). However, while the lens plane optimization may be slower, it yields more precise results. One can rapidly optimize on the source plane in order to understand the parameters range and then run on the lens plane to refine the model.

Another indicator of the goodness of the fit, besides the χ^2_{min}, is the root-mean-square (rms) between the observed and the model predicted positions of the multiple images, defined as (e.g. Caminha et al. 2016 [91], Bergamini et al. 2020 [90]):

$$\Delta_{rms} = \sqrt{\frac{1}{N_{tot}} \sum_{i=1}^{N_{tot}} \|\Delta_i\|^2}, \quad \text{with} \quad \Delta_i = \mathbf{x}_i^{obs} - \mathbf{x}_i^{pred} \tag{4.4}$$

where N_{tot} is the total number of multiple images and Δ_i represents the displacement of the i-th observed multiple image from the predicted position.

Moreover, two additional criteria can be adopted to help the selection among a set of models:

- the **Akaike information criterion** (AIC, Akaike 1974 [92]), which provides a means for model selection by evaluating the relative amount of information lost by a given model. It is defined as

$$\text{AIC} \equiv -2 \ln (\mathcal{L}_{\max}) + 2N_{\text{par}} \tag{4.5}$$

where $\mathcal{L}_{\max}$ is the maximum value of the likelihood function for the model. The criterion evaluates the goodness of fit by assessing the likelihood function, but penalize the introduction of new parameters in the model to avoid overfitting. The preferred model will be the one with the smaller AIC value.

- the **Bayesian information criterion** (BIC, Schwarz 1978 [93]), is closely related to the AIC for model selection. This criterion as well aims at discouraging overfitting by penalizing the introduction of extra parameters; is it defined as

$$\mathrm{BIC} \equiv -2\ln\left(\mathcal{L}_{\max}\right) + N_{\mathrm{par}}\ln\left(N_{\mathrm{con}}\right) \tag{4.6}$$

so in this case the penalty term is larger than in AIC. The model with the lowest value for the BIC is preferred.

4.2.1 Parametric mass profiles

The lens mass distribution can be divided into a sum of parametric mass components. For a galaxy cluster the components mostly fall in two main categories: the cluster-scale halos (including DM halos and ICM), and the galaxy-scale halos (sub-halos, including stars and DM in galaxies). Thus the gravitational potential can be writtes as

$$\phi_{tot} = \sum_i \phi_{CS_i} + \sum_j \phi_{SH_j} \tag{4.7}$$

where ϕ_{CS_i} are the cluster-scale potentials and ϕ_{SH_j} are the sub-halo potentials. It is often necessary to introduce additional potentials to describe the presence of massive structures not belonging to the cluster but nonetheless relevant to the lensing system, such as a foreground object (ϕ_{foreg}) or an extra term introducing an external shear (ϕ_{shear}).

Each halo is described by at least two parameters for the their position (x, y) in the sky, one for the projected ellipticity and one for a position angle; more parameters are specific for the potential profile used to describe the component. Usually, in the description of clusters as lenses, isothermal and PIEMD profiles are assumed: the isothermal profile, described in Sec.(2.3.3), is often not sufficient to describe a complex system, limited by the divergent mass and the central singularity in the density profile; on the other hand, the PIEMD profile, whose parameters are listed in Sec.(2.3.5), is a suitable choice for cluster-scale and galaxy-scale halos. In fact, the introduction of a truncation radius (r_{cut}) and a core radius (r_{core}) solve the limitations of the isothermal profile. In most cases for cluster-scale halos, the

truncation radius is assumed to be extremely large, far beyond the area covered by multiple images, while the core radius is a free parameter. On the contrary, for galaxy-scale halo the core radius is fixed and the truncation radius is optimized, since it is unlikely to find multiple images near the galaxy core. The galaxy-scale component is needed to achieve precise lensing models and to reproduce at best the multiple image distribution on the lens plane, therefore in LENSTOOL the number of sub-halos to include in a model needs to be quantified. Except for a set of perturbing sub-halos that can influence the formation of images, all other cluster members that increase the total mass need to be catalogued by usually selecting galaxies within the cluster red sequence brighter than a given luminosity threshold (Jullo et al., 2007 [88]). To reduce the number of sub-halo parameters we assume some scaling relations between an observable quantity (such as the luminosity L) and the parameters defining the lens mass distribution. Following the approach adopted by Brainerd et al. (1996 [79]) for modeling galaxy-galaxy lensing events, and assuming a PIEMD profile to describe the individual sub-halos, the parameters follow a scaling relation for the central velocity dispersion, the truncation radius and the core radius:

$$\begin{cases} \sigma_0 = \sigma_0^\star \left(\frac{L}{L^*}\right)^\alpha \\ r_{\rm core} \ = r_{\rm core}^* \ \left(\frac{L}{L^*}\right)^{1/2} \\ r_{cut} = r_{cut}^\star \left(\frac{L}{L^*}\right)^\beta \end{cases} \tag{4.8}$$

where L is the luminosity of the cluster member and L^* is the reference luminosity, of a galaxy at the cluster redshift. If we fix the slopes, the only free parameters to optimize are the normalizations of the velocity dispersion, and the truncation, and core radius relations corresponding to the reference luminosity.

4.2.2 Bayesian statistics

LENSTOOL latest implementations include the Bayesian MCMC package BayeSys (Skilling, 2004 [94]) to perform the lens model fitting. We refer to Keeton, C. (2016 [95]) for a brief review of the important aspects of Bayesian statistics. Given two random variables x and y, we can define three different probability distributions: the marginal distribution *p(x)* describes the values that x can assume independently on the values of y and it is obtained by integrating over y; the joint distribution *p(x,y)* describes all the possible values that x and y can have; the conditional distribution *p(x/y)* describes the values that x can assume when y assume a specific value.

In the modeling framework, we can consider the variables are the data **d** and the model parameters **m**, we can write the Bayes theorem:

$$p(\mathbf{m} \mid \mathbf{d}) = \frac{p(\mathbf{d} \mid \mathbf{m})p(\mathbf{m})}{p(\mathbf{d})} \tag{4.9}$$

where $p(\mathbf{d} \mid \mathbf{m})$ is the probability of the data given the model, hence the likelihood $\mathcal{L}$, $p(\boldsymbol{m})$ is the *prior* probability distribution, hence informations on $\mathbf{m}$ before updating with new constraints, $p(\boldsymbol{d})$ is the model *evidence*, i.e. the probability of obtaining the data for a class of models. Finally, $p(\mathbf{m} \mid \mathbf{d})$ is the *posterior* probability distribution for the model constrained by the data. During the optimization of a model, it we proceed by drawing points from the parameter space and computing the posterior distribution, using the technique called Monte Carlo Markov Chains (MCMC), which works in steps. In particular, for the largely used Metropolis-Hastings algorithm (Metropolis et al. 1953 [96], Hastings 1970 [97]) we can summarize the steps as (see e.g. Trotta 2017 [98]):

- Start from a random point θ^0 with associated posterior $p_0 \equiv p(\theta^0 \mid d)$;
- Draw a candidate point θ^c from a trial distribution $q\left(\theta^0, \theta^c\right)$;
- The candidate point is accepted only if its posterior $p_c = p\left(\theta^c \mid d\right)$ satisfies

$$P = \min\left(\frac{p_c q\left(\theta^{(c)}, \theta^{(0)}\right)}{p_0 q\left(\theta^{(0)}, \theta^{(c)}\right)}, 1\right); \tag{4.10}$$

- If accepted, it is added to the chain and the chain iterates from there. If rejected, the chain stays at the old point.

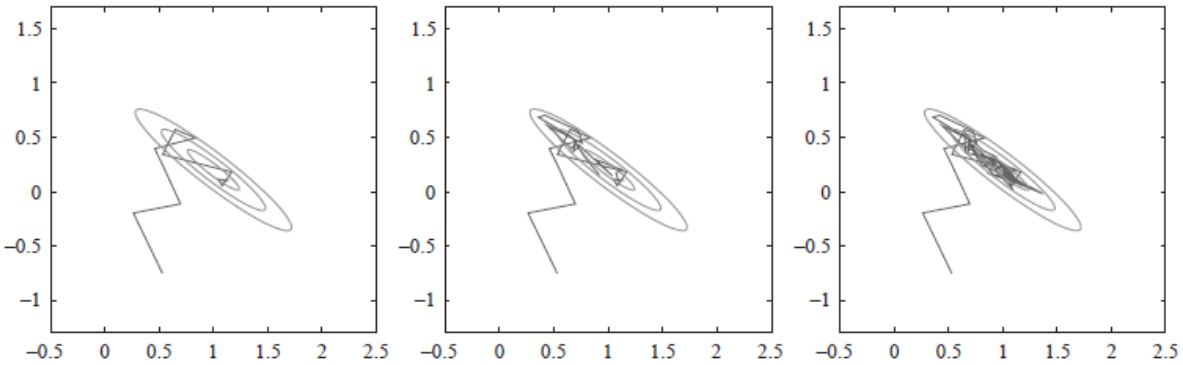

Fig. 4.1. The figure shows the superimposition of a MCMC on the contours of the likelihood function, for 50 (left panel), 150 (middle panel) and 250 (right panel) steps. Figure from C. Keeton, 2016 [95].

The Bayesian MCMC sampler is one of the possible runmodes in LENSTOOL. It explores the parameters space to sample the real shape of the posterior PDF (for flat priors, posterior and likelihood have the same shape). The set of parameters that gives the best-fit and it is consistent with the prior

PDF, have the highest posterior PDF. When starting, LENSTOOL initializes the so-called "walkers" that sample the parameters space by evaluating the posterior probability distribution function (PDF) for every point. They move towards areas with higher posterior. BayeSys run 10 interlinked MCMC so at each step the walkers draw 10 new samples from the current posterior PDF (which is the prior PDF at the beginning). The samples are then weighted on their likelihood and selected with a variant of the Metropolis-Hastings criteria. The MCMC convergence to the posterior PDF is performed with a techinque called *selective annealing* (O Ruanaidh & Fitzgerald, 1996 [99]).

In LENSTOOL are implemented different optimization methods besides BayeSys, such as the *maximum likelihood* runmode: LENSTOOL increases in time a cooling coefficient to maximize the likelihood, hence driving the walkers towards the peak in the likelihood surface. This method looks for the best sample of parameters. Both these modes have a *burn-in* phase: the walkers are randomly initialized in the parameter space, hence to help the search of the peak of the likelihood during this phase, the likelihood influence is raised progressively through a cooling factor λ, which is increased in steps from $\lambda = 0$ to $\lambda = 1$ (see Jullo et al., 2007 [88]). The burn-in phase is then eliminated to disregard the starting point of each chain.

4.2.3 Degeneracies

The posterior PDF gives the probability of the model parameters given the data, hence in order to focus on a subset of the parameters we need to obtain the marginal distribution, integrating over the other "nuisance" parameters, not of interest for the problem. The integration gives the probability distribution functions of only the parameters of interest, however possible correlations between the parameters can be hid from the marginalization. A useful way to visualize the one- and two-dimensional projections of the posterior PDF of the parameters is a corner plot: this shows, for each parameter, the marginalized two dimensional posterior distribution and the one dimensional one for each parameter independently along the diagonal. From the confidence regions in the corner plots it is easy to see the inner degeneracies presents between the model parameters involved in the computation of the mass. For example, the DM halos can balance their mass in different ways to reproduce the same multiple images configuration. We consider the case of a single PIEMD model: the position of the multiple images is dependent from the mass inside the Einstein radius. The mass is determined by the truncation radius and the velocity dispersion, where we can see the degeneracy. The definition of the PIEMD potential leads to the expression of the projected mass enclosed in an aperture of radius R. Rewriting Eq.(2.57) as in Limousin et al. (2005 [24]), we obtain:

$$M(<R) = \frac{\pi r_{cut}\sigma_0^2}{G}\left(1 - \frac{\sqrt{r_{cut}^2+R^2} - \sqrt{r_{core}^2+R^2}}{r_{cut}-r_{core}}\right) \tag{4.11}$$

For a vanishing core $r_{core} = 0$ the enclosed mass becomes

$$M(<R) = \frac{\pi\sigma_0^2}{G}\left(R + r_{cut} - \sqrt{r_{cut}^2+R^2}\right) \tag{4.12}$$

which shows how the same value for the enclosed mass can be recovered from different combinations of r_{cut} and σ_0.

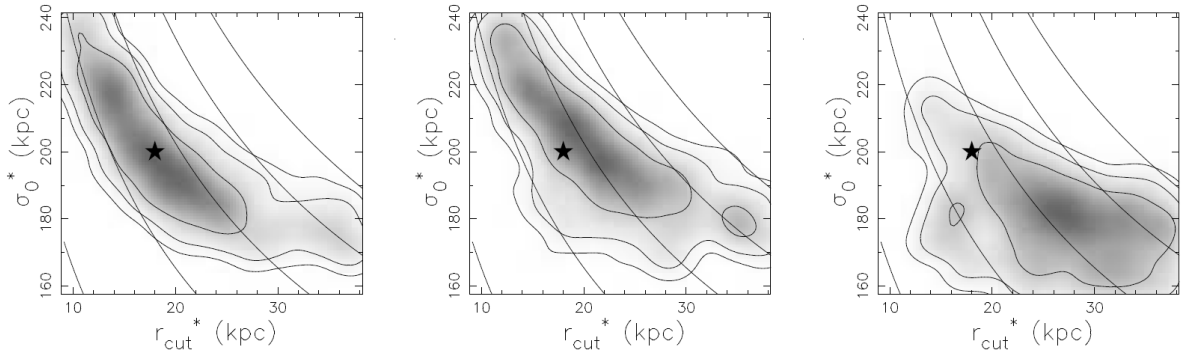

Fig. 4.2. Confidence regions that shows the two-dimensional marginalized posterior PDFs of a cluster-scale halo described with a PIEMD profile, for three different multiple image configurations. This highlights the strong degeneracy existing between the normalization parameters of the $\sigma - L$ and $r_{cut} - L$ relations The solid lines are the isodensity contours. Figure from Jullo et al., 2007 [88].

Chapter 5

PSZ1-G311 galaxy cluster

In this chapter we present the construction of the lens model for the galaxy cluster PSZ1 G311.65-18.48 (hereafter, PSZ1-G311), a peculiar object for its geometry and symmetry that produces an exceptionally extended arc magnifying a distant background galaxy. Such system configuration is worth exploring, being the arc the result of a rare almost perfect alignment between us, the cluster lens, and the distant source. Such alignment results in a very high magnification, unveiling details in the source that would be visible otherwise. The discovery of a transient object in the arc triggered the need for building a lens model, on whose accuracy depends the predictions of the next transient appearance. Moreover, the observation of a lensed transient object can be exploited to make independent measurements of cosmological parameters. In the first and second section of this chapter we introduce the object and the studies already conducted on his properties. In the third section we describe the steps for the construction of the lens model of PSZ1-G311, hence the available data-sets, the multiple images and cluster members selection and the parametrization of the cluster- and galaxy-scale halo components. Then we discuss the results by comparing different lens models reconstruction. Finally, in the fourth section, we draw the conclusions on the chosen reference model.

5.1 Detection and follow-ups

PSZ1-G311 was first detected in the ESA Planck survey and catalogued as a Sunyaev-Zel'dovich source (PSZ) (Planck Collaboration XXIX 2014 [100]). One of the ways to select a cluster is through the inverse Compton (IC) scattering of the Cosmic Microwave Background (CMB) photons by the high energy electrons in the intracluster medium, the Sunyaev-Zel'dovich (SZ) effect (Sunyaev & Zel'dovich, 1972 [101]; Sunyaev & Zel'dovich, 1980 [102]). The frequency shift caused by the IC scattering results

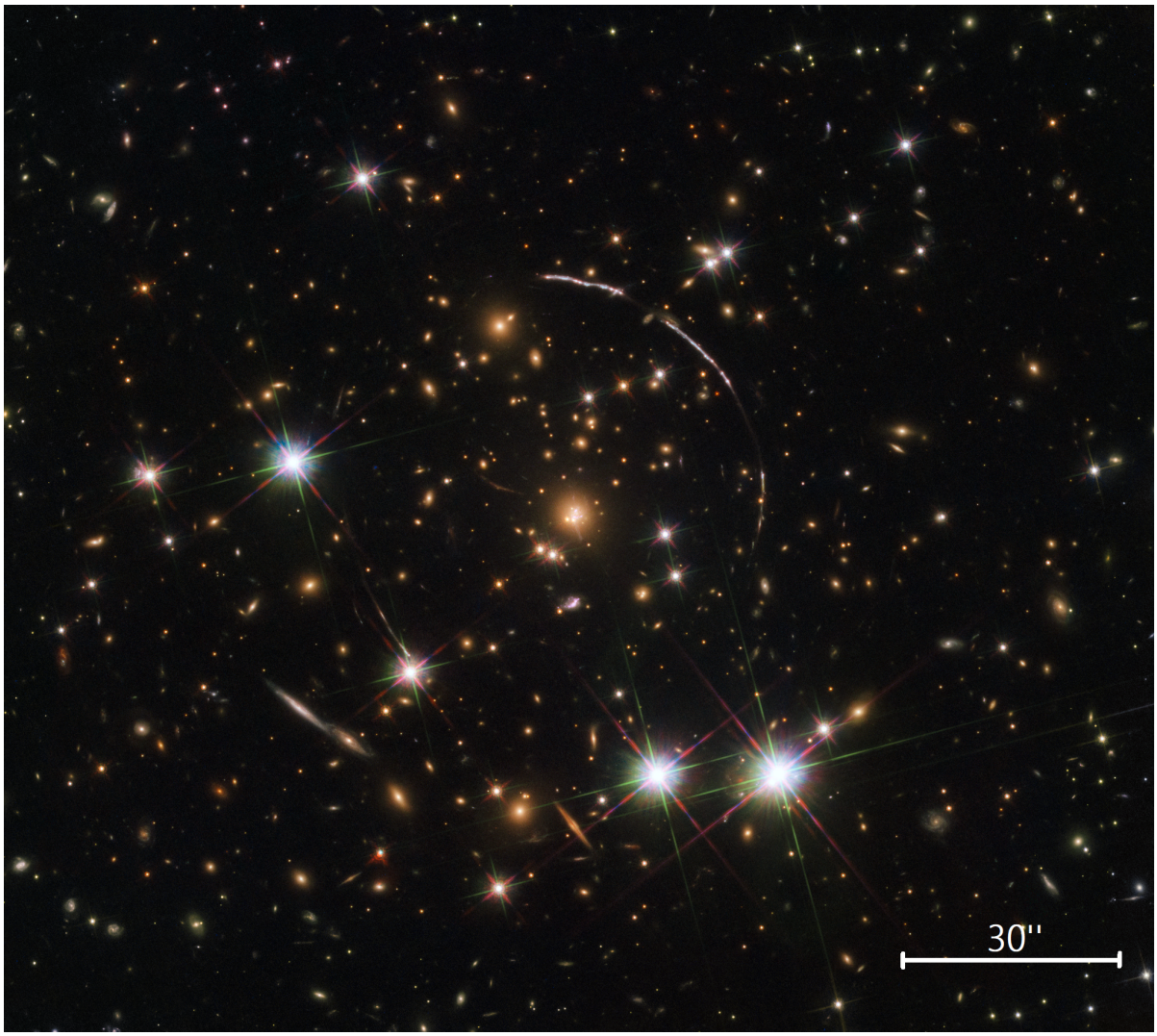

Fig. 5.1. PSZ1-G311 color composite image with HST observations.

in a decrement in the CMB intensity in the Rayleigh-Jeans part of the blackbody spectrum and an increment in the Wein tail, in particular this shift is uniquely identified at the frequency of 218 GHz. A galaxy cluster can thus be identified as an area where the CMB temperature appears higher than 2.7K. The SZ cluster PSZ1-G311 is part of the PSZ1 catalogue (Planck Collaboration XXIX 2014 [100]), detected at $S/N = 4.64$ close to the PSZ1 signal-to-noise threshold of $S/N = 4.5$. PSZ1-G311 is located in the southern hemisphere at coordinates $(\mathrm{RA}, \mathrm{Dec}) = (237.5291667d, -78.19222d)$ in a region of medium Galactic extinction but relatively high foreground star density (Dahle et al., 2016 [103]).

5.1.1 Optical observations

The Planck Collaboration undertook extensive follow-up programmes to confirm the SZ clusters detection with optical telescopes such as the Russian-Turkish 1.5m (RTT150) telescope and the Canary Islands observatories (Planck collaboration XXVI 2015 [104], Planck collaboration XXXVI 2016 [105]), with the aim of determining the redshifts of all PSZ1 sources. For the first time Dahle et al. 2016 [103] reported the discovery of the exceptional arc in the field of PSZ1-G311 as a result of inspection of the ESO New Technology Telescope (NTT) imaging data from an ongoing follow up programme on 120 PSZ1 cluster candidates (Aghanim et al., in prep.).

We refer to Dahle et al. (2016 [103]) for more details on the data analysis that follows. The instrument used for optical observations was the EFOSC2 at the NTT. In particular, they conducted 3×400s expositions in the Bessel R (central wavelength 6431 Å) and Gunn z (central wavelength 8400 Å) filter: the images show the Brightest Cluster Galaxy (BCG) with an offset of $36''$ from the Planck detection, and the arc feature north-west of the BCG (denoted as PSZ1-ARCG311.6602-18.4624 in Dahle et. al, 2016 [103], and nicknamed "Sunburst Arc" in Rivera-Thorsen et al., 2017 [106]). To measure the redshift of the BCG, which is then adopted as the cluster redshift, they took a long-slit spectrum (with spectral resolution R $\simeq$ 700) with 2×750s exposures with the IMACS instrument at the Magellan-I Baade Telescope: the slit covered simultaneously the BCG and the brighest R-band knot in the arc. The BCG spectrum showed strong [OII] doublet emission at $\lambda\lambda = 3727, 3729$ Å and absorption lines typical of ETGs, yielding a spectroscopic redshift for the cluster of $z = 0.44316 \pm 0.00035$. More details on optical follow-ups of the Sunburst Arc given in Sec.(5.2) and Sec.(5.2.1).

5.1.2 SZ mass estimate

The amplitude of the SZ effect is quantified by the Compton parameter $y(\theta)$. If l is the distance along the line of sight, we can define the Compton parameter as (see e.g. Morandi et al. 2007 [107], Rodriguez-Gonzalvez et al. 2017 [108]):

$$y(s) = \frac{\sigma_T}{m_e c^2} \int P_e(r) dl \tag{5.1}$$

where s is the projected distance from the cluster center such that $r^2 = s^2 + l^2$, $P_e(r) \equiv n_e(r) k T_e(r)$ is the ICM electron pressure at radius r, m_e is the electron mass, and σ_T is the Thomson scattering cross-section. If we integrate the signal over the solid angle Ω, we obtain the integrated-y parameter Y

as

$$Y \equiv \int_{\Omega} y(s) d\Omega \propto \frac{M_g T_e}{D_A^2(z)}, \tag{5.2}$$

which is proportional to the total gas mass. However, to define a proper scaling relation we need to introduce a distance-independent mass proxy. We can use the intrinsic integrated Compton parameter Υ, defined as

$$\Upsilon \equiv D_{\rm A}^2(z) Y. \tag{5.3}$$

The Planck SZ data revealed a degeneracy between the cluster size (θ) and the SZ flux (i.e. the integrated Compton parameter): breaking the degeneracy to use the flux measure as mass proxy would require fixing the cluster size. We define R_{500} as the cluster radius enclosing a density 500 times the critical density of the Universe at the cluster redshift. Then we denote θ_{500} as the cluster size and Y_{500} as the integrated Compton parameter within R_{500}. The size-flux degeneracy can be broken using the $M_{500} - \Upsilon_{500}$ relation: since Υ is defined as in Eq.(5.3), this offers a relationship between the θ_{500} and Y_{500} when the redshift is known. Intersecting the degeneracy curves of the size-flux relation and $M_{500} - \Upsilon_{500}$ relation we obtain a value for Y_{500}. The derived Y_{500} from the intersection is called Y_z and it is equivalent to the X-ray mass proxy. Hence it provides an estimate of the SZ-mass X-ray calibrated. Using the measured redshift and the Planck degeneracy curves, Dahle et al. (2006 [103]) obtained the mass estimate

$$M_{500\rm SZ} = 6.6^{+0.9}_{-1.0} \times 10^{14} M_{\odot}. \tag{5.4}$$

5.2 The Sunburst Arc

The first observations in the R- and z-band showed a very long gravitational arc north-west of the BCG, which appears composed of three segments. An additional arc is observed south-est of the BCG. All these arc segments are images of the same source. They show multiple bright knots, interpreted as sites of star formation in the lensed galaxy. The arc R-band magnitude is ~ 18, measured running `SExtractor` (Bertin & Arnouts 1996 [109]) on stacked R- and z-band images. The infrared imaging was obtained with one 2100s exposure in the J-band (reference wavelength 12361.81 Å) and one 800s exposure in the K_s band (reference wavelength 21414.86 Å) with the FourStar instrument on the Magellan-I Baade Telescope, yielding total magnitudes of $J \sim 17$ and $K_s \sim 15$ with `SExtractor`. Using the same long-slit observations discussed above, Dahle et al. (2016 [103]) obtained the arc spectrum in the wavelength range between 4000 Å and 6500 Å: the observed lines are reported in Tab.(5.1), from which they inferred a redshift of $z = 2.3686 \pm 0.0006$ for the lensed source.

