UNIVERSITÀ DEGLI STUDI DI MILANO

FACOLTÀ DI SCIENZE E TECNOLOGIE

CORSO DI LAUREA TRIENNALE IN FISICA

THE IMPACT OF SUPERMASSIVE BLACK HOLES IN STRONG LENSING MODELS OF GALAXY CLUSTERS

Relatore:
Prof. Claudio GRILLO
Correlatore:
Dr. Pietro BERGAMINI

Tesi di Laurea di:
Guglielmo DE TOMA
Matricola: 967168

DEDICATION

To Jupiter, Neptune, and Pluto.

Acknowledgments

Contents

Contents—*Continued*

List of Figures

List of Figures—*Continued*

ABSTRACT

In the strong gravitational lensing model of a galaxy cluster, the total mass distribution of the cluster normally includes three main mass components: extended dark-matter halos, extended hot-gas halos, and sub-halos corresponding to the cluster galaxies. In this thesis, we have also considered supermassive black holes located at the centers of the cluster galaxies. The aim of this thesis is to investigate the impact, in strong lensing models of galaxy clusters, of the presence of supermassive black holes at the centers of cluster galaxies. As a representative example, we have used the galaxy cluster RXC J2248.7−4431 (also known as Abell S1063), at a redshift of $z = 0.348$, studied by Bergamini et al. (2019). The masses of the black holes have been estimated using the observed $M_\bullet - \sigma$ scaling relation (McConnell & Ma 2013), which relates the masses ($M_\bullet$) of black holes to the velocity dispersions (σ) of their host galaxies. We have used the public software Lenstool (Kneib et al. 1996, Jullo et al. 2007) to quantify the differences, between mass models that include or ignore the presence of supermassive black holes. In particular, we have focused on the positions of the model-predicted multiple images and on the values, estimated through a strong lensing analysis, of the parameters describing the total mass distribution of the lens. We have concluded that the presence of supermassive black holes in the strong lensing model of Abell S1063 can cause the disappearance of the radial caustics (and the critical curves) produced by cluster galaxies, and it can modify the number of model-predicted multiple images from specific positions of the sources located near the caustics produced by the cluster galaxies. Moreover, we have found that the effect on the overall model precision of the presence of supermassive black holes in a strong lensing model is negligible compared to other simplifying assumptions in the model, such as neglected dark-matter clumps and unconsidered mass structures along the line of sight. In conclusion, we have observed that strong lensing models of galaxy

clusters that ignore the presence of supermassive black holes, predict more compact cluster galaxies, if the latter actually contain central black holes.

Chapter 1

Strong Gravitational Lensing

A light beam is deflected in the presence of a gravitational field. This physical phenomenon is accurately described in the theory of General Relativity (A. Einstein, 1915). According to the theory, a large mass can warp the space-time and, consequently, bend the light geodesics.

It is possible to observe the deflection of light when a massive object (e.g. a galaxy cluster or a galaxy) is located between a bright luminous source (e.g. a galaxy or a QSO) and the observer. In this case, the light rays are emitted from the source, then they are bent in the proximity of the mass and, finally, they reach the observer. Looking through the telescope, the observed image of the source is different from what it would have been in the absence of the deflecting mass. Its position, magnitude, shape, arrival time (the time that a source light ray takes to reach the observer) can change. Also, more than one image of the same background source can occur. These types of physical phenomena are often referred to by the name *gravitational lensing*. In particular, *strong* gravitational lensing studies the astrophysical systems where the light deflection leads to the occurrence of multiple images of a source. Gravitational lensing permits the study the properties of the deflecting mass (called *lens*) and of the background source. It also provides a method to estimate cosmological parameters from the positions of many observed multiple images at different redshifts (Caminha et al. 2015) or from the measured time delays between the images of a same time-variable source (Grillo et al. 2020, Suyu et al. 2018). Moreover, sometimes, the deflecting mass acts as a telescope in the sense that it magnifies sources that would have been too faint to be observed (E. Vanzella et al. 2012, Mestric et al. 2022, Roberts-Borsani et al. 2023).

This first chapter will discuss the theory of *strong gravitational lensing*. Sect. 1.1 will deal with the deflection angle of a light ray in the presence of a gravitational field. Then, in Sect. 1.2, the lens equation will be derived, showing the analogy between the deflecting mass and a lens. After defining some important quantities in Sect. 1.3, Sect. 1.4 will examine the distortion and the magnification, and will classify the different types of images. In Sect. 1.5, the critical curves and the caustics will be introduced. Finally, Sect. 1.6 will illustrate a simple category of lens models. The following sections refer mainly to the book entitled "Gravitational Lenses" by P. Schneider, 1992.