Lines	Rest-frame wavelength [Å]	Observed wavelength [Å]
Lyα	1215.67	4095.59
SiII	1259	4241.57
SiIII	1294	4359.48
C IV	1548	5215.2
C IV	1550	5221.95
He II	1640	5525.16
OIII]	1666	5612.75
SiIII]	1892	6374
CIII]	1907	6424.68
CIII]	1909	6431.42

Tab. 5.1. Emission and absorption lines revealed in the Sunburst arc spectrum, rest-frame wavelength and corresponding observed wavelength at the galaxy redshift.

The observed spectrum, in Fig. (5.2), shows strong Lyα emission, low-ionization absorption lines from neutral outflowing gas, and high-ionization lines with P-Cygni profile (CIV doublet) (Robinson K., 2007 [110]) from stellar winds: typical features of starbursts that results in strong outflows in the interstellar medium (ISM).

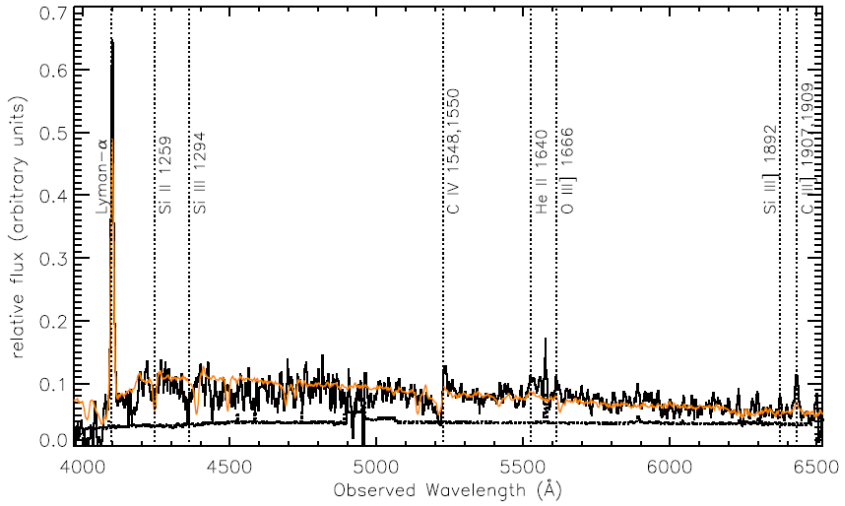

Fig. 5.2. Observed spectrum of the Sunburst Arc from IMACS at the Magellan-I Baade Telescope. Figure from Dahle et al., 2016 [103].

5.2.1 Ly-C knot

The Sunburst Arc galaxy can be classified as a Lyα emitter (LAE) from the strength of the measured rest-frame equivalent width of the line. In this framework, it is important to define the the phenomenon of *reionization*, an epoch in the Universe evolution that started less than a billion years after the Big Bang. Between $z \sim 14$ and $z \sim 6$, the hydrogen in the intergalactic medium, which was neutral since

the recombination, is "reionized" due to ultraviolet radiation from the first luminous objects (see e.g. Fan et al., 2006 [111]). In particular, this radiation corresponds to photons with wavelengths shorter than 912 Å, that contribute to the Lyman continuum (LyC). To escape into the intergalactic medium (IGM) and be observed, the LyC photons need to avoid the absorption from neutral hydrogen within the host galaxy. There are numerous ways for these photons to escape from galaxies, mostly allowed by radiative or mechanical feedback from young stars that can carve a path into the optically thick IGM ionizing most of the gas (see e.g. Zackrisson et al. 2013 [112], Bik et al., 2018 [113]). The kinematics and geometry of the escape scenario is reflected in the spectral shape of the Lyα line at rest-frame wavelength of 1216 Å. Specifically, escape from an optically thin medium yield a narrow double-peaked Lyα, while the escape from neutral gas-free channels carved in an optically thick medium yields a narrow and bright single line. Finally, if there are few narrow channels, the Lyα most probably will be trapped in resonant scattering the dense neutral medium, producing a triple-peaked profile with a narrow brighter peak at the line center over the typical broader double-peaked profile (Rivera-Thorsen et al., 2017 [106]).

This last feature was observed in the spectrum of the Sunburst Arc, thus indicating the presence of

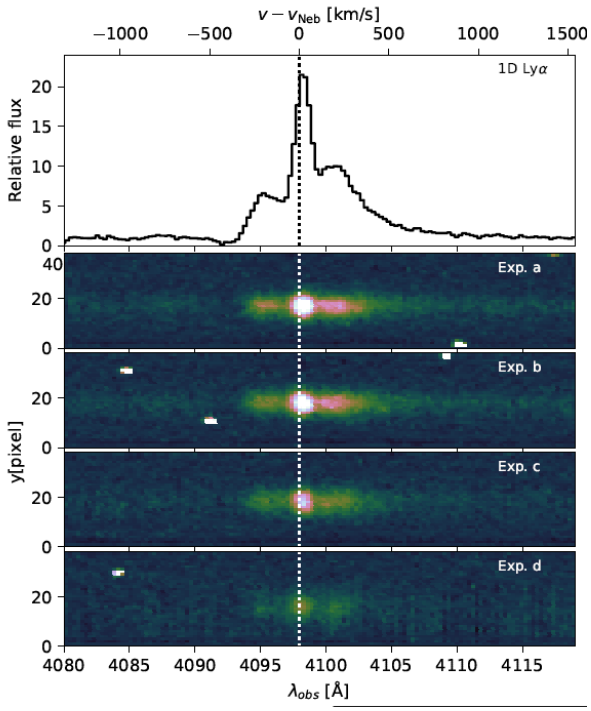

Fig. 5.3. The top panel shows the 1D profile of the Lyα in the LyC knot of the Sunburst Arc. Note the triple peaked shape, caused by espace through a clear channel in a optically thick neutral medium. The lower panels represents the 2D Lyα profiles of the four MagE exposures. Figure from Rivera-Thorsen et al., 2017 [106].

Ly-C leakage from one of the brightest knots.

Rivera-Thorsen et al. (2019 [114]), later presented Hubble Space Telescope (HST) observations of the Sunburst Arc, showing that the Ly-C radiation leakes from 12 multiple images of the same knot: the Lyα profile is triple-peaked, as shown in the spectrum in Fig.(5.3) obtained with the Magellan Echellette (MagE) spectrograph on the Magellan-I telescope. The arc was observed in the UVIS channel of HST Wide Field Camera 3 (WFC3) in the F275W filter, whose wavelength cut-off matches the lowest-energy limit of the LyC (3072 Åat the galaxy redshift) , for a total exposure time of 9422s. It was also observed with the HST Advanced Camera for Surveys (ACS) in the filter F814W, sensitive to near-UV non-ionizing radiation not absorbed by the neutral hydrogen, with eight exposures of total time 5280s.

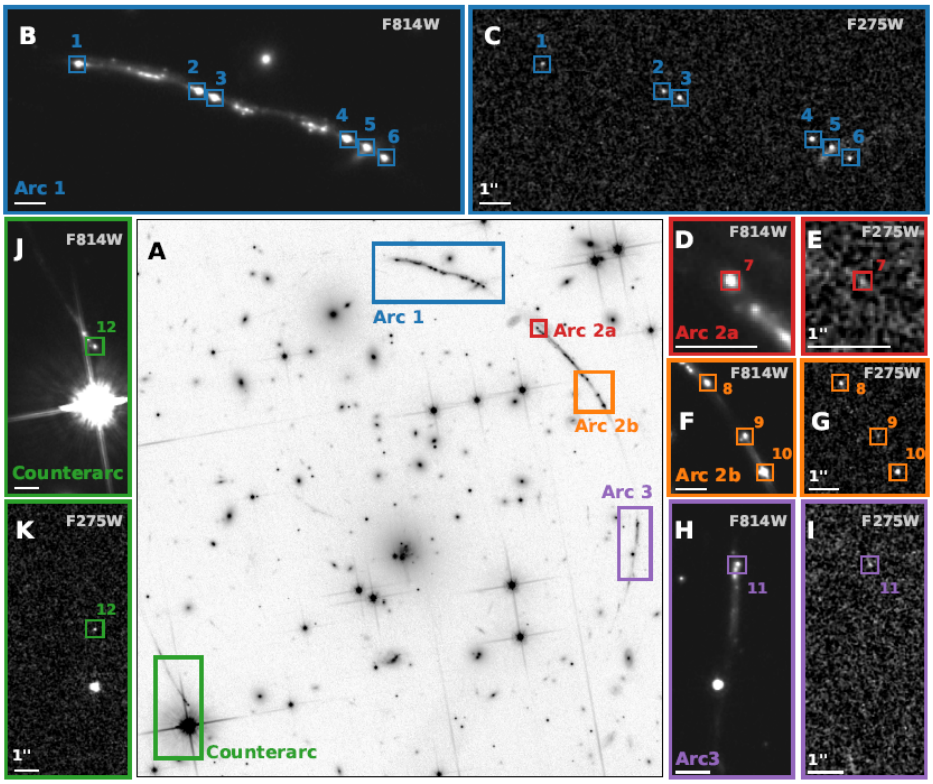

Fig. 5.4. The image shows an overview of the Sunburst Arc, with different cutouts to further highlight the morphology of the galaxy. The bright LyC knot is identified 12 times in all the four subsegments of the arc. The F814W filter is more sensitive to the ionizing photons, while the F814W filter allow a comparison with the non-ionizing near-UV radiation coming from the same stars. Figure from Rivera-Thorsen et al., 2019 [114].

Vanzella et al. (2019 [115]) further investigated the LyC continuum emission from the brightest knot, presenting VLT/MUSE observations conducted with 3×1483s exposures. They argue that this knot could be a gravitationally bound system such as a star cluster, hence they address the determination of quantities as the stellar mass and physical size of the emitting region. The stellar mass is estimated

to be ranging between $10^6 - 10^7 M_\odot$ depending on the chosen initial stellar mass function (IMF) and the magnification. The effective radius of the region is then smaller than 20pc if $\mu_{tot} > 25$, and smaller than 9pc if $\mu_{tot} > 50$.

5.2.2 Transient

Vanzella et al. (2020 [1]) discovered a possible transient stellar object between a couple of multiple images of Ly-C knot, using the VLT/MUSE observations. Then, they took additional spectroscopic observations with VLT/X-Shooter at resolution R=11400 and VLT/ESPRESSO at R=70000. The pe-

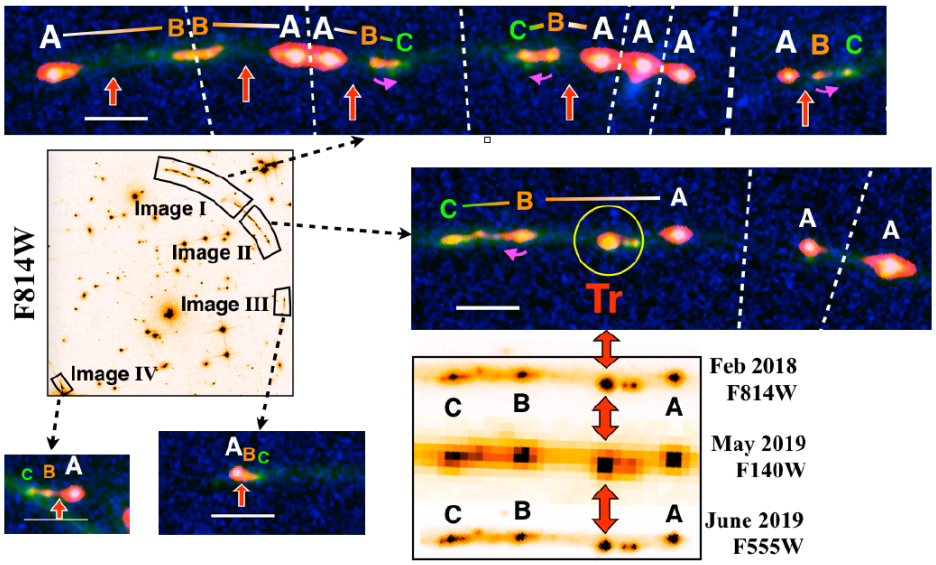

Fig. 5.5. Cutouts of the four multiple arcs are shown in RGB color images (R=F160W, G=F814W, B=F606W) from HST. Highlited the position of the transient in a yellow circle and its expected position in the other families of multiple images with a red arrow. Figure from Vanzella et al., 2020 [1].

culiar feature that led to the identification of this object is the Bowen fluorescence (Bowen, 1934 [116]). This emission mechanism is characterized by photoexcitation by accidental resonance, namely the strong emission at a given wavelength can excite different chemical elements that have a transition close to this wavelength and when they return to the ground state, they emit a group of fluorescing lines. In Fig.(5.5) we show the four segments of the arc, where the knots A,B,C are identified as the images of three different star-forming knots in the host galaxy. The transient observed and expected positions are indicated with a yellow circle and arrows, respectively. One characteristic corroborating the transient hypothesis is the appearance of the object only in one segment of the arc, in between knots A and B; its magnitude ($mag_{F814W} \simeq 22$) would imply its detection 8 times between the multiple images of the same knots in the other arc segments. The detailed analysis of the transient spectrum lines and its magnitude,

led to the hypothesis that it could be a luminous supernova (SN) explosion. In fact, fluorescence lines have been already detected in the spectra of type II SNe with strong interaction between the ejecta and the surrounding gas, idea further supported by the long-lasting light curve of the transient. This could also point to a giant eruption of a massive star. However, the lack of Balmer lines Hα and Hβ and HeI questions the hypothesis of a SN explosion in an hydrogen-rich medium. An alternative explanation could be the observation of an H-poor interacting SN, or an extremely stripped core-collapse SN powered by the interaction between the ejecta and the surrounding medium without showing H or HeI.
The importance of lensed transient lies in the possibility of exploiting measures of time-delays between their multiple images to provide an independent measure of the Hubble constant H_0 and other cosmological parameters such as Ω_M (see e.g. Refsdal 1964 [117], Grillo et al., 2015 [62]). For this purpose it is fundamental to develop a lens model to predict the reappearance of the transient in other positions in the arcs.

5.3 Lens model

5.3.1 HST and MUSE data-sets

This work is based on the HST observations of PSZ1-G311 conducted in different programs between February 2018 and June 2019. The cluster was observed in the UVIS channel of HST WFC3 in the filters F555W, F275W (program ID 15101 and 15418 respectively, PI: H. Dahle) and F606W (program ID 15377, PI: M. Bayliss); in the IR channel of HST WFC3 in the filters F140W, F105W (program ID 15101, PI: H. Dahle) and F160W (program ID 15377, PI: M. Bayliss); finally, with the HST ACS using the filters F275W and F814W (program ID 15418 and 15101 respectively, PI: H. Dahle). The lens model based on the multiple images constraints is further strengthened by VLT/MUSE spectroscopic observations, that provided a set of velocity dispersion measurements: MUSE data-cubes have a field-of-view of 1 arcmin^2 with a spatial sampling of 0.2"; the wavelength range covers from 4700 Å to 9350 Å with a dispersion of 1.25 Å/pix, and a spectral resolution of ∼2.6 Å, rather constant across the entire spectral range. A total integration time of 1.2 hours with was used to acquire the data during May-August 2016 with seeing 0.5″-0.8″, based on the DDT programme 297.A-5012(A) (PI. Aghanim) and first presented in Vanzella et al. (2020 [115]). We show how we can identify the multiple images through MUSE observations in Fig.(5.6).

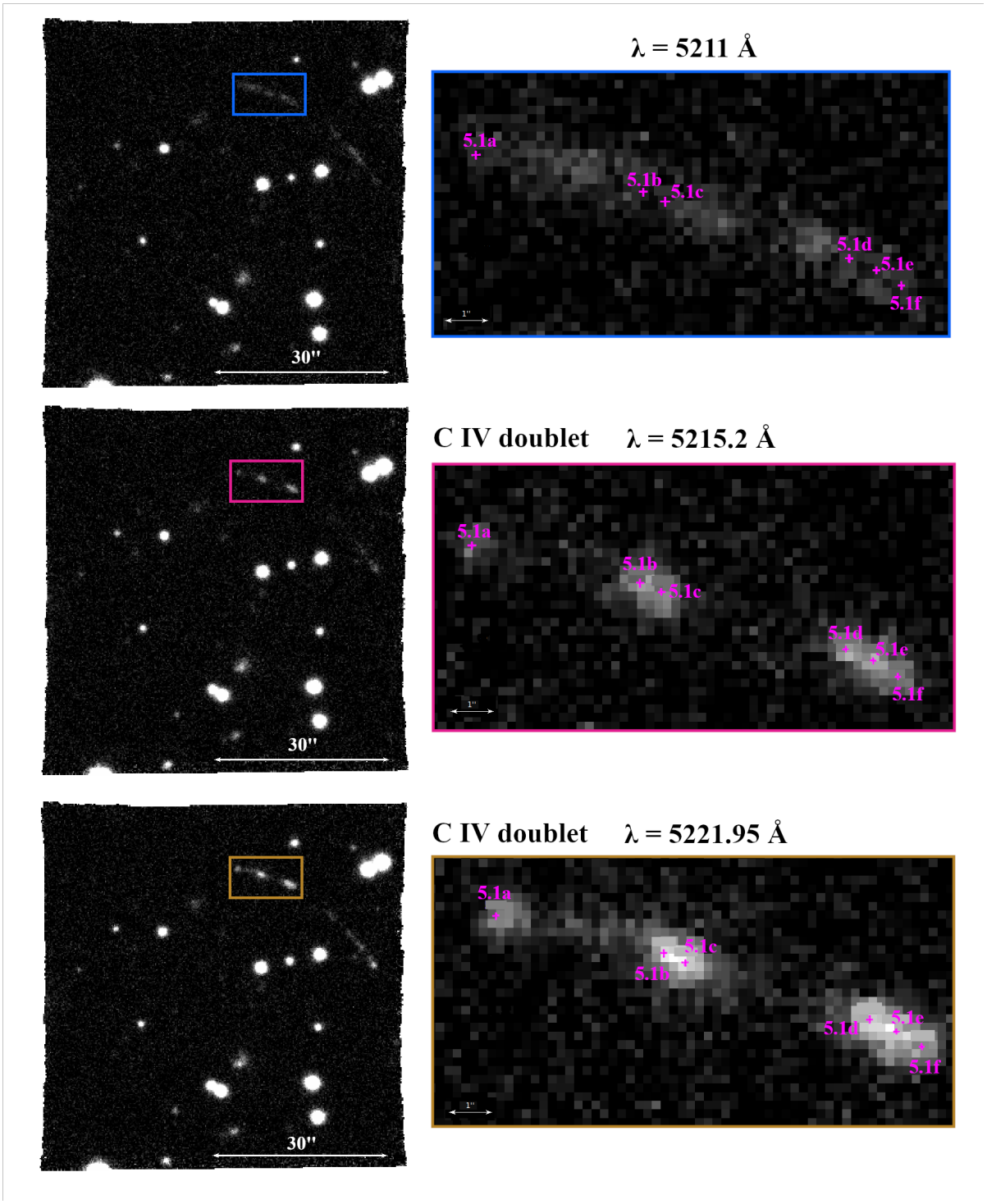

Fig. 5.6. MUSE observations in three different wavelengths. In the wavelength $\lambda = 5211$ Å(blue cut-out) there is no significant emission from the knots, as opposed to the wavelengths $\lambda\lambda = 5215.2, 5221.95$ Å, where we can see the C IV doublet emission. This allows the matching of different knots by recognizing the same emission features.

5.3.2 Cluster members selection

One of the main ingredients for the construction of a lens model is the selection of cluster galaxies. This is essentially based on the identification of the cluster red-sequence: in fact, cluster galaxies are typically early type galaxies with similar colors, independently of their magnitude (de Vaucouleurs 1961 [118], Holden et al. 2004 [119], Lidman et al. 2008 [120]). For this purpose we used `Source Extractor` (SExstractor, Bertin & Arnouts 1996 [109]), a software for the automated detection and photometric analysis of sources in a given FITS (Flexible Image Transport System) file. `SExtractor` works in step regulated by an input configuration file containing parameters defining image and background informations, detection and photometry limits and outputs. After scanning the entire image, `SExtractor` measures the background and subtracts it to isolate the luminous objects over the given detection threshold; then it performs shape and positions measurements and photometry, all informations listed in the output final catalogues. To enhance the detection process, `SExtractor` allows a runmode based on two

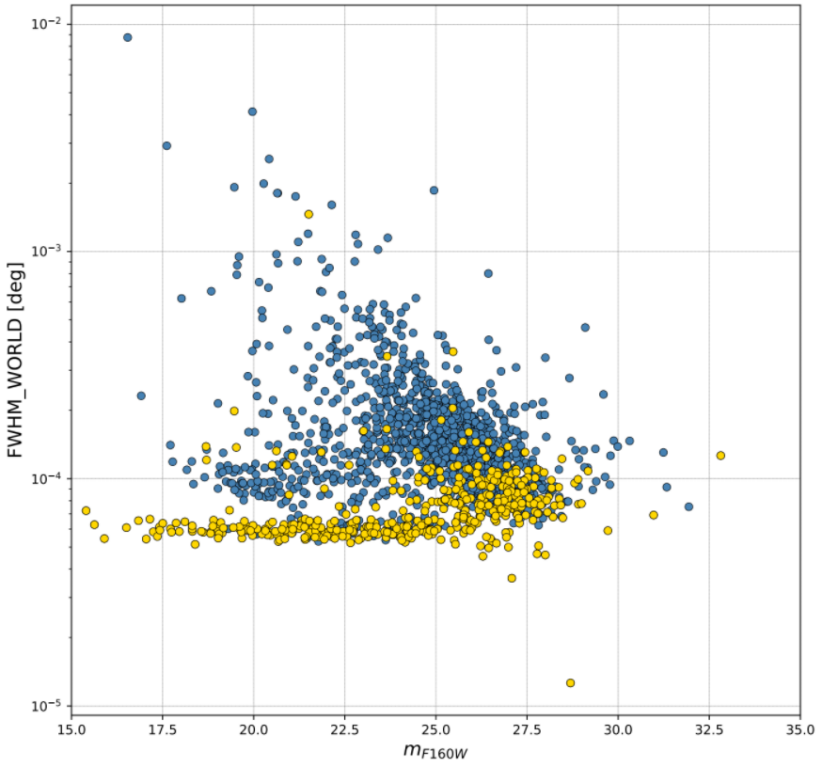

Fig. 5.7. Plot relating the dimension of objects (`FWHM_WORLD`) identified with `SExtractor` and their magnitude in the filter F160W; the yellow points are the stars, approximately of the same dimension indipendently on their magnitude.

input images, one usually is a deeper observation used as reference for the detection, and one is in the filter of interest to perform the photometry. We applied the analysis on the HST observations, taking

a stacked IR image as the detection reference for PSZ1-G311 while the photometry was performed on the F160W, F814W and F606WU filters.

The output parameters used to delimit the cluster red-sequence are essentially the magnitude and the object classification. The magnitude is measured from a flexible elliptical aperture around every detected object (`MAG_AUTO` parameter), as described in Kron (1980, [121]). The object classification is described by the parameter `CLASS_STAR`, which can assume a value between 0 and 1, being 0 an extended object and 1 a star. As shown in Fig.(5.7) the point-like objects, independently of their magnitude, have roughly the same dimension, defined as the full-width at half maximum (FWHM) when assuming a Gaussian profile for the object (`FWHM_WORLD` parameter). To avoid contamination, all objects attributed with a `CLASS_STAR` parameter higher than 0.1 are excluded from the final selection. The last step is constructing a color-magnitude diagram, as shown in Fig.(5.8), of the remaining extended objects in the field: we can see a clear sequence in the redder colors, which we can enclose in a range of $0.7 < m_{F814W} - m_{F160W} < 2.0$. Adopting these selection criterion the cluster members catalogue contains 273 objects brighter than $m_{F160W} = 24$. Following a visual inspection of this first catalogue, 73 objects are excluded because possible stars or foreground sources. Further spectroscopic analysis

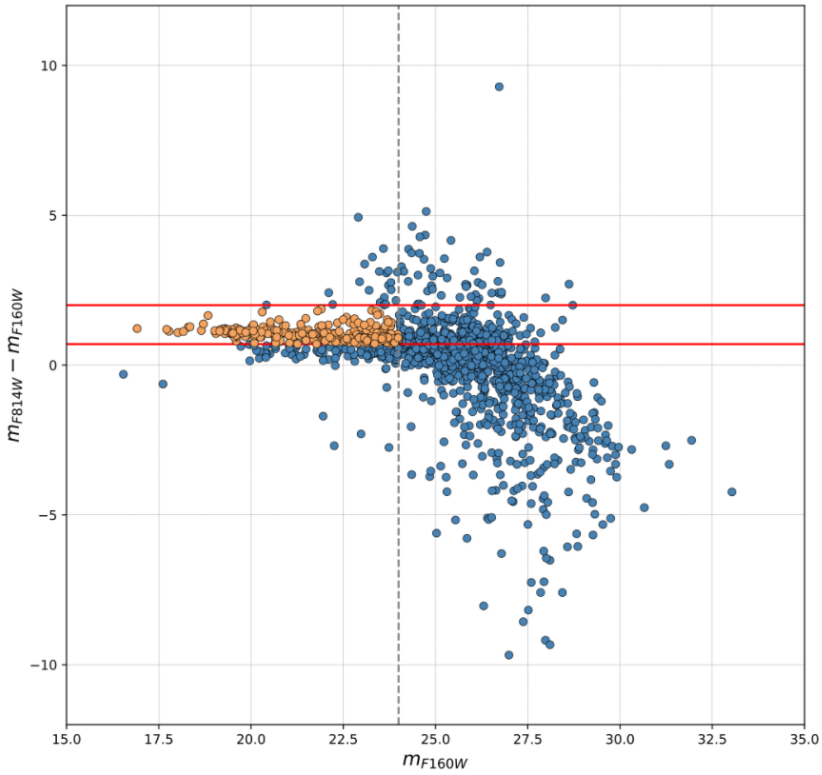

Fig. 5.8. Color-magnitude diagram for PSZ1-G311 cluster members: the orange points indicate the red-sequence, delimited by the red horizontal lines at $m_{F814W} - m_{F160W} = 0.7$ and $m_{F814W} - m_{F160W} = 2.0$ and the grey dotted vertical line at $m_{F160W} = 24$.

is conducted on this catalogue, obtaining the redshifts for 54 objects. Of those, we considered only the spectroscopically identified members within 3000km/s from the median cluster redshift, z=0.4436; this second catalogue is concatenated with the photometric one, obtaining a final selection of 205 total cluster members, including the BCG. Of these, 155 are photometrically identified without a redshift, 46 are spectroscopically confirmed with an associated magnitude value and 4 without one. As described in Sec.(4.2.1), the sub-halo mass component in LENSTOOL is described both by a small set of small-scale mass components whose parameters are optimized individually, and by all the other cluster members which are generally modeled adopting scaling relations to reduce the number of free parameters. In this work the catalogue of cluster galaxies for LENSTOOL contains 200 galaxies, indicated in Fig.(5.9) with red circles, excluding the 4 without a magnitude measurements and the BCG.

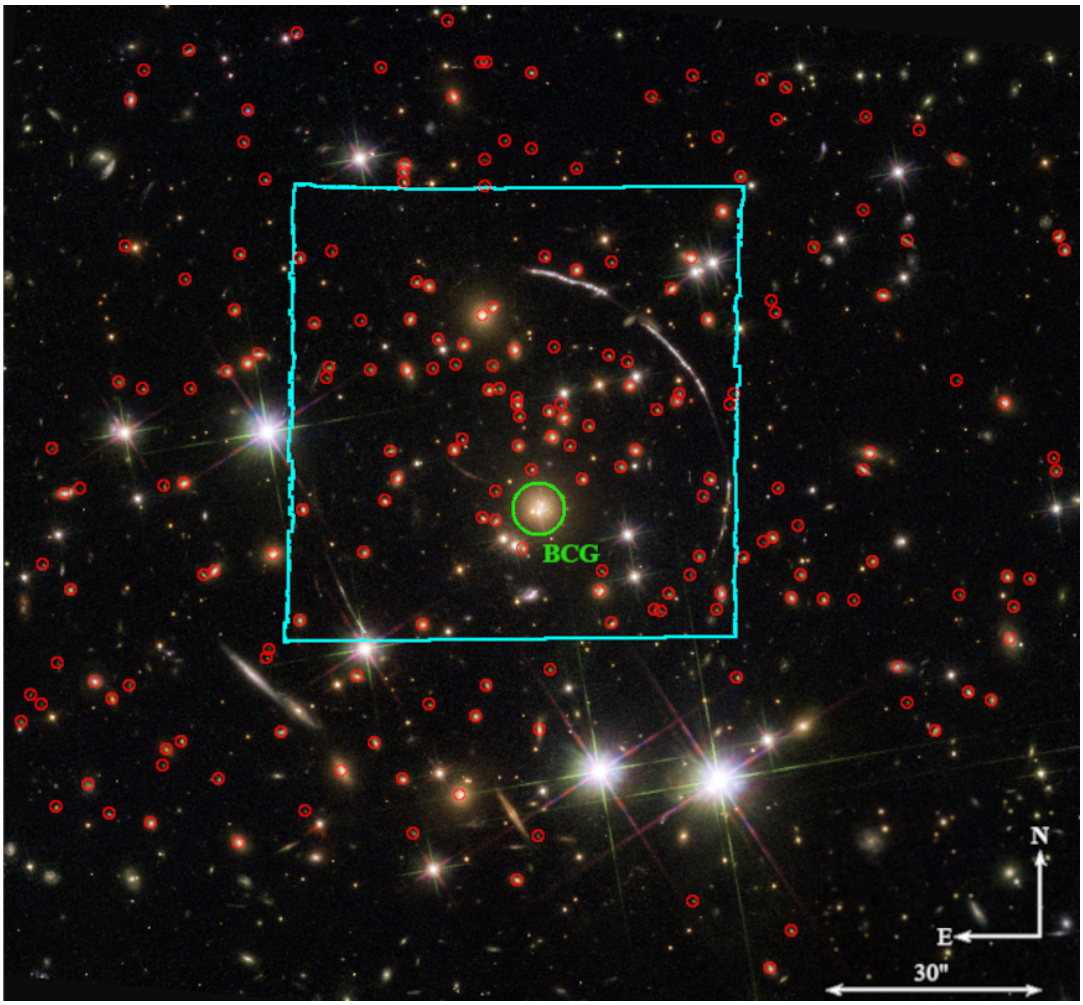

Fig. 5.9. PSZ1-G311 determined cluster members from the final catalogue are circled in red, while the BCG is circled in green. Only the galaxies inside the the cyan contours of the MUSE data-cube are considered to extract the l.o.s. velocity dispersion.

5.3.3 Measured velocity dispersions

Following the approach of Bergamini et al. (2019 [122]), we use the VLT/MUSE observations to extract the line-of-sight (l.o.s.) stellar velocity dispersions (σ_{ap}^{gal}) of a sample of the cluster galaxies. These measurements are used to calibrate the normalization and slope of the $\sigma - mag$ scaling relation for the cluster galaxies, parameters that can be used as prior for the scaling relations of the sub-halo component in the lens model. We measure the average luminosity-weighted value of the l.o.s. projected velocity dispersion, obtained extracting the spectra of the galaxies from the MUSE data-cube inside an aperture of $R_{ap} = 0.8''$. The choice of the aperture radius is the optimal compromise between a low contamination from other sources and an high signal-to-noise ratio ($\langle S/N \rangle$). To measure the velocity dispersion consider only 46 of the cluster galaxies inside the MUSE data-cube for which we could obtain a spectrum (see Fig.(5.9)). We use the software `pPXF` (Cappellari & Emsellem 2004 [123]; Cappellari 2017 [124]) which can extract the stellar kinematics from the galaxy spectra by performing a full spectrum fitting: in particular, `pPXF` cross-correlates the input observed spectrum with a set of stellar templates and returns a measure of the velocity dispersion from the best-fit. Each produced spectrum from this procedure is visually inspected to check that the fit was performed correctly. The selected wavelength range is [3600-5000]Å rest-frame, in order to avoid the red wavelengths of the MUSE spectra, which are the most contaminated by skylines residuals, and at the same time to include the galaxy absorption lines (such as the Ca doublet). If the skyline residuals are high, we mask the corresponding pixels to exclude them from the fitting procedure and avoid the introduction of bias in the measurements (see Fig.(5.11)). For the i-th cluster member galaxy we obtain a velocity dispersion value ($\sigma_{ap,i}^{gal} \pm \delta\sigma_{ap,i}^{gal}$) and a $\langle S/N \rangle_i$. To ensure a robust measurement we only consider galaxies with $\sigma_{ap}^{gal} > 60$ km/s and a $\langle S/N \rangle > 10$. The final sample includes 14 galaxies.

We can now write the $\sigma - mag$ scaling relation for the sample as

$$\hat{\sigma}_{ap,i}^{gal} = \sigma_{ap}^{ref} 10^{0.4\left(m_{F160W}^{ref} - m_{F160W,i}^{gal}\right)\alpha} \tag{5.5}$$

where the $\hat{\sigma}_{ap,i}^{gal}$ is the predicted velocity dispersion from the Faber-Jackson relation (Faber and Jackson, 1976 [74]) for the i-th cluster member with magnitude $m_{F160W,i}^{gal}$. To determine the values of the normalization σ_{ap}^{ref} and the slope α that better fit our sample of cluster galaxies, we adopt a Bayesian approach. We also include the parameter $\Delta\sigma_{ap}$, the intrinsic scatter of the galaxies around the Faber-Jackson. Using the Bayes theorem (see Sec.(4.2.2)), the maximized posterior is defined as the product of a likelihood and a prior which are respectively defined as

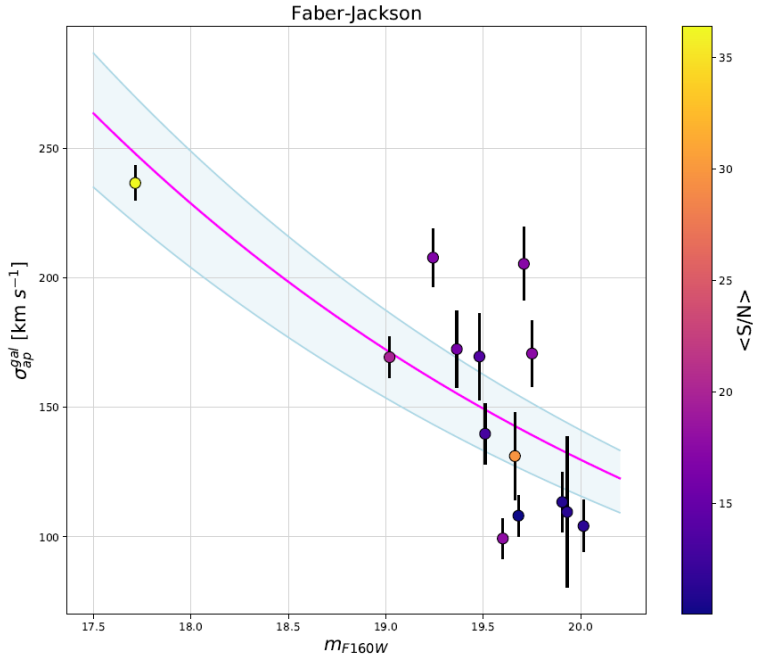

Fig. 5.10. Plot relating the measured internal stellar velocity dispersion of 14 cluster galaxies and their HST/F160W magnitude. The galaxies are represented by the filled circles, where their color varies as a function of the mean signal-to-noise ratio. The magenta line is the best-fit $\sigma - mag$ relation obtained, while the light blue area corresponds to the intrinsic $\Delta\sigma_{ap}$ scatter.

$$\begin{aligned} &\ln\left[p\left(\sigma_{ap}^{gal} \mid m^{gal}, \delta\sigma_{ap}^{gal}, \sigma_{ap}^{ref}, \alpha, \Delta\sigma_{ap}\right)\right] = \\ &= -\tfrac{1}{2}\sum_{i=1}^{N}\left\{\frac{\left(\sigma_{ap,i}^{gal} - \hat{\sigma}_{ap,i}^{gal}\right)^2}{\left(\delta\sigma_{ap,i}^{gal}\right)^2 + \Delta\sigma_{ap}^2} + \ln\left[2\pi\left(\left(\delta\sigma_{ap,i}^{gal}\right)^2 + \Delta\sigma_{ap}^2\right)\right]\right\}, \end{aligned} \tag{5.6}$$

and

$$\ln\left[p\left(\sigma_{ap}^{ref}, \alpha, \Delta\sigma_{ap}\right)\right] = \begin{cases} -\ln\left(\Delta\sigma_{ap}\right), & \text{if } \sigma_{\min}^{ref} < \sigma_{ap}^{ref} < \sigma_{\max}^{ref} \\ & \text{and } \alpha_{\min} < \alpha < \alpha_{\max} \\ & \text{and } \left(\Delta\sigma_{ap}\right)_{\min} < \Delta\sigma_{ap} < \left(\Delta\sigma_{ap}\right)_{\max} \\ -\infty, & \text{otherwise} \end{cases} . \tag{5.7}$$

The limits in Eq.(5.7) are chosen based on the measured velocity dispersions. The sampling of the posterior is based on the Affine-Invariant MCMC Ensemble sampler (see Goodman & Weare 2010 [125], Foreman-Mackey et al. 2013 [126]). It is performed by 100 walkers that sample 5000 points each, and we remove 80 steps in the burn-in phase. The marginalized posterior distribution for the fitting parameter, along with the median values and [16-th, 84-th] percentiles of their marginalized posterior distributions are shown in Fig.(5.12). The resulting best-fit scaling relation is the solid magenta line in Fig.(5.11), while the light-blue band represent the fitted intrinsic scatter ($\Delta\sigma_{ap}$) of the relation. These results are useful in our lens model, since to reduce the number of free-parameters we adopt the following scaling

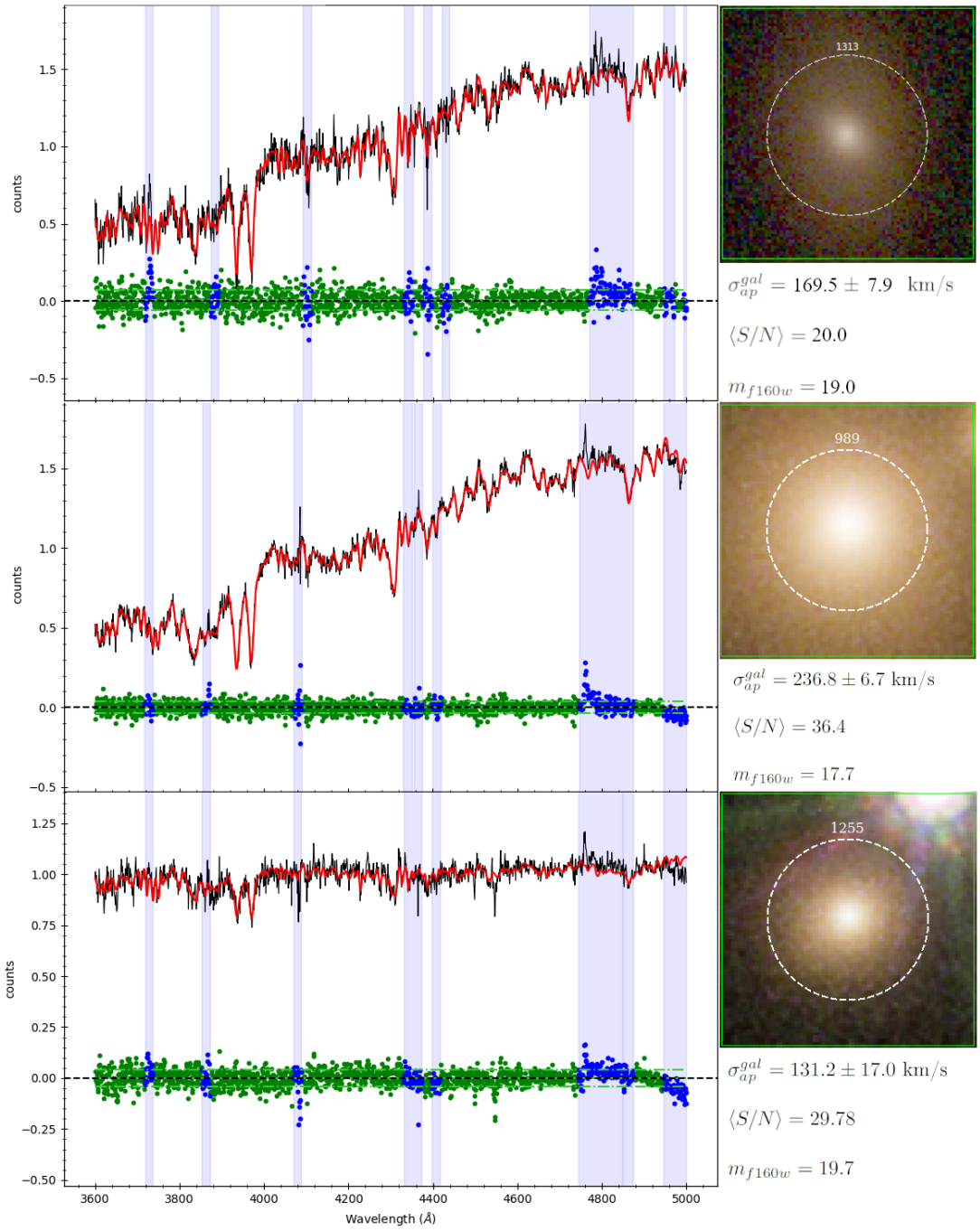

Fig. 5.11. Plots showing the fitting procedure performed by `pPXF` for three cluster members of PSZ1-G311, and the resulting velocity dispersions (σ_{ap}^{gal}), $\langle S/N \rangle$ and F160W magnitudes. The spectral extraction is performed in the white dashed apertures of 0.8" radius as shown in the HST RGB image cut-outs. The `pPXF` best-fit models are shown in red over the black galaxy spectra. The green points represent the data-model residuals, and the blue bands (with corresponding blue points) represent the excluded rest-frame wavelength ranges due to high sky contamination. The three galaxies are the highest signal-to-noise cluster members in our sample.

relations for each sub-halo component:

$$\begin{cases} \sigma^{gal}_{LT,i} = \sigma^{ref}_{LT} \left(\frac{L_i}{L_{ref}}\right)^{\alpha} \\ r^{gal}_{cut,i} = r^{ref}_{cut} \left(\frac{L_i}{L_{ref}}\right)^{\beta_{cut}} \end{cases} \tag{5.8}$$

where σ_{LT} is the fiducial velocity dispersion defined in LENSTOOL, related to the central velocity dispersion of the dPIE as $\sigma_0 = \sqrt{\frac{3}{2}}\sigma_{LT}$; L_i is the luminosity of the i-th cluster member and L_{ref} is the reference luminosity; the two normalizations are computed at the reference luminosity. From the values of the measured velocity dispersions, in our lens models, we can fix the the value of the slope α to the fitted one. Also, the value of the slope β_{cut} is inferred by assuming that the galaxy mass-to-light ratio scales as a function of the luminosity as $M_{tot,i}/L_i \propto L_i^{\gamma}$. In this case, we can show that

$$\beta_{cut} = \gamma - 2\alpha + 1. \tag{5.9}$$

We adopt $\gamma = 0.2$ in according to the observed fundamental plane (Faber et al. 1987 [127], Bender et al. 1992 [128]). Since $\alpha = 0.311^{+0.076}_{-0.079}$, we obtain $\beta_{cut} = 0.577^{+0.151}_{-0.158}$. For the normalization σ^{ref}_{LT} we use a flat prior with limits evaluated on the measured velocity dispersions of our sample, and for the normalization r^{ref}_{cut} we adopt a large flat prior between $0.1''$ and $25''$.

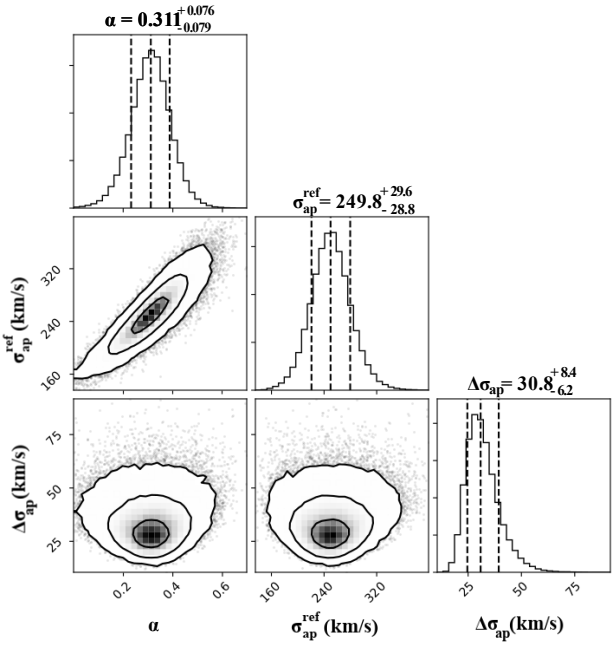

Fig. 5.12. Marginalized 2D and 1D posterior PDF for the scaling relation parameters obtained from the measured velocity dispersions of 14 cluster member galaxies. We report the median value of the parameter with the 16-th and 84-th percentiles (also shown as dashed lines in each histogram) of the marginalized distribution.

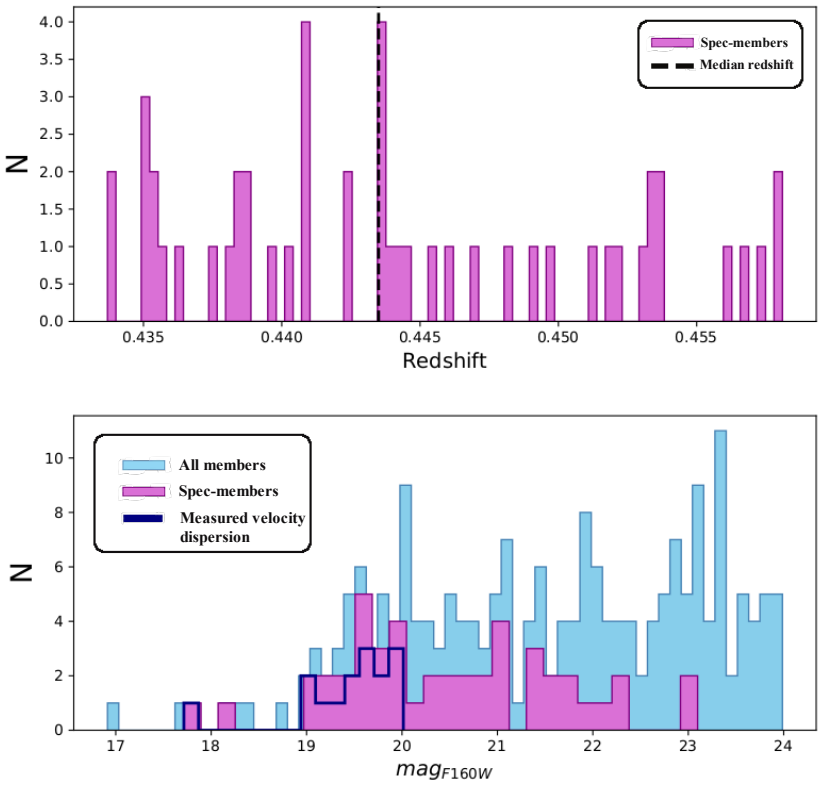

Fig. 5.13. *Top:* redshift distribution of the spectroscopically identified cluster member galaxies with a redshift within 3000 km/s from the median cluster redshift, z=0.4436 (black dashed line). *Bottom:* distribution of the cluster member galaxies as a function of their F160W band magnitude; the light blue histogram represents all 201 identified cluster members, the magenta histogram represents the 46 spectroscopic members, and the 15 cluster members with a measurement of their velocity dispersion are in dark blue.