1.1   The deflection angle

The Fermat's principle applied on a curved space-time implies that the geodesic of a photon, with the initial and the final points fixed, satisfies the following equation:

$$\delta\tau = 0, \tag{1.1.1}$$

where τ is the time that the light needs to go from the initial fixed point to the final one. The Eq. (1.1.1) means that the null geodesic (light geodesic) corresponds to the path for which the time τ is stationary under variations of the smooth null curves with fixed end points. In geometrical optics, the time passed between the two fixed points is usually written with respect to the refraction index n:

$$\tau(\boldsymbol{x}(l)) = \int_a^b \frac{n(\boldsymbol{x}(l))}{c}\, dl \;, \tag{1.1.2}$$

where $\boldsymbol{x}(l)$ is a curve with fixed end points $\boldsymbol{x}(a)$, $\boldsymbol{x}(b)$ and the refraction index is

$$n(\boldsymbol{x}(l)) = \frac{c}{c'(\boldsymbol{x}(l))}, \tag{1.1.3}$$

c' is the actual speed of light along the curve $\boldsymbol{x}(l)$ and c is the speed of light in vacuum. The typical setup of a gravitational lensing system consists of a source of

light (the first fixed point), an observer (the second fixed point) and, in between, a mass distribution (the lens). The latter is supposed to have a velocity relative to the rest of the system much smaller than c. In order to apply the Fermat's principle, we have to make a couple of assumptions: firstly, the Newtonian gravitational potential[1] ϕ_g of the lens must be much smaller than c^2; Secondly, the deflection of the light ray[2] must occur in a region close to the lens, whose extent is comparable with the value of the smallest distance between the center of the lens and the light ray. Now, the gravitational potential perturbs[3] the Minkowski's metric in this manner:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} = \begin{pmatrix} -(1 + 2\frac{\phi_g}{c^2}) & 0 & 0 & 0 \\ 0 & (1 - 2\frac{\phi_g}{c^2}) & 0 & 0 \\ 0 & 0 & (1 - 2\frac{\phi_g}{c^2}) & 0 \\ 0 & 0 & 0 & (1 - 2\frac{\phi_g}{c^2}) \end{pmatrix}, \quad h_{\mu\nu} << 1,$$

(1.1.4)

where $h_{\mu\nu}$ is the small perturbation of the Minkoskian metric. Since, the line element for a light ray vanishes: $ds^2 = g_{\mu\nu}dx^\mu dx^\nu = 0$, it is straightforward to compute the velocity of light $c' := \frac{|d\boldsymbol{x}|}{dt}$:

$$c' = \frac{|d\boldsymbol{x}|}{dt} = c\sqrt{\frac{1 + 2\frac{\phi_g}{c^2}}{1 - 2\frac{\phi_g}{c^2}}} \sim c\left(1 + 2\frac{\phi_g}{c^2}\right), \quad \frac{\phi_g}{c^2} << 1.$$

(1.1.5)

In the last passage, the assumption on the weakness of ϕ_g is used. Accordingly, the refraction index n is

$$n = \frac{c}{c'} = \frac{1}{1 + 2\frac{\phi_g}{c^2}} \sim 1 - 2\frac{\phi_g}{c^2}, \quad \frac{\phi_g}{c^2} << 1.$$

(1.1.6)

Here, it can be noticed that n does not depend on the wavelength of the photon. In fact, gravitational lensing is said to be *achromatic*. As a result, geodesics of light

---

[1] In the case of galaxy clusters (the type of lens discussed in this Thesis), $\phi_g << 1$. If we assume the mass distribution to be in virial equilibrium, ϕ_g has the same order of magnitude of v^2. Therefore, $\frac{\phi_g}{c^2} \sim \frac{v^2}{c^2} << 1$ for a cluster where its galaxies have a typical speed of 1000 km/s.

[2] In gravitational lensing theory, light rays are assumed to propagate along straight lines as in geometrical optics.

[3] The metric shown in Eq. (1.1.4) is derived by linearizing in $h_{\mu\nu}$ the Einstein equation.

rays with different frequencies but with the same source (and observer) do not differ. Moreover, since $\phi_g \leq 0$ and consequently $n \geq 1$, the velocity of light under the influence of a gravitational field is smaller than the one in vacuum.

After having found n, the time τ a photon spends going from the source to the observer must be varied with respect to the null curves with fixed end points (at the source and at the observer). Then, the actual path of the photon will be the one for which τ is stationary under variations of these curves

$$\delta\tau = \delta \int_a^b \frac{n(\boldsymbol{x}(l))}{c}\, dl \ = 0. \tag{1.1.7}$$

This variational problem can be solved finding the solutions of the following Euler equation:

$$\frac{d}{d\lambda}(n\boldsymbol{e}) - \nabla n = 0, \tag{1.1.8}$$

where $\boldsymbol{e}$ is the tangent unit vector to the light path and λ is a suitable parameter such that the following relation holds:

$$\boldsymbol{e} := \frac{d\boldsymbol{x}}{|d\boldsymbol{x}|} = \frac{d\boldsymbol{x}}{d\lambda}. \tag{1.1.9}$$