5.3.4 Important deflectors

As already explained in Sec.(4.2.1), some of the catalogued cluster members and background/foreground objects can be parametrized separately from the scaling relations. This implies that they are described by a mass profile where its parameters are fixed or can vary between limits based on a choice of a prior. The values chosen for the flat priors limits in the different tested lens models are discussed in the next section. Here, we list the important deflectors of our catalogue:

- Deflector **1678**: described by a circular dPIE model ($e = 0$). In all models this deflector has a fixed position and core radius, while the velocity dispersion and cut radius are free parameters.

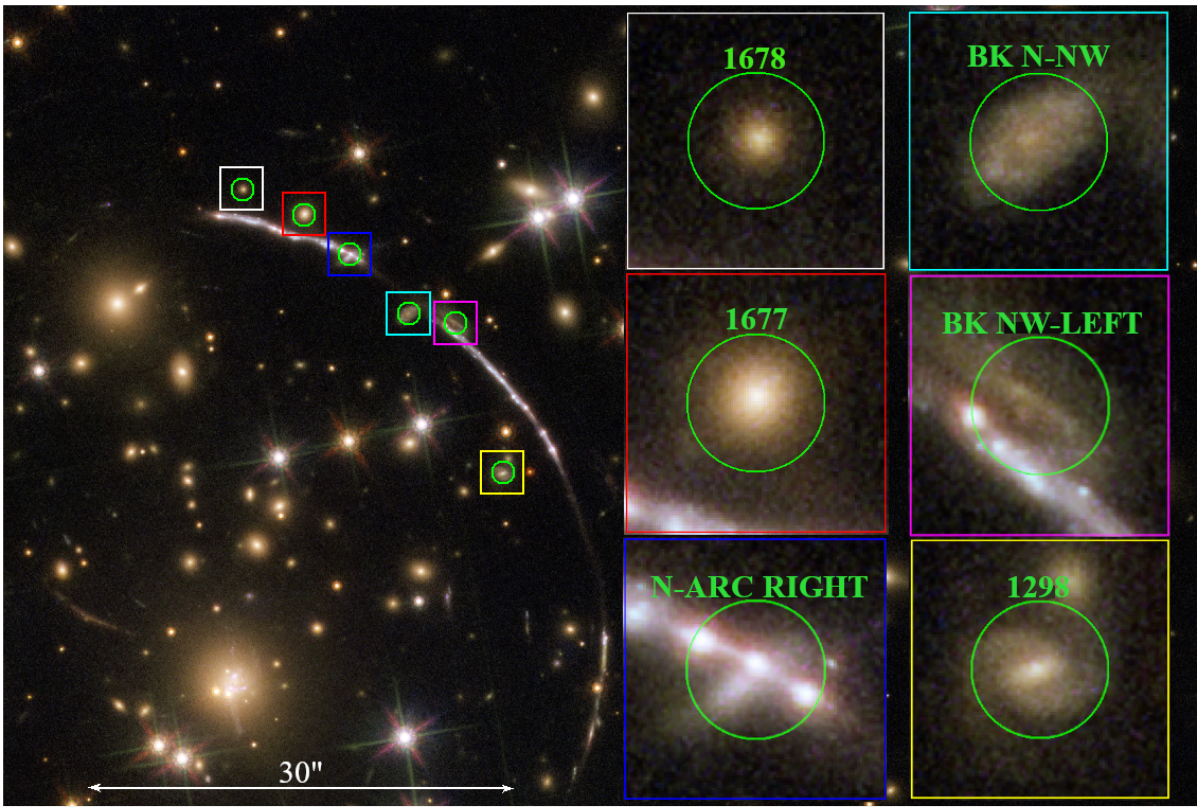

Fig. 5.14. HST RGB image of the northern segment of the arc, with cut-outs of all the important deflectors included in the lens models described in Sec.(5.3.4).

The multiple images in the north-eastern part of the arc are better reproduced with this cluster member outside the scaling relations. See the white cut-out of Fig.(5.14).

- Deflector **1677**: described by a circular dPIE model ($e = 0$). In all models this deflector has a fixed position and core radius, while the velocity dispersion and cut radius are free parameters. From its magnitude, we can deduce that its velocity dispersion is higher than that of deflector 1678. The multiple images in the central part of the arc are better reproduced with this cluster member outside the scaling relations. See the red cut-out of Fig.(5.14).

- Deflector **N-ARC RIGHT**: described by an elliptical dPIE mode. This object is clearly elliptical and its superposition with the arc creates three multiple images of the same knot. The ellipticity and position angle are additional free parameters, and the prior limits are changed during the testing. See the blue cut-out of Fig.(5.14).

- Deflector **BK N-NW**: described by a circular dPIE model ($e = 0$). This object is a background galaxy at redshift $z = 0.5578$. Since `LENSTOOL` can not perform multi-plane lensing, we can

describe the contribution of this deflector by parametrizing a fictitious halo at its position assuming it is at the same redshift of the cluster. See the cyan cut-out of Fig.(5.14).

- Deflector **BK NW-LEFT**: described by by an elliptical dPIE model. This object is a background galaxy at redshift $z = 0.7346$. Its contribution to the multiple images is accounted by parametrizing it as a fictitious halo at the cluster redshift. In some models, its position is left free to vary. See the magenta cut-out of Fig.(5.14).

- Deflector **1298**: described by an elliptical dPIE model. The peculiar configuration of families 5.1 (h,i,l) is only reproduced by leaving this cluster member out of the scaling relations. The ellipticity and position angle are additional free parameters, and the prior limits are changed during the testing. See the yellow cut-out of Fig.(5.14).

5.3.5 Multiple images identification and catalogue

Galaxy clusters are so powerful lenses that they can lens simultaneously several sources. In addition, given the extreme stretching that the background sources can experience when they are lenses by a cluster, we can often distinguish multiply imaged point-like knots in their images as described in Sec.(5.2) and Sec.(5.2.2). We define the multiple images of the same knot as a *family*, and all the images coming from the same source a *system*. The identification and pairing of the several multiple images can be carried out either by means of their color patterns, or by using their positions and morphology which must follow specific parity inversion rules. For example, we can see in Fig.(5.2.2) how the knot A show a clear emission in the F160W band and knots B and C configuration also shows a repeating color association that allows determining the position with less uncertainty. For those multiple images for which the color study is not definitive, we can speculate on the position of the critical lines by searching for sign of mirroring (see Fig.(5.15)) between pairs of knots and then test our hypothesis through subsequent steps in the modeling, as discussed in the next section.

System	Redshift (z)	Number of families	Total num. of multiple images
1	3.505	4	8
2	2.196*	2	4
3	2.393	1	2
4	1.186	1	2
5	2.369	14	60

Tab. 5.2. Identified multiple images as regrouped in different systems and families, for a catalogue of 76 total multiple images. *The redshift z_2 is the mean value obtained from the optimization of Model 10 in Sec.(5.4).

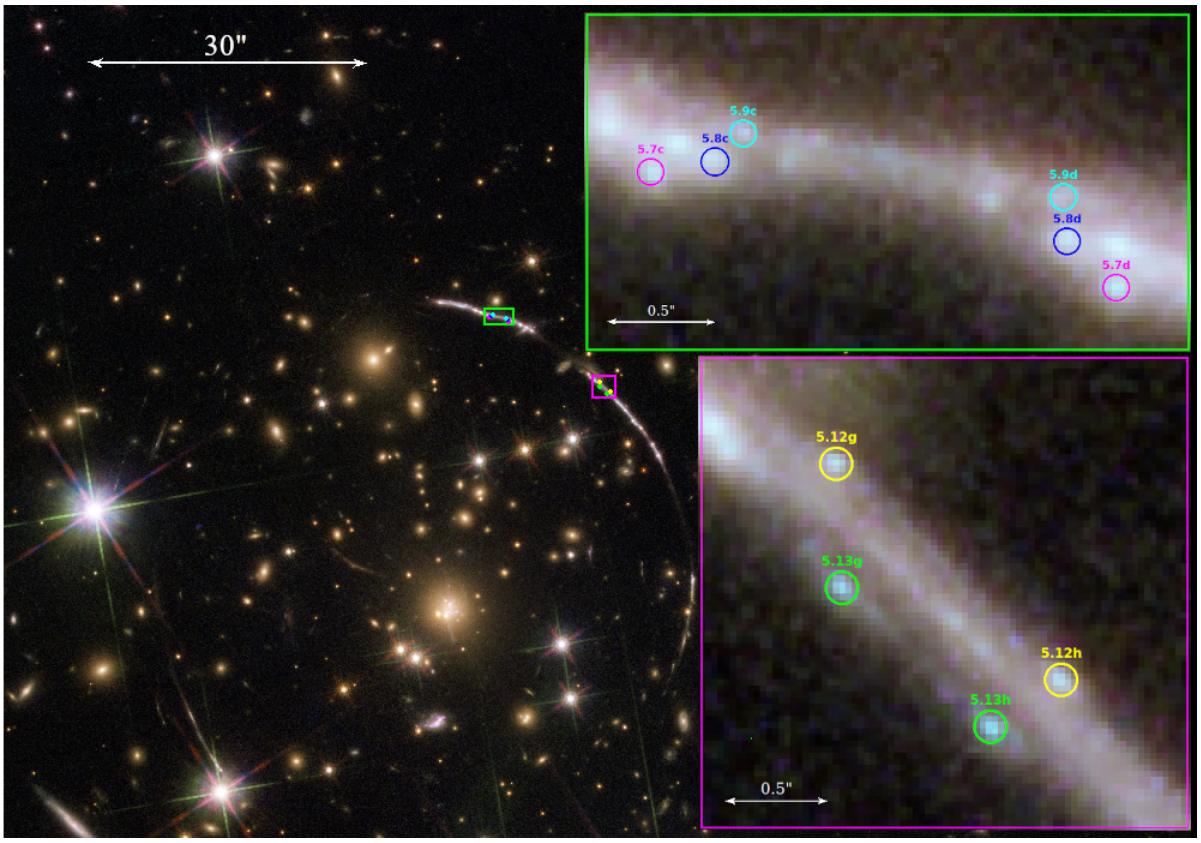

Fig. 5.15. Figure showing an example of the identification and pairing of multiple images by searching for mirrored configurations. The green cut-out highlights families 5.7, 5.8 and 5.9 *c* configuration matched with the respective *d* configuration, while the magenta cut-out highlights families 5.12 and 5.13 *g* and *h* configurations.

The construction of the catalogue of multiple images is based on the following nomenclature: each multiple image is identified through an ID containing a number and a letter, where the integer part of the number identifies the system, the fractional part of the number identifies the family, and the letter identifies all the images belonging to the same family. For example Sys-5, that corresponds to the Sunburst Arc, is composed by fourteen different image families (with IDs 5.1-5.9, 5.11-5.15): of these, family 5.1 contains twelve multiple images (from 5.1a to 5.1n) of the LyC knot, described in Sec.(5.2.1). The redshifts of the systems are measured using MUSE spectra. Altogether the catalogue used in our lens model contains a total of 76 multiple images coming from 5 background sources at different redshifts. Some details on these systems of multiple images are given in Tab.(5.2). For each of them we also report the measured redshift. While this was measured spectroscopically for Sys-1,3,4,5, for Sys-2 the redshift is estimated by LENSTOOL. Indeed, it is used as a free parameter during the model optimization process. All the systems and the multiple images are shown in Fig.(5.16). The coordinates of each image in our catalogue are reported in Tab.(5.3).

Multiple image	RA [deg]	Dec [deg]
1.1a	237.57392850	-78.18547050
1.1b	237.49811550	-78.19423620
1.2a	237.57721440	-78.18732470
1.2b	237.49737000	-78.19380270
1.3a	237.57816950	-78.18802790
1.3b	237.49685590	-78.19346700
1.4a	237.57833520	-78.18811950
1.4b	237.49684050	-78.19338580
2.1a	237.57126620	-78.19509460
2.1b	237.50142040	-78.19036930
2.2a	237.57208180	-78.19467860
2.2b	237.50141330	-78.19064780
3a	237.56717330	-78.19493140
3b	237.49528420	-78.19269000
4a	237.55932290	-78.19395080
4b	237.50903040	-78.19001330
5.1a	237.53086790	-78.18256990
5.1b	237.52561570	-78.18281030
5.1c	237.52493010	-78.18286890
5.1d	237.51916050	-78.18323520
5.1e	237.51830380	-78.18331130
5.1f	237.51751690	-78.18340990
5.1g	237.50925300	-78.18470480
5.1h	237.50157090	-78.18631220
5.1i	237.49984950	-78.18680110
5.1l	237.49897480	-78.18713330
5.1m	237.49392040	-78.19082100
5.1n	237.56288320	-78.19652950
5.2a	237.52815520	-78.18265610
5.2b	237.52747910	-78.18268700
5.2c	237.52396860	-78.18295440
5.2d	237.52020090	-78.18315250
5.2g	237.50884650	-78.18480390
5.2h	237.50384950	-78.18578400
5.2m	237.49400480	-78.19086520
5.2n	237.56327320	-78.19644360
5.3c	237.52353250	-78.18298000
5.3d	237.52075450	-78.18311290

Multiple image	RA [deg]	Dec [deg]
5.3g	237.50833380	-78.18489690
5.3h	237.50496350	-78.18555120
5.3m	237.49404200	-78.19091930
5.3n	237.56351420	-78.19638080
5.4a	237.53053750	-78.18258360
5.4b	237.52599170	-78.18277860
5.4c	237.52471000	-78.18289080
5.4d	237.51934170	-78.18322190
5.4e	237.51819130	-78.18332720
5.4f	237.51764370	-78.18339440
5.5a	237.52867080	-78.18261250
5.5b	237.52719090	-78.18267390
5.5c	237.52397540	-78.18292360
5.5d	237.52005380	-78.18312970
5.6a	237.52934380	-78.18253640
5.6b	237.52710830	-78.18262160
5.6c	237.52374540	-78.18289800
5.6d	237.52009070	-78.18306850
5.7c	237.52367960	-78.18301330
5.7d	237.52074440	-78.18316290
5.8c	237.52332440	-78.18299810
5.8d	237.52105850	-78.18310360
5.9c	237.52313140	-78.18296550
5.9d	237.52109980	-78.18304890
5.11c	237.52282320	-78.18299740
5.11d	237.52155460	-78.18305280
5.12g	237.50752850	-78.18494690
5.12h	237.50602110	-78.18524470
5.13g	237.50746700	-78.18511750
5.13h	237.50646360	-78.18530950
5.13m	237.49411630	-78.19072630
5.14i	237.50003930	-78.18683690
5.14l	237.49919580	-78.18714270
5.15a	237.53276520	-78.18241490
5.15f	237.51713490	-78.18331080
5.15h	237.50371900	-78.18565270
5.15m	237.49378830	-78.19152710
5.15n	237.56323270	-78.19629390

Tab. 5.3. Coordinates (RA,Dec) of the 76 multiple images in our catalogue.

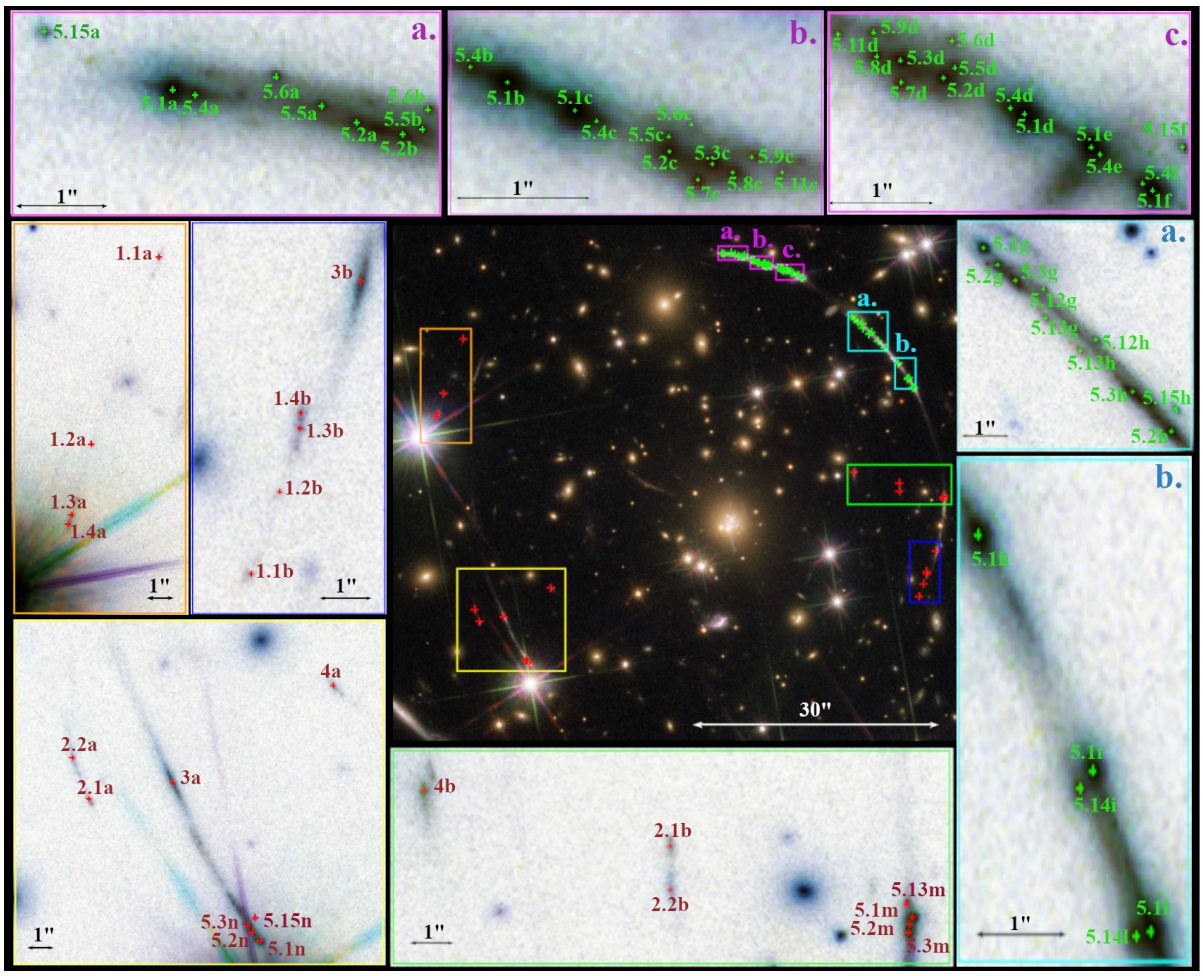

Fig. 5.16. HST color composite image of PSZ1-G311 with all catalogued multiple images shown. The first segment of the Sunburst Arc is highlighted in the magenta cut-outs *a*,*b* and *c*, and with the cyan cut-outs *a* and *b*, we can see almost all families of Sys-5. Sys-1 is visible in both the orange and blue cut-out. In the latter we can also see family 3b. Sys-2 is visible in both the yellow and green cut-out. In the yellow cut-out we can also see images 3b, 4a, 5.1n, 5.2n, 5.3n and 5.15n. Images 4b, 5.1m, 5.2m, 5.3m and 5.13m are visible in the green cut-out.

5.3.6 Models evolution

In this section we discuss some of the lens models built in different steps in order to obtain a final solution that best describes our galaxy cluster's mass distribution. In the following, we start from the simplest model and each improvement is characterized by increasing complexity and modifications in the limit range for the free parameters. The optimization is performed using the maximum-likelihood runmode.

The choice of the values of the limits for the free parameters' flat priors is dictated by different factors: for instance, the evaluation of the position of the multiple images guided the starting guess for the truncation and core radii as explained in Sec.(4.2.1); the measurements of the velocity dispersions, as described in Sec.(5.3.3), allowed to define the limits for the parameter in each profile; other parameters such as ellipticity and positions, are visually evaluated from the HST images. Each starting guess is then re-evaluated based on the outcome of the model, and changed when necessary.

- **Model 1:** the first model is based only on the most secure multiple image family, namely 5.1. It is composed by two DM halos, the BCG and the six deflectors outside the scaling relations. The main cluster-scale DM halo is described by an elliptical PIEMD profile centred on the BCG. The velocity dispersion is a free parameter, varying between 400 km/s and 1200 km/s, along with the ellipticity and core radius. The ellipticity is small, since the Sunburst Arc shape suggests that the cluster is quite circular, and the core radius varies between $0''$ and $20''$. The second DM halo is introduced to take into account asymmetries in the mass distribution not well described by the central elliptical halo. It is described by an elliptical PIEMD profile positioned in the north-east area near the first segment of the arc and its velocity dispersion varies between 250 km/s and 750 km/s. The BCG is described by a circular dPIE profile, since there are no multiple images to constrain the ellipticity, and its position is taken as the reference centre for all other halos. For the BCG we expect a small core radius. Therefore, we fix it at $0.01''$. We assume that the velocity dispersion is in the range 10 km/s and 600 km/s, and that the cut radius is between $0.1''$ and $20''$. In this first model all deflectors are described by circular dPIEs, with velocity dispersions lower limits at $50 \div 80$ km/s and higher limits at $120 \div 150$ km/s, fixed core radius at $0.01''$, and cut radius between $0.1''$ and $10''$. This is the simplest model and the total Δ_{rms} of this model is $0.31''$.

- **Model 2:** this model includes Sys-3, Sys-4 and families 5.1 to 5.3. It is composed by the two DM halos and the six deflectors, while the BCG is incorporated in the main DM halo to investigate whether its contribution is relevant to the model or we can obtain a good model with less free parameters by not including it. In this model we add the position angle and ellipticity to the

free parameters for deflector N-ARC RIGHT and for deflector 1298. The ellipticity of deflectors N-ARC RIGHT and 1298 varies between 0.0 and 0.7, and between 0.2 and 0.9, respectively. The images are all well reproduced by the model, except for 5.1a that is mapped with a great off-set within respect to the observed position (see Fig.(5.17), orange box). We try to solve this problem by modifying the priors on the position of the second DM halo. The total Δ_{rms} of this model is $0.12''$.

- **Model 3:** this model includes also families 5.7 to 5.9. It is composed by the two DM halos and the six deflectors, while the BCG is incorporated in the main DM halo. We modified the prior limits for the main DM halo velocity dispersion, to 700-1200 km/s and the prior limits for the second DM halo y-position to $15'' \div 31''$ from the reference point (i.e. the BCG position); this fixed the problem with image 5.1a (see Fig.(5.17), cyan box). The distribution of the resulting displacements between the observed and predicted multiple images in this model is shown in Fig.(5.18), for a total Δ_{rms} of $0.14''$.

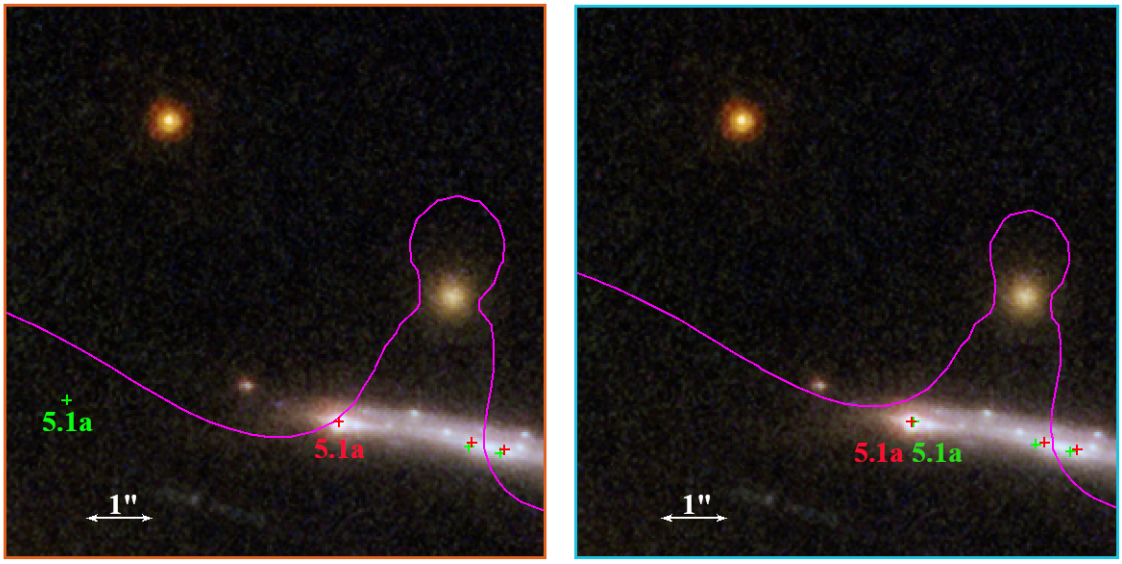

Fig. 5.17. This image shows the comparison between the position of the critical line in Model 2 (orange box) and 3 (cyan box), that influences the predicted position of image 5.1a. The observed image is indicated in red, while the model-predicted one is shown in green. Note how in Model 3 the positions of the observed and the predicted image match perfectly.

- **Model 4:** this model includes also families 5.4(a,b,c), 5.6 and 5.11. It is composed by the two DM halos and the six deflectors, while the BCG is incorporated in the main DM halo. Based on the last model results, we modified the priors for the main DM halo ellipticity, which we assume to be in the range $0.0 \div 0.5$, and the priors for deflector N-ARC RIGHT ellipticity, which we assume in the range $0.0 \div 0.9$. Moreover, we modified the priors for deflector BK NW-LEFT velocity dispersion ($10 \div 150$ km/s), and we added the position angle and ellipticity, with limits between 0.0 and 0.9,

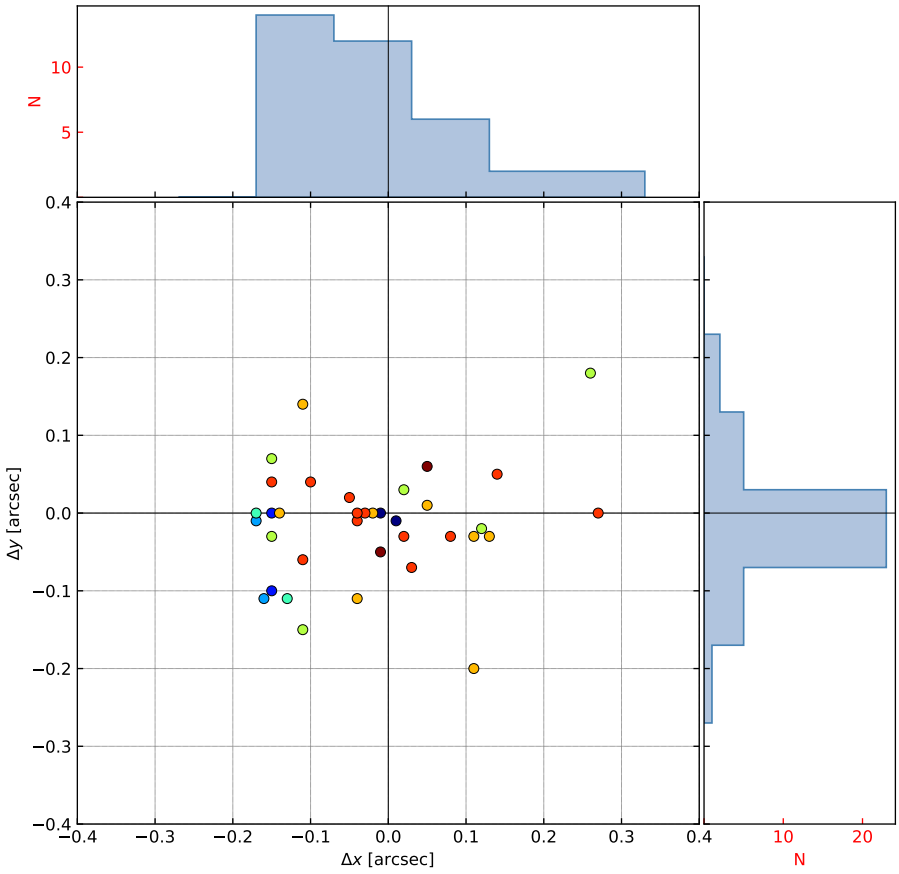

Fig. 5.18. The graph shows the 2D and 1D distribution of the displacements Δ_x and Δ_y between the observed and predicted position of the multiple images of Model 3. Each family is marked with a different color, for a total of 36 multiple images.

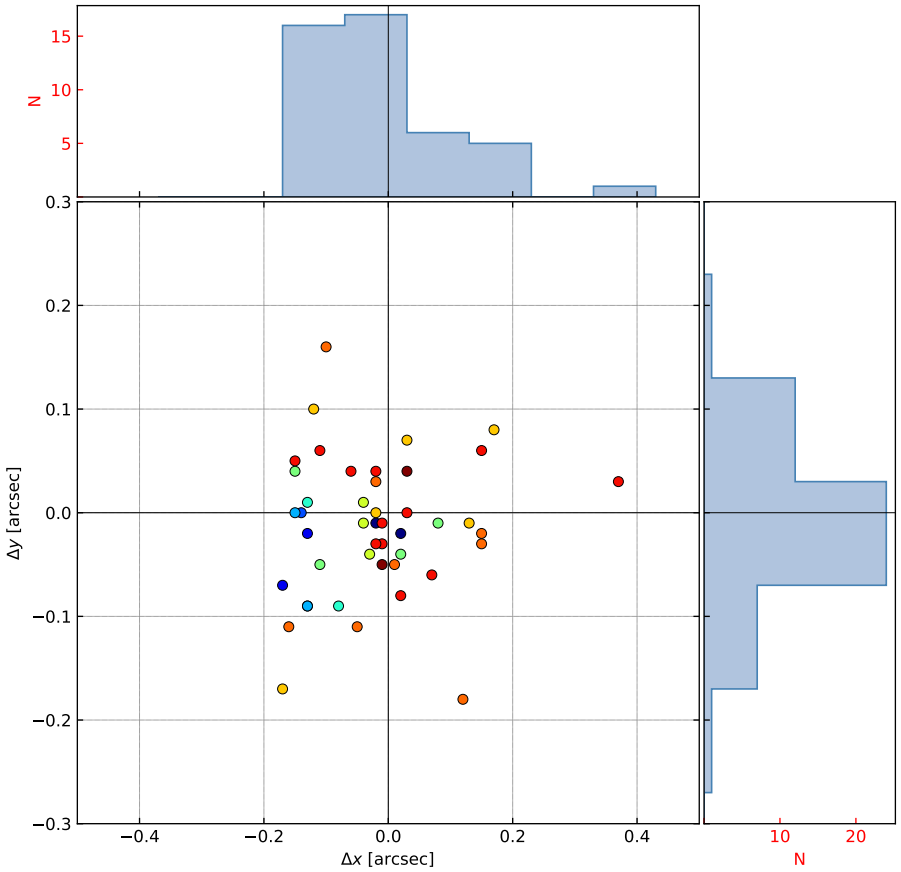

Fig. 5.19. The graph shows the 2D and 1D distribution of the displacements Δ_x and Δ_y between the observed and predicted position of the multiple images of Model 4. Each family is marked with a different color, for a total of 45 multiple images.

to its free parameters. As expected, the model predicts the presence of additional multiple images of family 5.4, that are already catalogued as 5.4(d,e,f). More multiple images of families 5.7, 5.8, 5.9 and 5.11 are predicted, for instance in the third segment of the arc, where the resolution did not allow us a secure identification. The distribution of the resulting displacements between the observed and predicted multiple image positions with this model is shown in Fig.(5.19). The total Δ_{rms} is 0.13″.

- **Model 5:** this model includes the additional families 5.4(d,e,f), and families 5.12, 5.13(g,h). The last two show a mirrored configuration (see Fig. (5.15), magenta box) which is not reproduced correctly by the previous models. In particular, the critical lines for those models were not passing at the expected positions based on the geometrical configurations of these images. We add these families now as additional constraint. The components are the two DM halos and the six deflectors,

while the BCG is incorporated in the main DM halo. This model results show how now the critical line position is where it is expected to be, passing in between the two 5.12, 5.13 families. However, this modification yields some problems with images 5.1h, 5.1i and 5.1l. In this area, the critical line should go around image 5.1i in order to reproduce the triplet, and to better achieve this configuration we can try adding family 5.14 to further bend the critical line (see Fig.(5.20), orange box). The total Δ_{rms} of this model is $0.17''$.

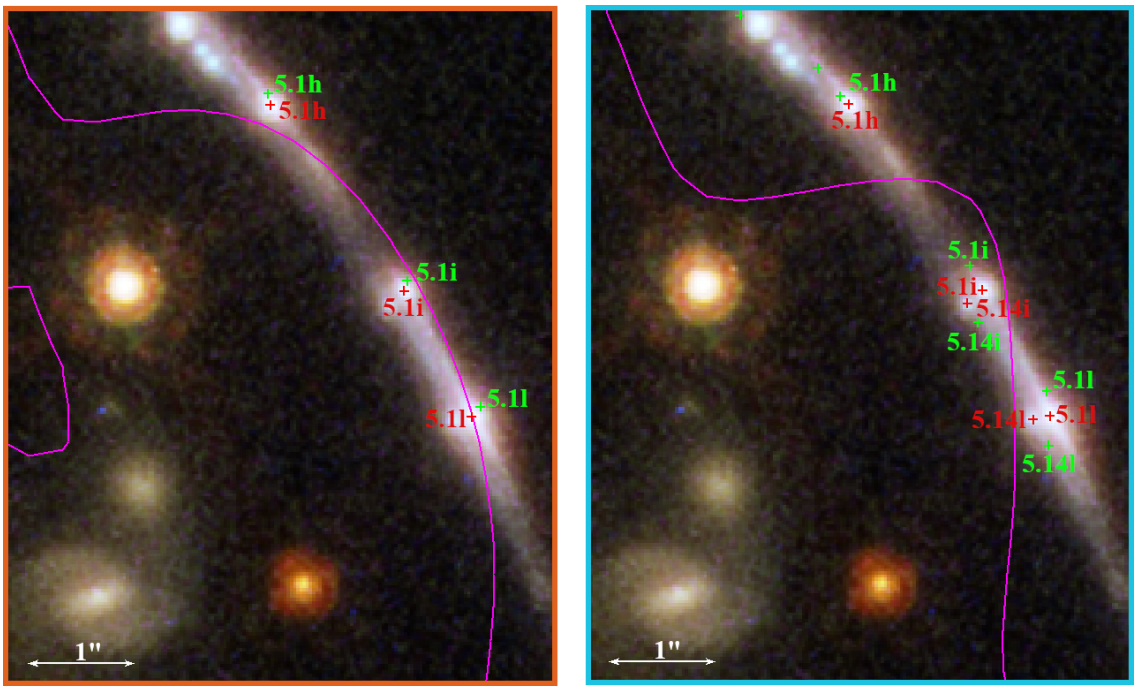

Fig. 5.20. The image shows the position of the critical line passing between families 5.1h, 5.1i and 5.1l: in Model 5 (orange box) the offset between observed (red) and predicted (green) images is small, but the critical line should go around image 5.1i to reproduce the symmetry of this configuration; in Model 6 (cyan box) this problem is fixed with the addition of family 5.14.

- **Model 6:** this model includes also family 5.14. It is composed by the two DM halos and the six deflectors, while the BCG is incorporated in the main DM halo. Based on the last model best fit results, we modified the priors for deflector 1298 velocity dispersion ($50 \div 150$ km/s). The addition of family 5.14 solved the problem of the critical line position in the area highlighted in Fig.(5.20, (cyan box)), but at the cost of an extreme ellipticity assigned to deflector 1298. In the next steps we can try adding more families or we can add the (x,y) position to the free parameters of deflector BK NW-LEFT. The distribution of the resulting displacements between the observed and predicted multiple images in this model is shown in Fig.(5.21), for a total Δ_{rms} of $0.19''$.

- **Model 7:** this model includes also family 5.5. It is composed by the two DM halos and the six deflectors, while the BCG is incorporated in the main DM halo. With the addition of this family the Δ_{rms} lowers but the critical line passing between families 5.12 and 5.13 is now misplaced.

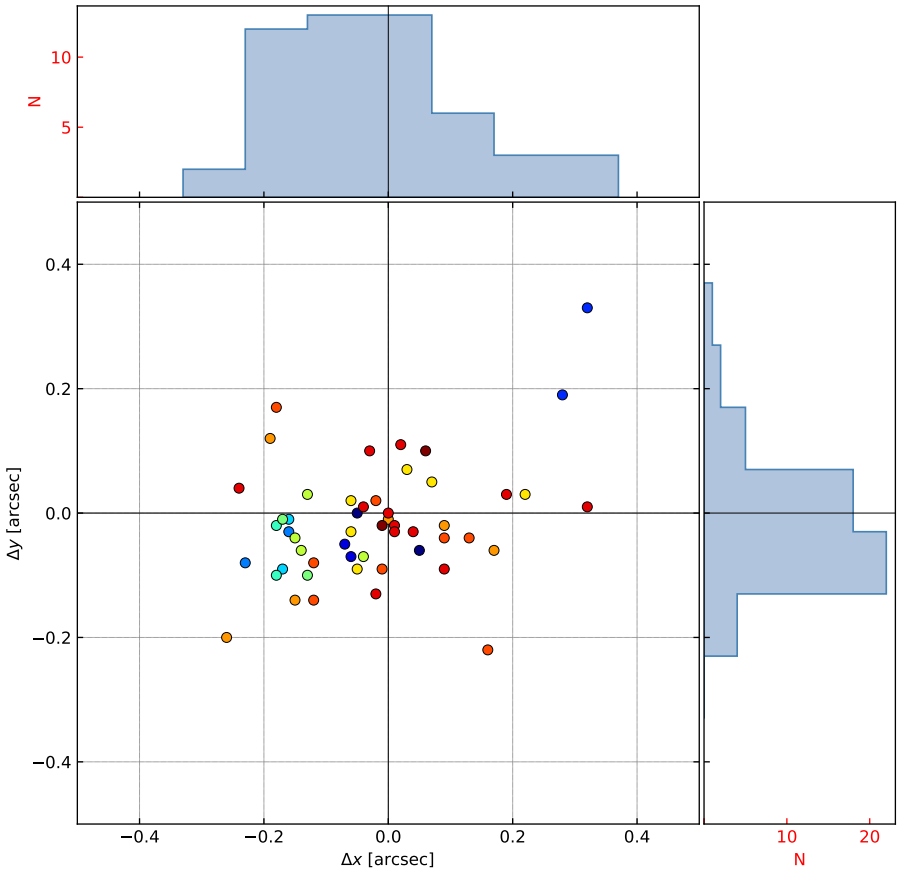

Fig. 5.21. The graph shows the 2D and 1D distribution of the displacements Δ_x and Δ_y between the observed and predicted position of the multiple images of Model 6. Each family is marked with a different color, for a total of 52 multiple images.

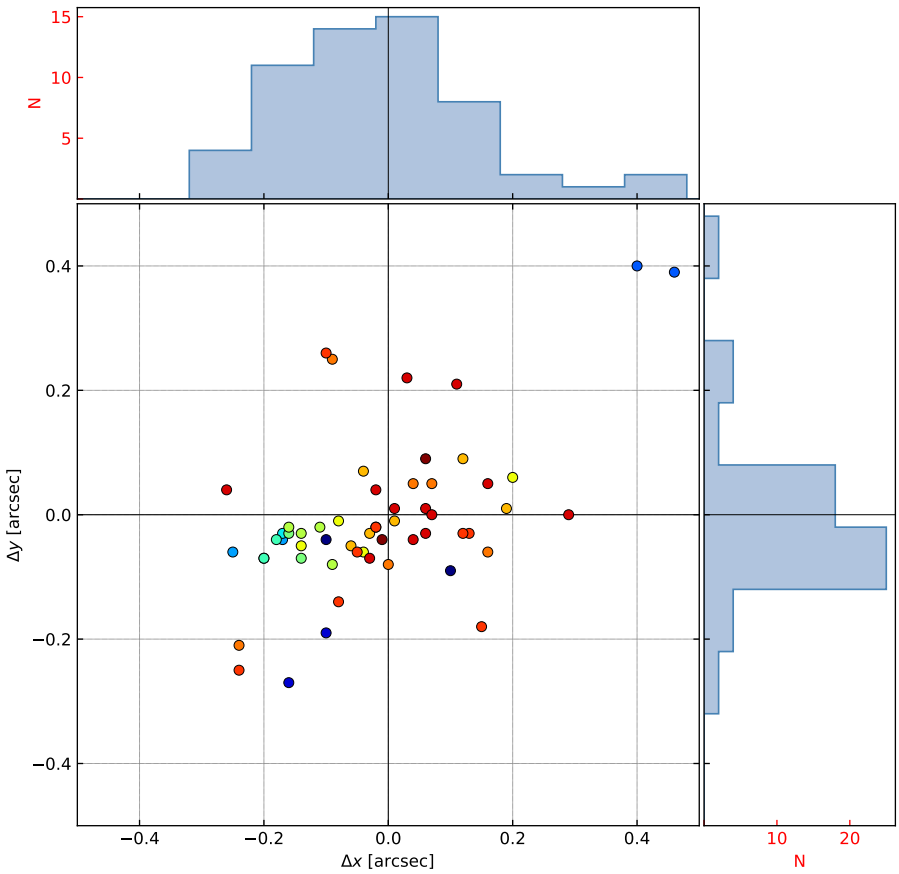

Fig. 5.22. The graph shows the 2D and 1D distribution of the displacements Δ_x and Δ_y between the observed and predicted position of the multiple images of Model 7. Each family is marked with a different color, for a total of 54 multiple images.

To fix this issue the last step is freeing the position of deflector BK NW-LEFT, so that the best fit can determine where its centre should be to better reproduce the critical line near that configuration. The distribution of the resulting displacements between the observed and predicted multiple images in this model is shown in Fig.(5.19), for a total Δ_{rms} of $0.14''$.

- **Model 8:** in this model the multiple image catalogue is the same as in Model 7, and the position of BK NW-LEFT is left free to vary. It is composed by the two DM halos and the six deflectors, while the BCG is incorporated in the main DM halo. The best fit now solved the problem with the position of the critical line between families 5.12 and 5.13, but the ellipticity of deflector BK NW-LEFT and 1298 is extreme. The total Δ_{rms} of this model is $0.15''$.

- **Model 9:** this model includes also families 5.15(h,f,m,n) and 5.13m, while the remaining 5.15a is excluded because the identification is not secure. It is composed by the two DM halos and the six

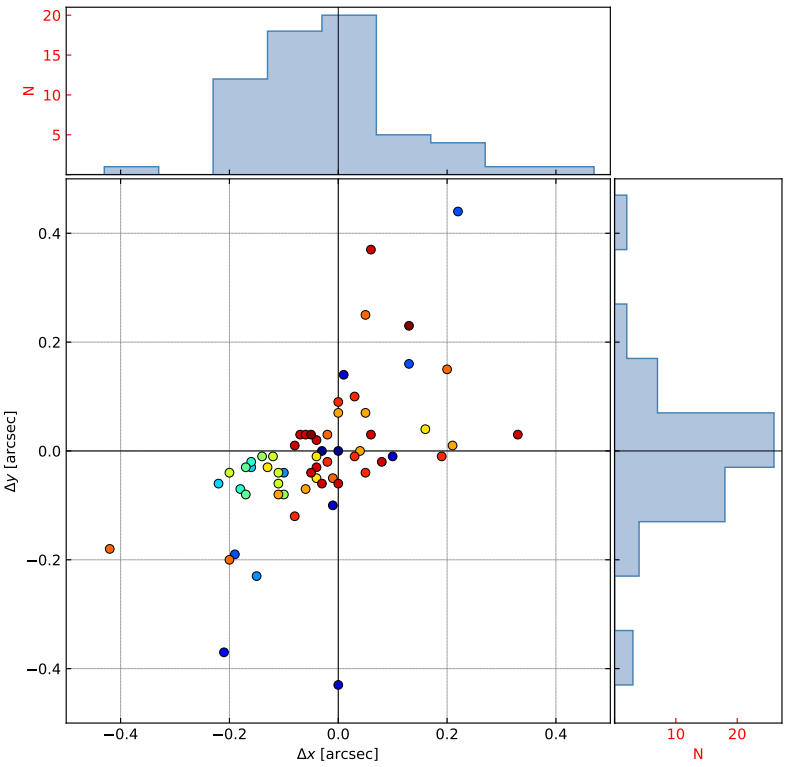

Fig. 5.23. The graph shows the 2D and 1D distribution of the displacements Δ_x and Δ_y between the observed and predicted position of the multiple images of Model 9. Each family is marked with a different color, for a total of 63 multiple images.

deflectors, while the BCG is incorporated in the main DM halo. The inclusion of these families introduces some problems such as the prediction of extra multiple images of Sys-4 that are not visible and a shift in the position of the critical line near the three images 5.1(h,i,l), as seen in Model 5. Since this modification results in a considerable increase of the Δ_{rms}, in the next model these families are excluded. The distribution of the resulting displacements between the observed and predicted multiple images in this model is shown in Fig.(5.19), for a total Δ_{rms} of $0.24''$.

Note that in all these models we never included Sys-1: this is the result of a series of test with this system, that yielded many problems such as predicting many invisible extra images of Sys-3 along the arc; this problem is caused by an unsecure matching of the multiple images, since the MUSE observations only covered part of the cluster. Thus, the identification, based on the Ly-α emission, can be verified only for the images inside the MUSE cube. This uncertainty led to the decision of excluding the system from the testing. All models' features are summarized in Tab.(5.4).

In Fig.(5.24) we show the magnification maps and the projected mass-density maps for each model: the

	Model ID	Description	M.I. Systems	Number of constraints	Free parameters	Δ_{rms} [arcsec]
Model 1	lensmodel_02_12_20	BCG + MAIN DMH + NORD DMH 6 Deflectors	5.1	22	28	0.31
Model 2	lensmodel_07_01_21	MAIN DMH + NORD DMH 6 Deflectors	3 + 4 + [5.1-5.3]	50	30	0.12
Model 3	lensmodel_10_01_21	MAIN DMH + NORD DMH 6 Deflectors	3 + 4 + [5.1-5.3] + [5.7-5.9]	56	30	0.14
Model 4	lensmodel_12_01_21	MAIN DMH + NORD DMH 6 Deflectors	3 + 4 + [5.1-5.3] + 5.4(a,b,c) + [5.6-5.9] + 5.11	68	32	0.13
Model 5	lensmodel_12_01_21	MAIN DMH + NORD DMH 6 Deflectors	3 + 4 + [5.1-5.4] + [5.6-5.9] + [5.11-5.13(g,h)]	78	32	0.17
Model 6	lensmodel_14_01_21	MAIN DMH + NORD DMH 6 Deflectors	3 + 4 + [5.1-5.4] + [5.6-5.9] + [5.11-5.13(g,h)] + 5.14	80	32	0.19
Model 7	lensmodel_18_01_21	MAIN DMH + NORD DMH 6 Deflectors	3 + 4 + [5.1-5.9] + [5.11-5.13(g,h)] + 5.14	86	32	0.14
Model 8	lensmodel_20_01_21	MAIN DMH + NORD DMH 6 Deflectors (free position pot.7)	3 + 4 + [5.1-5.9] + [5.11-5.13(g,h)] + 5.14	86	34	0.15
Model 9	lensmodel_22_01_21	MAIN DMH + NORD DMH 6 Deflectors (free position pot.7)	3 + 4 + [5.1-5.9] + [5.11-5.14] + 5.15(h,f,m,n)	94	34	0.24
Model 10	lensmodel_25_01_21	MAIN DMH + NORD DMH 6 Deflectors (free position pot.7)	3 + 4 + 2 + [5.1-5.9] + [5.11-5.13(g,h)] + 5.14	90	36	0.14

Tab. 5.4. The table summarizes the components and the resulting Δ_{rms} for each tested model. The description column lists the potential included in each model, and the multiple images systems column shows the increasing complexity of the models. Also listed, the number of constraints and of free parameters of each model.

magnification maps are created computing the absolute valute of μ at each position on the lens plane for the Sunburst-Arc redshift ($z = 2.369$), and the color scale is set between 0 and 100, so that the lighter color helps visualize where the magnification formally diverges in correspondence of the critical lines; the projected mass-density maps are shown in units of $10^{12} M_{\odot}/\mathrm{kpc}^2$, and the color scale varies between 0 and 0.005. The evolution of the models can be seen in these maps. For example, we note that the Model 1 magnification map does not show the radial critical line. In fact, this model's only constraint is family 5.1, hence we do not have constraints in the cluster core. This leads to very small core radius assigned to the components in the cluster core, such as the BCG. By adding multiple images systems at different redshifts the core is better described. Moreover, we note how the progressive introduction of different families results in a better constrained second DM halo. In the last models, we can see how there is more mass associated to the second DM halo.

The next and final model, Model 10, will be taken as reference model since it is optimized with the most complete catalogue of multiple images and it is characterized by a low Δ_{rms}. We discuss it in the next section.

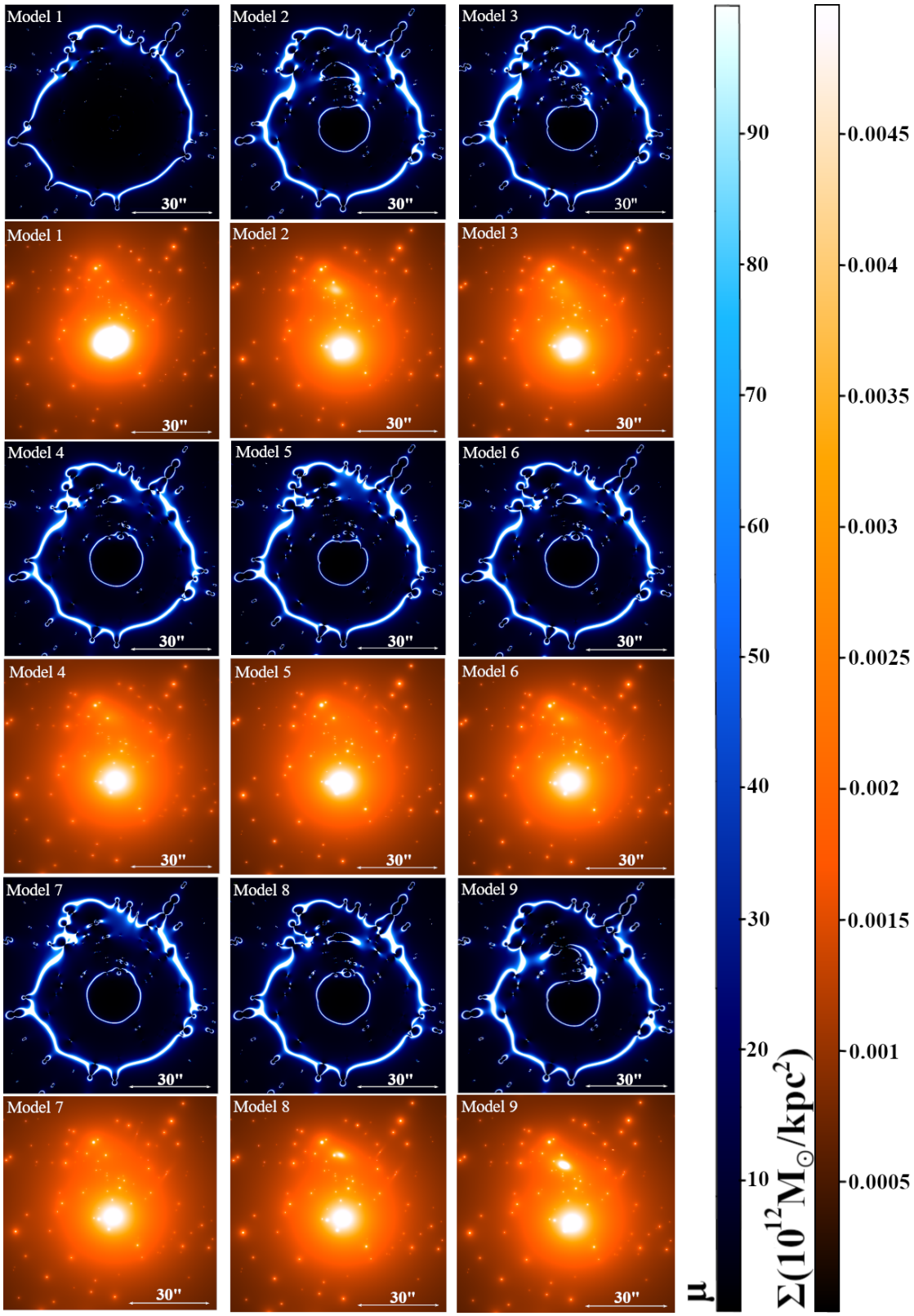

Fig. 5.24. Absolute magnification and projected mass-density maps for each model. The cool colormap ranges between 0 and 100, while the heat colormap ranges between 0 and 0.005 $10^{12} M_{\odot}/\mathrm{kpc}^2$

.

5.4 Reference Model - Model 10

We describe here the results of the reference model, Model 10. In this model the DM content of the cluster halo is parametrized by two PIEMD profiles with infinite truncation radius. The BCG is not included, and we include all the deflectors described in Sec.(5.3.4). In this model the multiple image catalogue includes all families except family 5.15, 5.13m and Sys-1 (see Tab.(5.4)), for a total of 62 multiple images. Since we were not able to determine the redshift of Sys-2, this is optimized by LENSTOOL: the input parameters and the optimized output of Sys-2 redshift are summarized in Tab.(5.5). The reference model predicts 117 multiple images associated to 4 background sources. There are 55 extra model-predicted multiple images for which there is not a secure identification. In Fig.(5.25)

Family	Input Redshift	Optimized Redshift
2.1	$1.9 \div 2.5$	$2.197^{+0.026}_{-0.025}$
2.2	$1.9 \div 2.5$	$2.194^{+0.022}_{-0.021}$

Tab. 5.5. Input interval values for the Sys-2 redshift determination and the optimized median value obtained from Model 10, with errors corresponding to the [16-th, 84-th] percentiles from the marginalized posterior distribution.

we show the 2D (scatter plot) and 1D (histograms) distributions of the difference in position in the x and y directions, between observed and predicted multiple images. We can see that the optimized positions are well determined, in fact most images show less than a $0.2''$ displacement in both directions. Indeed, the resulting $\Delta_{rms} = 0.14''$ from the best-fit is the lowest obtained for a model with this level of complexity. We can further describe the goodness of the fit by evaluating the absolute displacement ($\|\Delta_i\|$) between observed and predicted multiple images. In Fig.(5.26) we show the spatial distribution of the multiple images, with circles' sizes scaling proportionally to the absolute displacements. Fig.(5.26) also shows the surface density contour levels corresponding to 8.0, 12.0, 16.0 $\times 10^8 M_\odot/\mathrm{kpc}^2$.

The input parameters of our reference model are reported in Tab.(5.6). The free parameters are reported with the interval limits of the flat priors, while the fixed parameters are single numbers. In the last row of the table, we quote the input parameters for the optimization of the scaling relations, such as the number of cluster member galaxies (N_{gal}) included in the scaling relations, the reference magnitude (m^{ref}_{F160W}, the magnitude of the brightest cluster member galaxy in our catalogue), the slope (α, determined as in Sec.(5.3.3)) and the normalization (σ^{ref}_{LT}) of the $\sigma - mag$ relation, and finally the slope (β_{cut}, obtained from Eq.(5.9) and the quoted value of γ) and the normalization (r^{ref}_{cut}) of the cut radius-magnitude scaling relation. The best fit values of the free parameters of the lens model, along with the Sys-2 redshift, are obtained with a maximum-likelihood method.

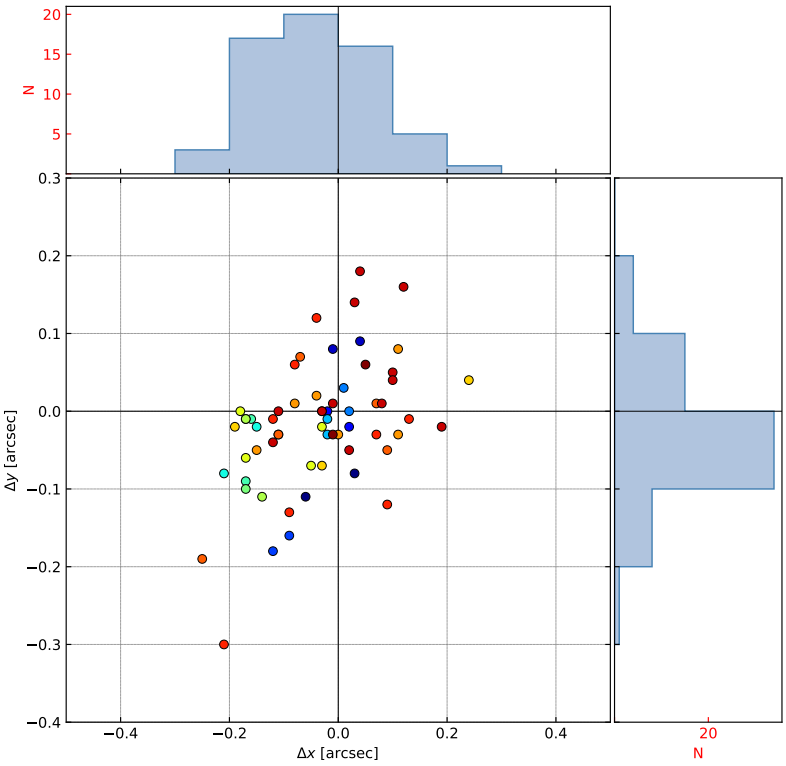

Fig. 5.25. The graph shows the 2D and 1D distribution of the displacements Δ_x and Δ_y between the observed and predicted position of the multiple images of Model 10. Each family is marked with a different color, for a total of 62 multiple images.

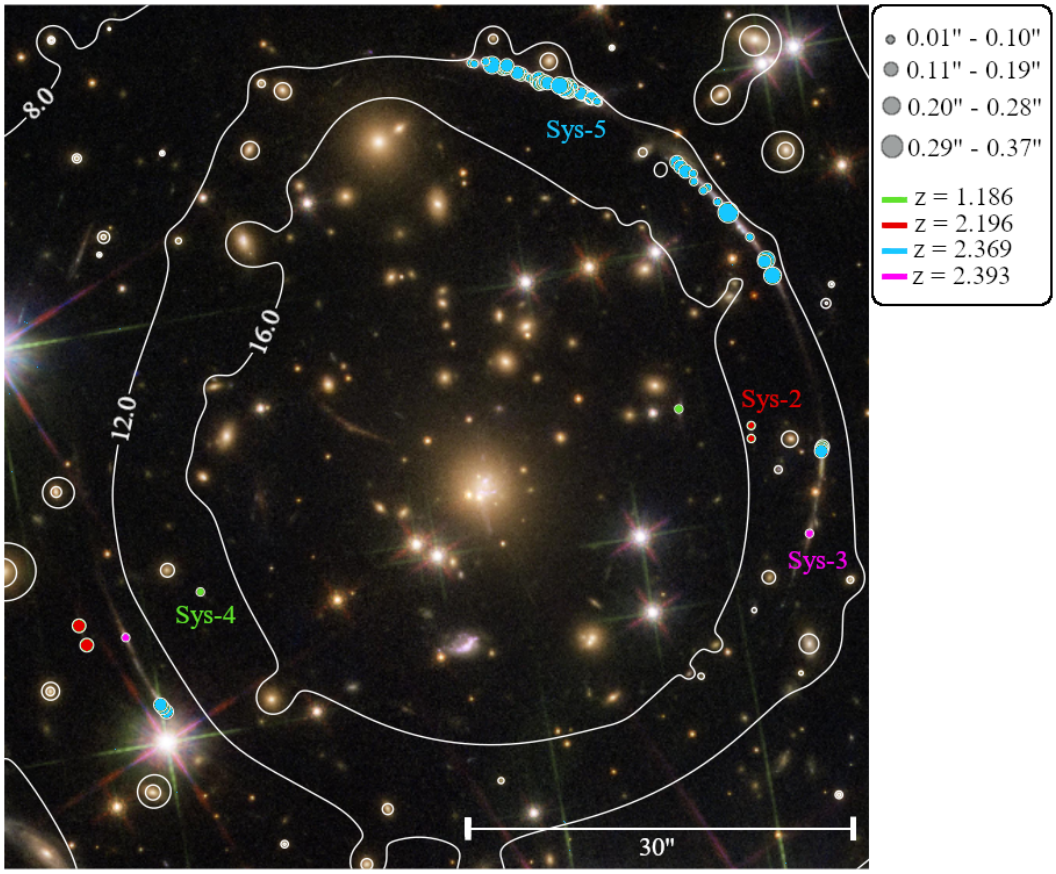

Fig. 5.26. RGB image of PSZ1-G311 with the observed positions of the 62 observed multiple images marked by circles. The size of the circles scales proportionally with the absolute displacements ($\|\Delta_i\|$) between observed and model-predicted positions. The different systems are marked with different colors corresponding to the different redshifts. The white overlaid contours represent the total projected mass-density distribution of Model 10 in units of $10^8 M_\odot/\mathrm{kpc}^2$.

	Input parameter values and intervals of Model 10						
	x [arcsec]	**y [arcsec]**	**e**	**θ [°]**	**σ_{LT} [km/s]**	**r_{core} [arcsec]**	**r_{cut} [arcsec]**
Main DM Halo	-3.0 ÷ 3.0	-3.0 ÷ 3-0	0.0 ÷ 0.3	-40.0 ÷ 40.0	800.0 ÷ 1100.0	2.0 ÷ 15.0	2000.0
Nord DM Halo	-16.0 ÷ 0.0	10.0 ÷ 34.0	0.0 ÷ 0.9	-90.0 ÷ 90.0	300.0 ÷ 800.0	0.01 ÷ 15	2000.0
1678	0.72	34.72	0.0	0.0	30.0 ÷ 120.0	0.01	0.1 ÷ 15.0
1677	5.04	33.0	0.0	0.0	50.0 ÷ 150	0.01	0.1 ÷ 20.0
N-ARC RIGHT	8.07	29.91	0.0 ÷ 0.9	0.0 ÷ 140.0	10.0 ÷ 150.0	0.01	0.0 ÷ 20.0
BK N-NW	12.31	25.96	0.0	0.0	30.0 ÷ 150.0	0.01	0.1 ÷ 15.0
BK NW LEFT	12.0 ÷ 16.0	22.0 ÷ 26.0	0.0 ÷ 0.9	-90.0 ÷ 90.0	10.0 ÷ 200.0	0.01	0.1 ÷ 15.0
1298	19.0	14.71	0.6 ÷ 0.9	0.0 ÷ 180.0	10.0 ÷ 200.0	0.01	0.0 ÷ 15.0
Scaling relations	$N_{gal} = 194$	m^{ref}_{F160W}=17.71	α=0.31	$\sigma^{ref}_{LT} = 180.0 \div 280.0$	β_{cut}=0.57	$r^{ref}_{cut} = 0.1 \div 25.0$	γ=0.2

Tab. 5.6. Input parameters values of Model 10. The singles numbers are fixed parameters, while for the free parameter we report the limits of the input flat priors. In the last row, we report the input parameters for the scaling relations.

The model uncertainties on the free parameters are evaluated using the the Bayesian MCMC sampler of LENSTOOL, described in Sec.(4.2.2). The parameters space is sampled to obtain the posterior probability distribution function. The corner plots, such as in Fig.(5.27), show us the shape of the 2D and 1D marginalized posterior PDF of the parameters that allow an estimation of the uncertainties. In Tab.(5.7) we quote, for the optimized parameters of the reference model, the median values derived from the marginalized parameter distributions and the errors corresponding to the 16-th and 84-th percentiles of the distributions. The quoted errors on the parameters are obtained re-scaling the errors on the input positions of the multiple images (from 0.5″ to 0.1396″, and from 1″ to 0.2792″) in such a way that the *reduced* chi-squared[1] of the model is 1.

As explained in Sec.(4.2.3), we expect to see a degeneracy between the normalization parameters of the $\sigma_{LT} - L$ and $r_{cut} - L$ (Eqs.5.8) scaling relations. In Fig.(5.27) we show the marginalized posterior PDF for the scaling relations normalization parameters σ^{ref}_{LT} and r^{ref}_{cut}. The posterior PDF confirms the degeneracy between the parameters.

In Fig.(5.28) we show the magnification map for the reference model. The magnification map shows the absolute value of μ in each pixel for the Sunburst-Arc redshift ($z = 2.369$). The color scale is set between 0 and 100, so that the lighter color helps visualize where the magnification formally diverges along the critical lines (mapped in red). The yellow and green cut-outs show, respectively, the least and the most magnified multiple images of family 5.1. The least magnified image is 5.1g, with a factor

[1]The reduced chi-squared is defined as the chi-squared per degree of freedom (Eq.(4.3)): $\chi^2_{dof} = \frac{\chi^2}{DoF}$.

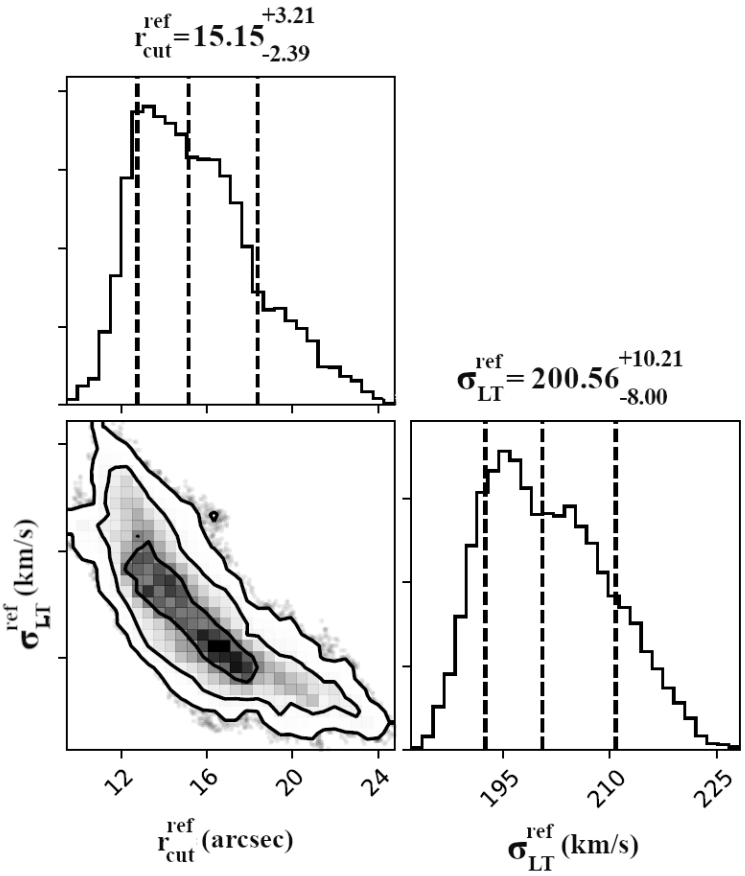

Fig. 5.27. Marginalized 2D and 1D posterior probability distributions for the normalization parameters of the $\sigma_{LT} - L$ and $r_{cut} - L$ scaling relations obtained from the optimization of Model 10. We quote the [16-th,84-th] percentiles values, also represented as the black dashed lines in the histograms. As expected, we can see a degeneracy between the two parameters.

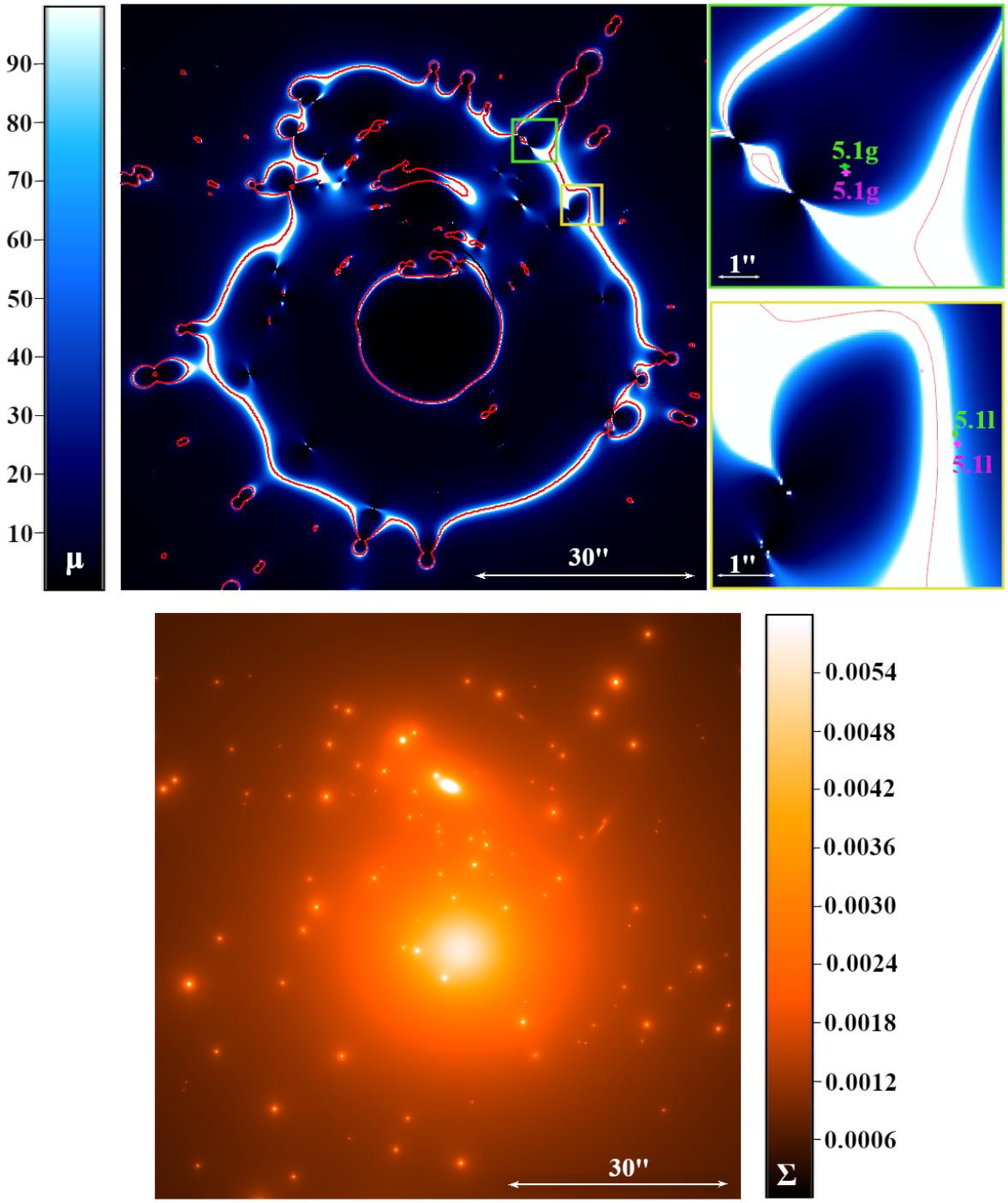

Fig. 5.28. *Top*: Absolute magnification map for the reference model at the Sunburst Arc redshift ($z = 0.369$). The scale is set between 0 and 100, and the critical lines are overlaid in red. *Green cut-out*: we show image 5.1g, the least magnified image of family 5.1, in magenta the observed and in green the model-predicted position. *Yellow cut-out*: we show image 5.1l, the most magnified image of family 5.1, in magenta the observed and in green the model-predicted position. *Bottom*: Projected mass-density map for the reference model. The color scale is set between 0 and 0.006 in units of $10^{12} M_{\odot}/\mathrm{kpc}^2$.

	Optimized output parameters of Model 10						
	x [arcsec]	**y [arcsec]**	**e**	**θ [°]**	**σ_{LT} [km/s]**	**r_{core} [arcsec]**	**r_{cut}[arcsec]**
Main DM Halo	$0.12^{+0.17}_{-0.20}$	$-2-61^{+0.39}_{-0.28}$	$0.10^{+0.01}_{-0.01}$	$-1.11^{+3.10}_{-2.98}$	$942.73^{+21.90}_{-15.13}$	$5.16^{+0.33}_{-0.34}$	2000.0
Nord DM Halo	$-3.27^{+0.81}_{-0.77}$	$17.80^{+1.57}_{-1.35}$	$0.41^{+0.08}_{-0.07}$	$-26.38^{+3.38}_{-5.03}$	$448.36^{+31.77}_{-40.48}$	$1.50^{+1.02}_{-0.83}$	2000.0
1678	0.72	34.72	0.0	0.0	$90.22^{+5.77}_{-6.06}$	0.01	$12.99^{+1.37}_{-2.22}$
1677	5.04	33.0	0.0	0.0	$116.57^{+7.31}_{-6.05}$	0.01	$8.77^{+6.54}_{-5.24}$
N-ARC RIGHT	8.07	29.91	$0.33^{+0.18}_{-0.17}$	$88.31^{+36.18}_{-44.72}$	$61.98^{+4.04}_{-3.22}$	0.01	$4.56^{+2.05}_{-1.85}$
BK N-NW	12.31	25.96	0.0	0.0	$59.55^{+16.49}_{-14.67}$	0.01	$10.79^{+2.90}_{-5.17}$
BK NW LEFT	$12.95^{+0.51}_{-0.52}$	$24.47^{+0.30}_{-0.27}$	$0.64^{+.12}_{-0.12}$	$23.02^{+20.62}_{-18.83}$	$96.10^{+17.69}_{-16.83}$	0.01	$5.81^{+1.67}_{-2.04}$
1298	19.0	14.71	$0.89^{+0.01}_{-0.01}$	$54.63^{+1.10}_{-1.11}$	$93.23^{+6.29}_{-3.67}$	0.01	$10.20^{+2.92}_{-3.60}$
Scaling relations	N_{gal}=194	m^{ref}_{F160W}=17.71	α=0.31	$\sigma^{ref}_{LT} = 200.56^{+10.21}_{+8.00}$	β_{cut}=0.57	$r^{ref}_{cut} = 15.15^{+3.21}_{-2.39}$	γ=0.2

Tab. 5.7. Output parameters values from the optimization of Model 10. We report the median value of the parameter with errors corresponding to the [16-th, 84-th] percentiles from the marginalized posterior distribution.

$\mu = 10.8$, while the most magnified image is $5.1l$, with a factor $\mu = 80.8$. In the same figure, bottom panel, we show the projected mass-density map for our model. The displayed values are in units of $10^{12} M_{\odot}/\mathrm{kpc}^2$, and the color scale is set between 0 and 0.006.

Finally, in Fig.(5.29) we plot in magenta the cumulative projected total mass distribution of PSZ1-G311 as a function of the projected distance from the BCG, R. We also plot in green the contribution from the cluster member galaxies (i.e. sub-halos). For both profiles, we show the 1-σ uncertainties, computed by extracting parameters from the MCMC chains of 500 random realizations of the reference model. We note that the profile is better constrained at the radii corresponding to the multiple images positions, marked with black vertical lines. Furthermore, we show the ratio between the sub-halo and the total mass profile, f_M (blue line). The sub-halo contribution is small at all projected distances from the BCG. The cluster members contribute to less than 2% to the total mass near the BCG and, while the contribution increases at larger radii, it never exceeds 10% at $R > 100$kpc.

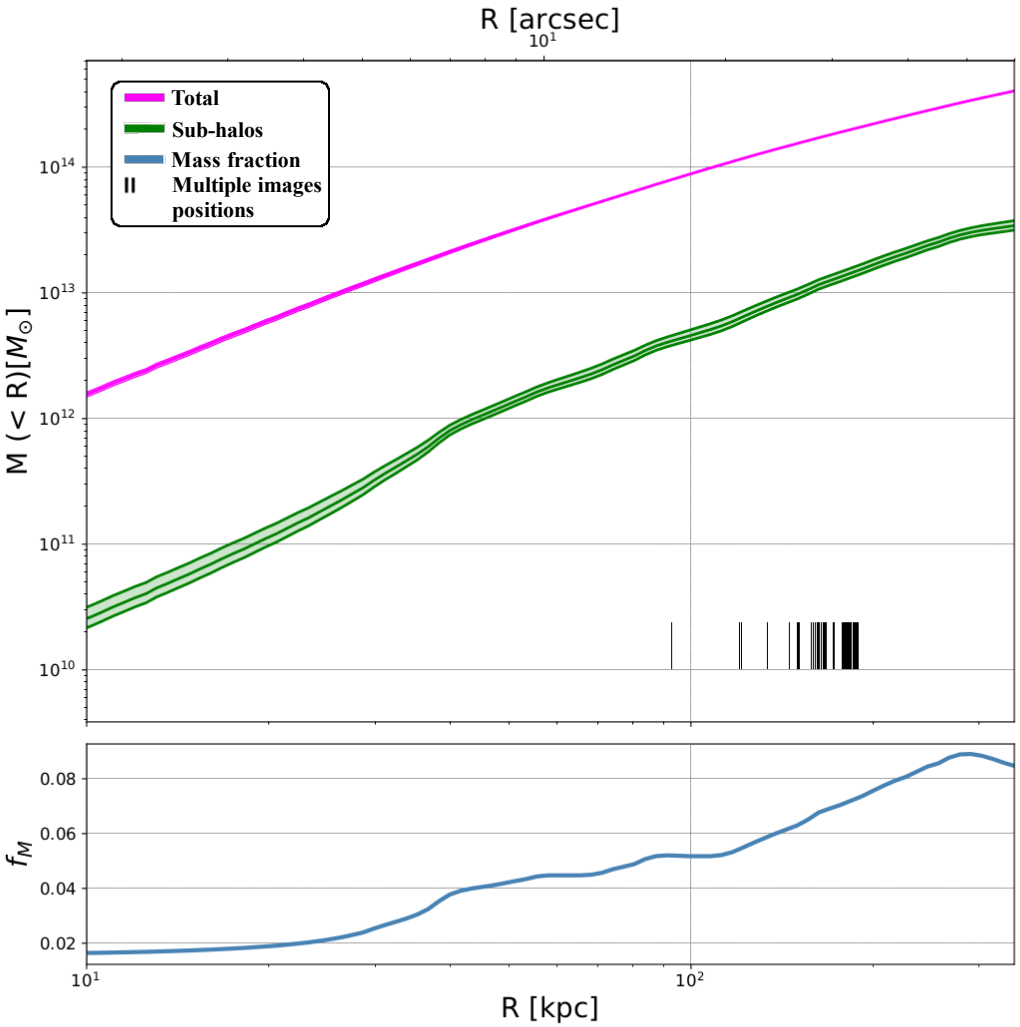

Fig. 5.29. *Top*: Projected cumulative mass profiles of PSZ1-G311 as a function of the projected distance from the BCG (R). In magenta we show the median value and the 1-σ confidence levels for the total mass profile. In green we show the median value and the 1-σ confidence levels for the sub-halo component mass profile. The positions of the multiple images are marked with black segments. *Bottom*: Fractional contribution of the cluster member galaxies mass to the total cumulative mass shown in blue.

5.4.1 Time-delay map

In Fig.(5.30) we show the time-delay map for the Sunburst Arc redshift computed from the Model 10 best fit. The map contains for each pixel the arrival time (in days) on the image plane of the photons that were emitted by a background source at the same time. As explained in Sec.(5.2.2), we exploited the best fit lens model and the observed position of the recently discovered transient object to predict its expected future appearances. Following our nomenclature, the family associated to the transient is family 5.16 and the observed image of the transient is 5.16d. This image is taken as reference to compute the relative arrival time of the other multiple images. The model-predicted multiple images of the transient are displayed with crosses in the two panels of Fig.(5.30). In Fig.(5.31) we show in detail the model-predicted positions of the transient object on the color-composite image of the cluster.

To obtain the time-delay ($\Delta\tau$) between two images of the same source we need to compute the difference

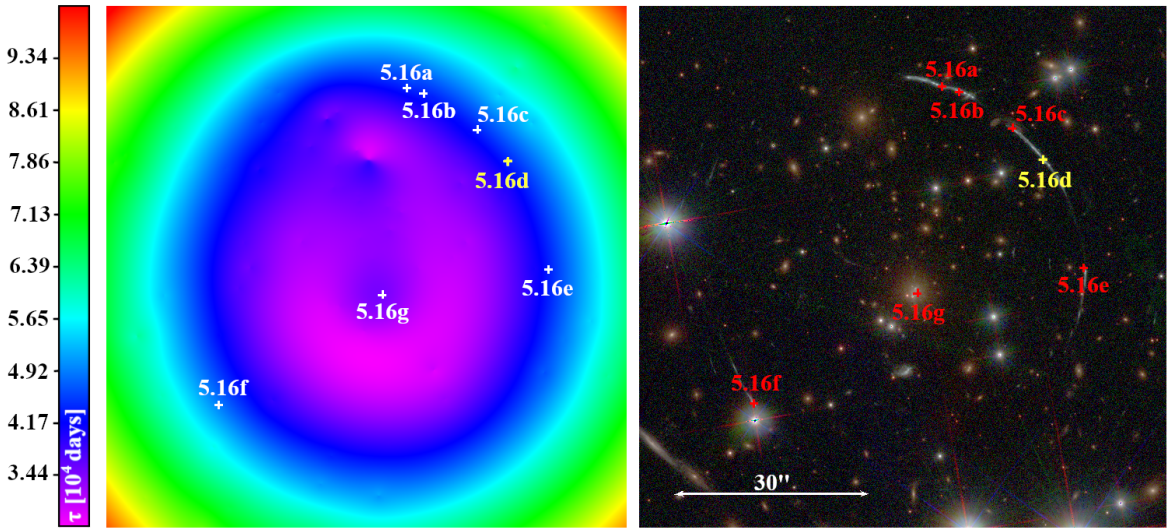

Fig. 5.30. *Left:* Time-delay map for the Sunburst Arc redshift. The map show for each pixel the time of arrival in units of 10^4 days. The position of the predicted multiple images of the transient object are marked in white, while the observed transient object is marked in yellow. *Right:* Color-composite image of PSZ1-G311 with the position of the model-predicted multiple images of the transient object in red.

in arrival time between the two images positions. Taking image 5.16d as the reference position, the time-delays in days within respect to the observed transient are reported in Tab.(5.8). If the resulting $\Delta\tau > 0$ the appearance of the image is in the future within respect to the appearance of 5.16d. If the resulting $\Delta\tau < 0$ the appearance of the image has already happened in the past. In Tab.(5.8) we also report the magnification factor (μ) of each model-predicted image. We note that image 5.16g, while representing a future appearance of the transient, will be highly de-magnified.

ID	RA [deg]	Dec [deg]	μ	$\Delta\tau$ [days]
5.16a	237.5237354	-78.1829845	12.99	-97
5.16b	237.5202441	-78.1832065	15.83	-106.5
5.16c	237.5090686	-78.1847487	14.90	17.3
5.16d	237.5027200	-78.1861004	55.97	-
5.16e	237.4942090	-78.1907123	14.59	122.7
5.16f	237.5630487	-78.1964912	13.73	-5614.2
5.16g	237.5289044	-78.1917973	0.36	8903.6

Tab. 5.8. Time-delay of the model-predicted images of the observed transient object (5.16d, yellow row). For each image we report the ID, the (RA,Dec) position, the magnification factor (μ) and the time-delay ($\Delta\mu$) in days within respect to the reference image 5.16d. A positive time-delay points to a future appearance, while a negative time-delay points to a past event.

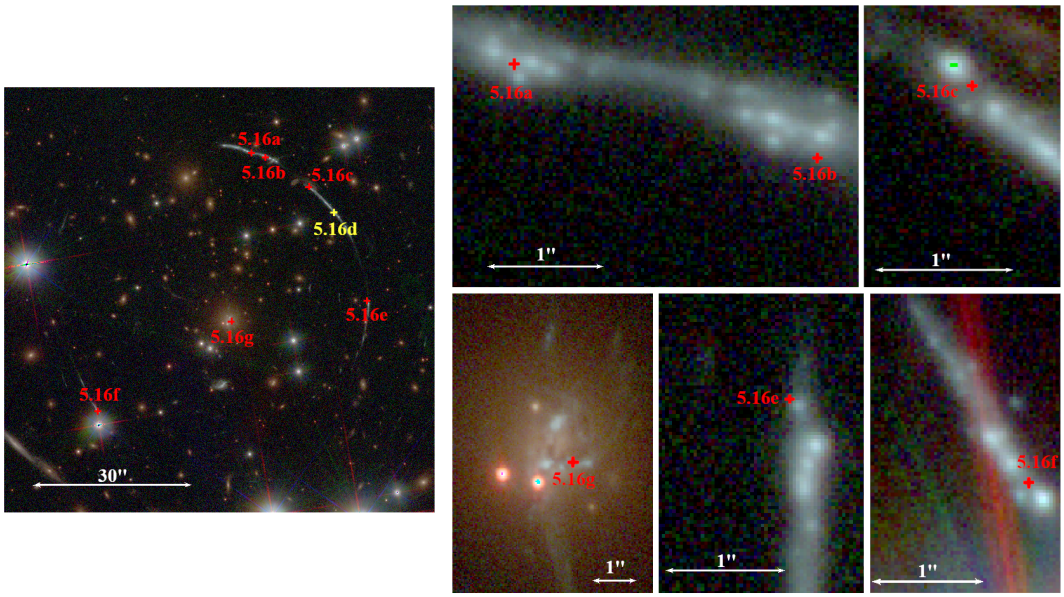

Fig. 5.31. Color-composite image of PSZ1-G311 with the position of the model-predicted multiple images of the transient object in red. In the different cut-outs we show in detail the position of the predicted appearances of the transient object along the arc.

Chapter 6

Conclusions

In this thesis we produced a lens model for the galaxy cluster PSZ1-G311. No other models of this clusters exists in the literature. Our work is based on HST imaging and MUSE spectroscopic observations. For the optimization of our models we used the parametric gravitational lensing software `LENSTOOL`. The construction of the lens model presented in this thesis can be summarized as follows:

- We identified the cluster member galaxies to construct a catalogue. To isolate the cluster red-sequence, first we conducted a photometric analysis on HST data using `SExtractor`; then, we cross-matched the photometric catalogue with the spectroscopically identified members. Our final catalogue contains 205 cluster member galaxies, of which 155 are photometrically identified with a magnitude value, 46 are also spectroscopically identified, and 4 are without a magnitude value.

- We measured the inner velocity dispersion of the spectroscopically identified cluster member galaxies to construct the $\sigma - L$ scaling relation for the galaxies. To extract the velocity dispersions we used the software `pPXF`. To ensure a robust measurement we limited the sample to 14 galaxies with resulting measured velocity dispersions higher than 60 km/s and a $\langle S/N \rangle$ higher than 10. This sample is used to fit the normalization and slope of the scaling relation that is incorporated in our lens model. We found a slope of $\alpha = 0.311^{+0.076}_{-0.079}$ and a normalization of $\sigma_{ap}^{ref} = 249.8^{+29.6}_{-28.8}$ km/s.

- We identified 6 important deflectors that are parametrized separately from the scaling relations. These contribute the most on the multiple images configurations.

- We constructed a multiple image catalogue. The identification and pairing of the several multiple images was performed inspecting the color pattern when possible, or by comparing similar

arrangements in position. The final catalogue used in our lens model contains a total of 76 multiple images coming from 5 background sources. We determined the spectroscopic redshift for 4 systems: $z_1 = 3.505$, $z_3 = 2.393$, $z_4 = 1.186$, $z_5 = 2.369$. Sys-2 redshift was optimized in our lens model, finding a mean value of $z_2 = 2.196^{+0.024}_{-0.023}$.

- We tested several different models of increasing complexity. We reported the 9 most significant models, which are the steps to the construction of the reference model. During testing, we found that we can exclude the BCG from the optimized halos and still reach a low Δ_{rms}. We also excluded Sys-1 due to unsure matching.

- We reported the results of our reference model. This model parametrizes 2 DH Halos and 6 Deflectors outside the scaling relations, and the multiple images catalogue includes a total of 62 images. It reaches a low Δ_{rms} of $0.14''$. We showed the total cumulative projected mass profile of the cluster as a function of the projected distance from the BCG. We found that the sub-halo component contribution to the total mass is small at all radii, never exceeding 10%.

- Finally, we produced the time-delay maps for the Sunburst Arc redshift. This allowed a prediction on the future appearances of a transient object discovered in the arc.

The Sunburst Arc configuration is the product of a rare alignment of observer, lens, and source along the line of sight, and it is a complex system to model. On the other hand, this extended arc stretches the shape of the source to the point where we could identify many different star-forming knots that are resolved and multiply-imaged. Such resolution at the source redshift is not reachable with our technology yet. The powerful strong lensing effect allows us to explore in detail both the background source and the mass-distribution of the lens, is the main motivation for the building a robust lens model for this galaxy cluster. Our lens model reproduces the observed multiple images with very high accuracy. Moreover, it will be extremely useful in future works to further characterize the physical properties of the source originating the Sunburst Arc, as well as to make further time-delay analysis on a recently discovered transient object.

The usage of galaxy clusters as cosmic telescopes, as illustrated in this thesis, allows to observe distant sources in the primordial Universe with a level of details that is not reachable from any existing facility.

Bibliography

[1] E. o. Vanzella, "Probing the circum-stellar medium 2.8 Gyr after the Big Bang: detection of Bowen fluorescence in the Sunburst arc," *Monthly Notices of the Royal Astronomical Society: Letters*, vol. 499, pp. L67–L71, Sept. 2020. arXiv: 2004.08400.

[2] S. L. Weinberg, "Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity | Wiley," 1972.

[3] Coles and Lucchin, "(PDF) Cosmology, The Origin and Evolution of Cosmic Structure."

[4] P. Collaboration *et al.*, "Planck 2018 results. VI. Cosmological parameters," *Astronomy & Astrophysics*, vol. 641, p. A6, Sept. 2020.

[5] W. L. Freedman, B. F. Madore, T. Hoyt, I. S. Jang, R. Beaton, M. G. Lee, A. Monson, J. Neeley, and J. Rich, "Calibration of the Tip of the Red Giant Branch," *The Astrophysical Journal*, vol. 891, p. 57, Mar. 2020.

[6] Wong *et al.*, "H0LiCOW XIII. A 2.4% measurement of H_0 from lensed quasars: 5.3σ tension between early and late-Universe probes," *Monthly Notices of the Royal Astronomical Society*, p. stz3094, June 2020.

[7] E. Macaulay *et al.*, "First cosmological results using Type Ia supernovae from the Dark Energy Survey: measurement of the Hubble constant," *Monthly Notices of the Royal Astronomical Society*, vol. 486, pp. 2184–2196, June 2019.

[8] G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S. S. Meyer, S. H. Moseley, T. Murdock, L. Rokke, R. F. Silverberg, L. Tenorio, R. Weiss, and D. T. Wilkinson, "Structure in the COBE Differential Microwave Radiometer First-Year Maps," , vol. 396, p. L1, Sept. 1992.

[9] J. H. Jeans and G. H. Darwin, "I. The stability of a spherical nebula," *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 199, pp. 1–53, Jan. 1902.

[10] "Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation," *http://adsabs.harvard.edu/abs/1974ApJ...187..425P*, vol. 187.

[11] R. Narayan and M. Bartelmann, "Lectures on Gravitational Lensing," *arXiv:astro-ph/9606001*, Oct. 1997.

[12] M. Meneghetti, "maxmen/LensingLectures ," July 2020.

[13] S. H. Suyu, "Lensing basics," in *Astrophysical Applications of Gravitational Lensing* (E. Mediavilla, F. Garzón, J. A. Muñoz, and T. J. Mahoney, eds.), Canary Islands Winter School of Astrophysics, pp. 1–30, Cambridge: Cambridge University Press, 2016.

[14] P. Schneider, J. Ehlers, and E. E. Falco, *Gravitational Lenses*. Astronomy and Astrophysics Library, Berlin Heidelberg: Springer-Verlag, 1992.

[15] G. B. Caminha *et al.*, "Mass distribution in the core of MACS J1206: Robust modeling from an exceptionally large sample of central multiple images," *Astronomy & Astrophysics*, vol. 607, p. A93, Nov. 2017.

[16] T. Treu *et al.*, ""REFSDAL" MEETS POPPER: COMPARING PREDICTIONS OF THE RE-APPEARANCE OF THE MULTIPLY IMAGED SUPERNOVA BEHIND MACSJ1149.5+2223," *The Astrophysical Journal*, vol. 817, p. 60, Jan. 2016.

[17] E. Jullo, P. Natarajan, J.-P. Kneib, A. D'Aloisio, M. Limousin, J. Richard, and C. Schimd, "Cosmological constraints from strong gravitational lensing in clusters of galaxies," *Science*, vol. 329, no. 5994, pp. 924–927, 2010.

[18] G. Golse, J. P. Kneib, and G. Soucail, "Constraining the cosmological parameters using strong lensing," , vol. 387, pp. 788–803, June 2002.

[19] "(PDF) Introduction to Gravitational Lensing - Lecture scripts."

[20] C. Jensen, *Theoretical Models for the Gravitational Lensing of a Quasar by Dwarf Galaxy*. PhD thesis, May 2012.

[21] R. Kormann, P. Schneider, and M. Bartelmann, "Isothermal elliptical gravitational lens models," *Astronomy and Astrophysics*, vol. 284, pp. 285–299, Apr. 1994.

[22] Elíasdóttir, M. Limousin, J. Richard, J. Hjorth, J.-P. Kneib, P. Natarajan, K. Pedersen, E. Jullo, and D. Paraficz, "Where is the matter in the Merging Cluster Abell 2218?," *arXiv:0710.5636 [astro-ph]*, Oct. 2007.

[23] A. Kassiola and I. Kovner, "Elliptic Mass Distributions versus Elliptic Potentials in Gravitational Lenses," *The Astrophysical Journal*, vol. 417, p. 450, Nov. 1993.

[24] M. Limousin, J.-P. Kneib, and P. Natarajan, "Constraining the mass distribution of galaxies using galaxy-galaxy lensing in clusters and in the field," *Mon Not R Astron Soc*, vol. 356, pp. 309–322, Jan. 2005. Publisher: Oxford Academic.

[25] J. F. Navarro, C. S. Frenk, and S. D. M. White, "The Structure of Cold Dark Matter Halos," *The Astrophysical Journal*, vol. 462, p. 563, May 1996.

[26] M. Bartelmann, "Arcs from a universal dark-matter halo profile.," *Astronomy and Astrophysics*, vol. 313, pp. 697–702, Sept. 1996.

[27] G. O. Abell, "The Distribution of Rich Clusters of Galaxies.," *The Astrophysical Journal Supplement Series*, vol. 3, p. 211, May 1958.

[28] N. A. Bahcall, "Clusters and superclusters of galaxies," *arXiv:astro-ph/9611148*, Nov. 1996. arXiv: astro-ph/9611148.

[29] J. E. Gunn and J. R. Gott, "On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution," *The Astrophysical Journal*, vol. 176, p. 1, Aug. 1972.

[30] L. L. Cowie and A. Songaila, "Thermal evaporation of gas within galaxies by a hot intergalactic medium," *Nature*, vol. 266, pp. 501–503, Apr. 1977.

[31] J. H. van Gorkom, "Interaction of Galaxies with the ICM," *arXiv:astro-ph/0308209*, Aug. 2003. arXiv: astro-ph/0308209.

[32] P. E. J. Nulsen, "Transport processes and the stripping of cluster galaxies," *Mon Not R Astron Soc*, vol. 198, pp. 1007–1016, Apr. 1982. Publisher: Oxford Academic.

[33] A. Dressler, "Galaxy morphology in rich clusters - Implications for the formation and evolution of galaxies," *The Astrophysical Journal*, vol. 236, pp. 351–365, Mar. 1980.

[34] A. Sandage, G. A. Tammann, and A. Yahil, "The velocity field of bright nearby galaxies. I - The variation of mean absolute magnitude with redshift for galaxies in a magnitude-limited sample," *The Astrophysical Journal*, vol. 232, pp. 352–364, Sept. 1979.

[35] D. W. Hogg, M. R. Blanton, J. Brinchmann, D. J. Eisenstein, D. J. Schlegel, J. E. Gunn, T. A. McKay, H.-W. Rix, N. A. Bahcall, J. Brinkmann, and A. Meiksin, "The dependence on environment of the color-magnitude relation of galaxies," *ApJ*, vol. 601, pp. L29–L32, Jan. 2004. arXiv: astro-ph/0307336.

[36] D. N. Spergel, L. Verde, H. V. Peiris, E. Komatsu, M. R. Nolta, C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright, "First Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters," *ASTROPHYS J SUPPL S*, vol. 148, pp. 175–194, Sept. 2003. arXiv: astro-ph/0302209.

[37] Y.-T. Lin, J. J. Mohr, and S. A. Stanford, "Near-Infrared Properties of Galaxy Clusters: Luminosity as a Binding Mass Predictor and the State of Cluster Baryons," *The Astrophysical Journal*, vol. 591, pp. 749–763, July 2003.

[38] C. L. Sarazin, "X-ray emission from clusters of galaxies," *Cambridge Astrophysics Series, Cambridge: Cambridge University Press, 1988*, 1988.

[39] R. K. Smith and J. P. Hughes, "IONIZATION EQUILIBRIUM TIMESCALES IN COLLISIONAL PLASMAS," *ApJ*, vol. 718, pp. 583–585, July 2010.

[40] M. Gitti, F. Brighenti, and B. R. McNamara, "Evidence for AGN Feedback in Galaxy Clusters and Groups," Jan. 2012. ISSN: 1687-7969 Pages: e950641 Publisher: Hindawi Volume: 2012.

[41] M. Gitti, R. Piffaretti, and S. Schindler, "Mass distribution in the most X-ray-luminous galaxy cluster RX J1347.5-1145 studied with *XMM-Newton*," *A&A*, vol. 472, pp. 383–394, Sept. 2007.

[42] L. M. Voigt and A. C. Fabian, "Galaxy cluster mass profiles," *Monthly Notices of the Royal Astronomical Society*, vol. 368, pp. 518–533, May 2006. arXiv: astro-ph/0602373.

[43] M. Gitti, B. R. McNamara, P. E. J. Nulsen, and M. W. Wise, "Cosmological Effects of Powerful AGN Outbursts in Galaxy Clusters: Insights from an XMM-Newton Observation of MS0735+7421," *ApJ*, vol. 660, pp. 1118–1136, May 2007. arXiv: astro-ph/0701386.

[44] A. Cavaliere and R. Fusco-Femiano, "X-rays from hot plasma in clusters of galaxies," *Astronomy and Astrophysics*, vol. 49, pp. 137–144, May 1976.

[45] I. King, "The structure of star clusters. I. an empirical density law," *The Astronomical Journal*, vol. 67, p. 471, Oct. 1962.

[46] S. Ettori and F. Brighenti, "On the evolution of cooling cores in X-ray galaxy clusters," *Monthly Notices RAS*, vol. 387, pp. 631–638, June 2008.

[47] A. D. Lewis, D. A. Buote, and J. T. Stocke, "*Chandra* Observations of A2029: The Dark Matter Profile Down to below 0.01 $r_{\rm vir}$ in an Unusually Relaxed Cluster," *ApJ*, vol. 586, pp. 135–142, Mar. 2003.

[48] F. Zwicky, "Republication of: The redshift of extragalactic nebulae," *Gen Relativ Gravit*, vol. 41, pp. 207–224, Jan. 2009.

[49] R. H. Sanders, "Clusters of galaxies with modified Newtonian dynamics," *Monthly Notices of the Royal Astronomical Society*, vol. 342, pp. 901–908, July 2003.

[50] G. M. Voit, "Tracing cosmic evolution with clusters of galaxies," *Rev. Mod. Phys.*, vol. 77, pp. 207–258, Apr. 2005. arXiv: astro-ph/0410173.

[51] S. W. Allen, A. E. Evrard, and A. B. Mantz, "Cosmological Parameters from Observations of Galaxy Clusters," , vol. 49, pp. 409–470, Sept. 2011.

[52] L. Gao, J. F. Navarro, C. S. Frenk, A. Jenkins, V. Springel, and S. D. M. White, "The Phoenix Project: the dark side of rich Galaxy clusters," , vol. 425, pp. 2169–2186, Sept. 2012.

[53] A. B. Newman, T. Treu, R. S. Ellis, D. J. Sand, C. Nipoti, J. Richard, and E. Jullo, "The Density Profiles of Massive, Relaxed Galaxy Clusters. I. The Total Density Over Three Decades in Radius," , vol. 765, p. 24, Mar. 2013.

[54] A. B. Newman, T. Treu, R. S. Ellis, and D. J. Sand, "The Density Profiles of Massive, Relaxed Galaxy Clusters. II. Separating Luminous and Dark Matter in Cluster Cores," , vol. 765, p. 25, Mar. 2013.

[55] Y. P. Jing, "The Density Profile of Equilibrium and Nonequilibrium Dark Matter Halos," *ApJ*, vol. 535, p. 30, May 2000. Publisher: IOP Publishing.

[56] J. F. Navarro, C. S. Frenk, and S. D. M. White, "A Universal Density Profile from Hierarchical Clustering," *The Astrophysical Journal*, vol. 490, pp. 493–508, Dec. 1997.

[57] M. Meneghetti *et al.*, "THE MUSIC OF CLASH: PREDICTIONS ON THE CONCENTRATION-MASS RELATION," *ApJ*, vol. 797, p. 34, Nov. 2014. Publisher: IOP Publishing.

[58] J. Einasto, "On the Construction of a Composite Model for the Galaxy and on the Determination of the System of Galactic Parameters," *Trudy Astrofizicheskogo Instituta Alma-Ata*, vol. 5, pp. 87–100, 1965.

[59] A. S. Bolton, S. Burles, L. V. E. Koopmans, T. Treu, R. Gavazzi, L. A. Moustakas, R. Wayth, and D. J. Schlegel, "The Sloan Lens ACS Survey. V. The Full ACS Strong-Lens Sample*," *ApJ*, vol. 682, p. 964, Aug. 2008. Publisher: IOP Publishing.

[60] S. T. Myers *et al.*, "The Cosmic Lens All-Sky Survey: I. Source selection and observations," *Monthly Notices of the Royal Astronomical Society*, vol. 341, pp. 1–12, May 2003. arXiv: astro-ph/0211073.

[61] I. W. A. Browne *et al.*, "The Cosmic Lens All-Sky Survey:II. Gravitational lens candidate selection and follow-up," *Monthly Notices of the Royal Astronomical Society*, vol. 341, pp. 13–32, May 2003. arXiv: astro-ph/0211069.

[62] C. Grillo *et al.*, "CLASH-VLT: INSIGHTS ON THE MASS SUBSTRUCTURES IN THE FRONTIER FIELDS CLUSTER MACS J0416.1–2403 THROUGH ACCURATE STRONG LENS MODELING," *ApJ*, vol. 800, p. 38, Feb. 2015.

[63] M. Postman *et al.*, "Cluster Lensing And Supernova survey with Hubble (CLASH): An Overview," *ApJS*, vol. 199, p. 25, Apr. 2012. arXiv: 1106.3328.

[64] H. Ferguson, *Hubble Deep Fields*, p. 2628. 2000.

[65] J. Lotz *et al.*, "The HST Frontier Fields," vol. 223, p. 254.01, Jan. 2014. Conference Name: American Astronomical Society Meeting Abstracts #223.

[66] A. M. Koekemoer *et al.*, "The HST Frontier Fields: Science Data Pipeline, Products, and First Data Release," vol. 223, p. 254.02, Jan. 2014. Conference Name: American Astronomical Society Meeting Abstracts #223.

[67] D. Coe *et al.*, "RELICS: Reionization Lensing Cluster Survey," *ApJ*, vol. 884, p. 85, Oct. 2019. arXiv: 1903.02002.

[68] C. L. Steinhardt *et al.*, "The BUFFALO HST Survey," , vol. 247, p. 64, Apr. 2020.

[69] J. Merten, M. Cacciato, M. Meneghetti, C. Mignone, and M. Bartelmann, "Combining weak and strong cluster lensing: applications to simulations and MS 2137," , vol. 500, pp. 681–691, June 2009.

[70] J. M. Diego, P. Protopapas, H. B. Sandvik, and M. Tegmark, "Non-parametric inversion of strong lensing systems," *Mon Not R Astron Soc*, vol. 360, pp. 477–491, June 2005. Publisher: Oxford Academic.

[71] D. Lam, "A New Approach to Free-Form Cluster Lens Modeling Inspired by the JPEG Image Compression Method," *PASP*, vol. 131, p. 114505, Nov. 2019. arXiv: 1906.00006.

[72] P. Natarajan and J.-P. Kneib, "Lensing by galaxy haloes in clusters of galaxies," *Mon Not R Astron Soc*, vol. 287, pp. 833–847, June 1997. Publisher: Oxford Academic.

[73] K. Sharon, M. D. Gladders, J. R. Rigby, E. Wuyts, B. P. Koester, M. B. Bayliss, and L. F. Barrientos, "Source-plane Reconstruction of the Bright Lensed Galaxy RCSGA 032727-132609," *The Astrophysical Journal*, vol. 746, p. 161, Feb. 2012.

[74] S. M. Faber and R. E. Jackson, "Velocity dispersions and mass-to-light ratios for elliptical galaxies," *The Astrophysical Journal*, vol. 204, pp. 668–683, Mar. 1976.

[75] M. Meneghetti *et al.*, "The Frontier Fields Lens Modeling Comparison Project," *Monthly Notices of the Royal Astronomical Society*, vol. 472, pp. 3177–3216, Dec. 2017. arXiv: 1606.04548.

[76] D. Clowe, M. Bradač, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky, "A Direct Empirical Proof of the Existence of Dark Matter," , vol. 648, pp. L109–L113, Sept. 2006.

[77] N. Kaiser and G. Squires, "Mapping the dark matter with weak gravitational lensing," *The Astrophysical Journal*, vol. 404, pp. 441–450, Feb. 1993.

[78] L. C. Parker, H. Hoekstra, M. J. Hudson, L. v. Waerbeke, and Y. Mellier, "The Masses and Shapes of Dark Matter Halos from Galaxy-Galaxy Lensing in the CFHT Legacy Survey," *ApJ*, vol. 669, p. 21, Nov. 2007. Publisher: IOP Publishing.

[79] T. G. Brainerd, R. D. Blandford, and I. Smail, "Weak Gravitational Lensing by Galaxies," *The Astrophysical Journal*, vol. 466, p. 623, Aug. 1996.

[80] P. Natarajan *et al.*, "Mapping substructure in the HST Frontier Fields cluster lenses and in cosmological simulations," *Monthly Notices of the Royal Astronomical Society*, vol. 468, no. 2, pp. 1962–1980, 2017. _eprint: https://academic.oup.com/mnras/article-pdf/468/2/1962/11210742/stw3385.pdf.

[81] M. Meneghetti *et al.*, "An excess of small-scale gravitational lenses observed in galaxy clusters," *Science*, vol. 369, no. 6509, pp. 1347–1351, 2020. Publisher: American Association for the Advancement of Science _eprint: https://science.sciencemag.org/content/369/6509/1347.full.pdf.

[82] Broadhurst *et al.*, "The Surprisingly Steep Mass Profile of A1689, from a Lensing Analysis of Subaru Images," *ApJ*, vol. 619, pp. L143–L146, Jan. 2005. Publisher: IOP Publishing.

[83] M. Bradač, P. Schneider, M. Lombardi, and T. Erben, "Strong and weak lensing united: I. The combined strong and weak lensing cluster mass reconstruction method," *A&A*, vol. 437, pp. 39–48, July 2005.

[84] D. Coe *et al.*, "LensPerfect: Gravitational Lens Mass Map Reconstructions Yielding Exact Reproduction of All Multiple Images," *The Astrophysical Journal*, vol. 681, pp. 814–830, July 2008.

[85] D. Coe, N. Benítez, T. Broadhurst, and L. A. Moustakas, "A HIGH-RESOLUTION MASS MAP OF GALAXY CLUSTER SUBSTRUCTURE: LensPerfect ANALYSIS OF A1689," *ApJ*, vol. 723, pp. 1678–1702, Oct. 2010. Publisher: IOP Publishing.

[86] J.-P. Kneib and P. Natarajan, "Cluster Lenses," *Astron Astrophys Rev*, vol. 19, p. 47, Dec. 2011. arXiv: 1202.0185.

[87] J.-P. Kneib, R. S. Ellis, I. Smail, W. J. Couch, and R. M. Sharples, "Hubble Space Telescope Observations of the Lensing Cluster Abell 2218," *ApJ*, vol. 471, pp. 643–656, Nov. 1996. arXiv: astro-ph/9511015.

[88] E. Jullo, J.-P. Kneib, M. Limousin, Elíasdóttir, P. J. Marshall, and T. Verdugo, "A Bayesian approach to strong lensing modelling of galaxy clusters," *New J. Phys.*, vol. 9, pp. 447–447, Dec. 2007. Publisher: IOP Publishing.

[89] E. Jullo and J.-P. Kneib, "Multiscale cluster lens mass mapping - I. Strong lensing modelling," *Monthly Notices of the Royal Astronomical Society*, vol. 395, pp. 1319–1332, May 2009.

[90] P. Bergamini *et al.*, "A new high-precision strong lensing model of the galaxy cluster MACS J0416.1-2403," *arXiv:2010.00027 [astro-ph]*, Sept. 2020. arXiv: 2010.00027.

[91] G. B. Caminha *et al.*, "CLASH-VLT: A highly precise strong lensing model of the galaxy cluster RXC J2248.74431 (Abell S1063) and prospects for cosmography," *A&A*, vol. 587, p. A80, Mar. 2016.

[92] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.

[93] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist.*, vol. 6, pp. 461–464, 03 1978.

[94] J. Skilling, "John Skilling / David MacKay: BayeSys / Nested sampling," 2004.

[95] C. Keeton, "Methods for strong lens modelling," in *Astrophysical Applications of Gravitational Lensing* (E. Mediavilla, F. Garzón, J. A. Muñoz, and T. J. Mahoney, eds.), Canary Islands Winter School of Astrophysics, pp. 213–250, Cambridge: Cambridge University Press, 2016.

[96] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of State Calculations by Fast Computing Machines," *J. Chem. Phys.*, vol. 21, pp. 1087–1092, June 1953. Publisher: American Institute of Physics.

[97] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, pp. 97–109, Apr. 1970. Publisher: Oxford Academic.

[98] R. Trotta, "Bayesian Methods in Cosmology," *arXiv e-prints*, p. arXiv:1701.01467, Jan. 2017.

[99] J. Ruanaidh and W. Fitzgerald, *Numerical Bayesian Methods Applied to Signal Processing*. Statistics and Computing, Springer New York, 1996.

[100] Planck Collaboration *et al.*, "Planck 2013 results. XXIX. The Planck catalogue of Sunyaev-Zeldovich sources," *Astronomy and Astrophysics*, vol. 571, p. A29, Nov. 2014.

[101] R. A. Sunyaev and Y. B. Zeldovich, "The Observations of Relic Radiation as a Test of the Nature of X-Ray Radiation from the Clusters of Galaxies," *Comments on Astrophysics and Space Physics*, vol. 4, p. 173, Nov. 1972.

[102] R. A. Sunyaev and I. B. Zeldovich, "Microwave background radiation as a probe of the contemporary structure and history of the universe," *Annual Review of Astronomy and Astrophysics*, vol. 18, pp. 537–560, 1980.

[103] H. o. Dahle, "Discovery of an exceptionally bright giant arc at z = 2.369, gravitationally lensed by the Planck cluster PSZ1 G311.65 18.48," *A&A*, vol. 590, p. L4, June 2016. Publisher: EDP Sciences.

[104] Planck Collaboration *et al.*, "Planck intermediate results. XXVI. Optical identification and redshifts of Planck clusters with the RTT150 telescope," *Astronomy and Astrophysics*, vol. 582, p. A29, Oct. 2015.

[105] Planck Collaboration *et al.*, "Planck intermediate results. XXXVI. Optical identification and redshifts of Planck SZ sources with telescopes at the Canary Islands observatories," *Astronomy and Astrophysics*, vol. 586, p. A139, Feb. 2016.

[106] T. E. Rivera-Thorsen *et al.*, "The Sunburst Arc: Direct Lyman escape observed in the brightest known lensed galaxy," *A&A*, vol. 608, p. L4, Dec. 2017.

[107] A. Morandi, S. Ettori, and L. Moscardini, "X-ray and Sunyaev-Zel'dovich scaling relations in galaxy clusters," *Monthly Notices of the Royal Astronomical Society*, vol. 379, pp. 518–534, Aug. 2007. arXiv: 0704.2678.

[108] C. Rodriguez-Gonzalvez, R. R. Chary, S. Muchovej, J. B. Melin, F. Feroz, M. Olamaie, and T. Shimwell, "CARMA observations of massive Planck -discovered cluster candidates at z $\gtrsim$ 0.5 associated with WISE overdensities: breaking the size-flux degeneracy," , vol. 464, pp. 2378–2395, Jan. 2017.

[109] E. Bertin and S. Arnouts, "SExtractor: Software for source extraction.," *Astronomy and Astrophysics Supplement Series*, vol. 117, pp. 393–404, June 1996.

[110] K. Robinson, *The P Cygni Profile and Friends*, pp. 119–125. New York, NY: Springer New York, 2007.

[111] X. Fan, C. Carilli, and B. Keating, "Observational Constraints on Cosmic Reionization," *Annual Review of Astronomy and Astrophysics*, vol. 44, no. 1, pp. 415–462, 2006. _eprint: https://doi.org/10.1146/annurev.astro.44.051905.092514.

[112] E. Zackrisson, A. K. Inoue, and H. Jensen, "THE SPECTRAL EVOLUTION OF THE FIRST GALAXIES. II. SPECTRAL SIGNATURES OF LYMAN CONTINUUM LEAKAGE FROM GALAXIES IN THE REIONIZATION EPOCH," *ApJ*, vol. 777, p. 39, Oct. 2013. Publisher: IOP Publishing.

[113] A. Bik *et al.*, "Super star cluster feedback driving ionization, shocks and outflows in the halo of the nearby starburst ESO 338-IG04," *A&A*, vol. 619, p. A131, Nov. 2018. Publisher: EDP Sciences.

[114] T. E. Rivera-Thorsen *et al.*, "Gravitational lensing reveals ionizing ultraviolet photons escaping from a distant galaxy," *Science*, vol. 366, pp. 738–741, Nov. 2019. arXiv: 1904.08186.

[115] E. Vanzella *et al.*, "Ionising the Intergalactic Medium by Star Clusters: The first empirical evidence," *arXiv:1904.07941 [astro-ph]*, Sept. 2019. arXiv: 1904.07941.

[116] I. S. Bowen, "THE EXCITATION OF THE PERMITTED O III NEBULAR LINES," *PASP*, vol. 46, p. 146, June 1934. Publisher: IOP Publishing.

[117] S. Refsdal, "The gravitational lens effect," *Monthly Notices of the Royal Astronomical Society*, vol. 128, p. 295, 1964.

[118] G. de Vaucouleurs, "Integrated Colors of Bright Galaxies in the u, b, V System.," , vol. 5, p. 233, Jan. 1961.

[119] B. P. Holden, S. A. Stanford, P. Eisenhardt, and M. Dickinson, "Evolution in the color-magnitude relation of early-type galaxies in clusters of galaxies atz $\simeq$ 1," *The Astronomical Journal*, vol. 127, pp. 2484–2510, may 2004.

[120] C. Lidman, P. Rosati, M. Tanaka, V. Strazzullo, R. Demarco, C. Mullis, N. Ageorges, M. Kissler-Patig, M. G. Petr-Gotzens, and F. Selman, "HAWK-I imaging of the X-ray luminous galaxy cluster XMMU J2235.3-2557. The red sequence at z = 1.39," , vol. 489, pp. 981–988, Oct. 2008.

[121] R. G. Kron, "Photometry of a complete sample of faint galaxies.," , vol. 43, pp. 305–325, June 1980.

[122] P. Bergamini *et al.*, "Enhanced cluster lensing models with measured galaxy kinematics," *A&A*, vol. 631, p. A130, Nov. 2019. arXiv: 1905.13236.

[123] M. Cappellari and E. Emsellem, "Parametric Recovery of Line-of-Sight Velocity Distributions from Absorption-Line Spectra of Galaxies via Penalized Likelihood," , vol. 116, pp. 138–147, Feb. 2004.

[124] M. Cappellari, "Improving the full spectrum fitting method: accurate convolution with Gauss-Hermite functions," , vol. 466, pp. 798–811, Apr. 2017.

[125] J. Goodman and J. Weare, "Ensemble samplers with affine invariance," *Communications in Applied Mathematics and Computational Science*, vol. 5, pp. 65–80, Jan. 2010.

[126] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, "emcee: The MCMC Hammer," , vol. 125, p. 306, Mar. 2013.

[127] S. M. Faber, A. Dressler, R. L. Davies, D. Burstein, D. Lynden Bell, R. Terlevich, and G. Wegner, "Global Scaling Relations for Elliptical Galaxies and Implications for Formation," in *Nearly Normal Galaxies. From the Planck Time to the Present* (S. M. Faber, ed.), p. 175, Jan. 1987.

[128] R. Bender, D. Burstein, and S. M. Faber, "Dynamically Hot Galaxies. I. Structural Properties," , vol. 399, p. 462, Nov. 1992.

Ringraziamenti

Desidero ringraziare innanzitutto il relatore di questa tesi, il professor Massimo Meneghetti, per la disponibilità e la professionalità con cui ha seguito la stesura del mio lavoro. Ci terrei a ringraziarla anche per la cura dedicata al suo corso di *Lensing Gravitazionale*, che ha saputo perfettamente creare in me un grande interesse e passione per questo argomento. Ringrazio il mio correlatore, il dottor Pietro Bergamini, per tutto il tempo e l'aiuto dedicatomi durante il lavoro, sempre con grande gentilezza e incoraggiamento.

Ringrazio tutta la mia famiglia che, tra chi mi ha sostenuto negli anni di università in una città nuova e chi mi ha incoraggiata da lontano, mi ha accompagnato affettuosamente in tutto il mio percorso. Grazie di cuore ai miei genitori e le mie sorelle, il mio fondamento. A Marina e Andrea, la mia seconda casa. A Gloria, perchè la tua presenza è stata per me indispensabile. Un particolare pensiero va ai miei carissimi nonni, affianco a me con amorevolezza in ogni passo e decisione.

Un grazie sincero alle mie amiche di sempre, Elisabetta, Martina e Luisa, per avermi dimostrato che anche chi è lontano può essere vicino. Ringrazio anche un grande amico che per fortuna è molto vicino, Francesco, con cui ho felicemente condiviso ogni singola risata e lacrima di questi anni.

Concludo ringraziando Lorenzo, il mio punto di riferimento fisso. Senza di te, non sarei dove sono oggi.

Grazie a tutti.