From Eq. (1.1.8), we can find a simple expression for $\dot{\boldsymbol{e}} := \frac{d\boldsymbol{e}}{d\lambda}$,

$$n\dot{\boldsymbol{e}} = \nabla n - \boldsymbol{e} \cdot (\nabla n \cdot \boldsymbol{e}) = \nabla_\perp n,$$

and finally,

$$\dot{\boldsymbol{e}} = \nabla_\perp \ln(n). \tag{1.1.10}$$

The term $\nabla_\perp n$ is the component of the gradient of n orthogonal to the light path. Recalling the expression of n (1.1.6) for $\frac{\phi_g}{c^2} << 1$,

$$\dot{\boldsymbol{e}} \sim -\frac{2}{c^2}\nabla_\perp \phi_g. \tag{1.1.11}$$

Finally, in order to find the deflection angle, $\dot{\boldsymbol{e}}$ must be integrated along the curve

$$\hat{\boldsymbol{\alpha}} = \frac{2}{c^2} \int_{\lambda_a}^{\lambda_b} \nabla_\perp \phi_g \, d\lambda. \tag{1.1.12}$$

Until this point, we have used a potential of a generic mass distribution. Now, we focus on the case of a point mass. To calculate the integral in Eq. (1.1.12) for a point-like lens, we apply the Born approximation[4] and, thus, compute the integral along the unperturbed light trajectory (the one in the absence of the gravitational field). This approximation is justified by the assumptions made above Eq. (1.1.4) which ensure that the deflection angle is small. Therefore, the deflection angle is

$$\hat{\boldsymbol{\alpha}} = \frac{2}{c^2} \int_{-\infty}^{+\infty} \nabla_\perp \phi_g \, dz, \tag{1.1.13}$$

where the z-axis is taken such that it lies along the unperturbed orbit of light (which is a straight line). From Eq. (1.1.13), the deflection angle corresponding to the point-like lens (with a potential of $\phi_g(r) = -\frac{GM}{r}$) can be obtained:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4GM}{c^2} \frac{\boldsymbol{\xi}}{|\xi|^2}, \tag{1.1.14}$$

where $\boldsymbol{\xi}$ is the *impact* vector. Its modulus is equal to the minimum distance between the point mass and the unperturbed light path and its direction is orthogonal to the unperturbed path.

Now, it is possible to make a step forward and find the deflection angle when the lens consists of a set of N point masses which lies on the same plane (the *lens plane*). In this case, the total deflection angle is a linear superposition of the deflection angles of every point mass. This holds because the considered metric $g_{\mu\nu}$ (see Eq. (1.1.4)) is linear in ϕ_g and so is the deflection angle as it can be seen in Eq. (1.1.12). Besides, ϕ_g is a *Newtonian* gravitational potential, therefore, the superposition principle is valid.

Finally, we can treat a generic deflecting mass distribution. To deal with this, an other assumption is needed (besides the ones made for the point mass): the lens has to be *thin*. The extent L of the lens in the direction of the incoming ray has to be small

---

[4]The Born approximation is a result of the first order perturbation theory.

enough that the gravitational force experiences a negligible variation (in the region where the deflection occurs) if calculated on the real orbit or on the unperturbed one. Practically, the extent of the lens[5] must be much smaller than the distances between the source and the lens and between the lens and the observer (see Fig. 1.1). If the mass distribution is thin, the Born approximation can be applied and the relation (1.1.13) is valid. Thanks to the thin-lens approximation, every infinitesimal cylinder of mass parallel to the non-deflected light ray has the same impact vector and is affected by the same gravitational force. Therefore, it is a good approximation to project all the cylinders on the same plane (the *lens plane*). This condition is now similar to the previous case of the set of point masses on a plane. The only difference is that, here, there are not points but two-dimensional mass elements $dm := \Sigma(x)d^2x$ (where d^2x is the surface element and Σ is the projected surface mass density at x). Thus,

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4G}{c^2} \int_{\mathbb{R}^2} \frac{(\boldsymbol{\xi} - \boldsymbol{\xi}')\Sigma(\boldsymbol{\xi}')}{|\boldsymbol{\xi} - \boldsymbol{\xi}'|^2} \, d^2\xi', \qquad (1.1.15)$$

where $\boldsymbol{\xi}'$ is the two-dimensional impact vector of the mass element $\Sigma(\boldsymbol{\xi}')d^2\xi'$ in the lens plane.

1.2 Lens equation

The typical gravitational lensing system is displayed in Fig. 1.1. The aim of this section is to derive the *lens equation* of a gravitational lensing system with a point-like source, a generic deflecting mass distribution and an observer. The source lies on the *source plane* and the lens on the *lens plane*. The *optical axis* is a straight line that connects the observer and a certain reference point in the lens plane. The origins of the frames of reference in the lens plane and in the source plane are identified by the points of intersection between the planes and the optical axis. Having set these

---

[5]When the lens is a galaxy cluster, the lens length scale is $\sim Mpc$, while, the order of magnitude of the distance between the source and the lens or between the lens and the observer is $\sim Gpc$. Therefore, a galaxy cluster is a *thin* lens.

15
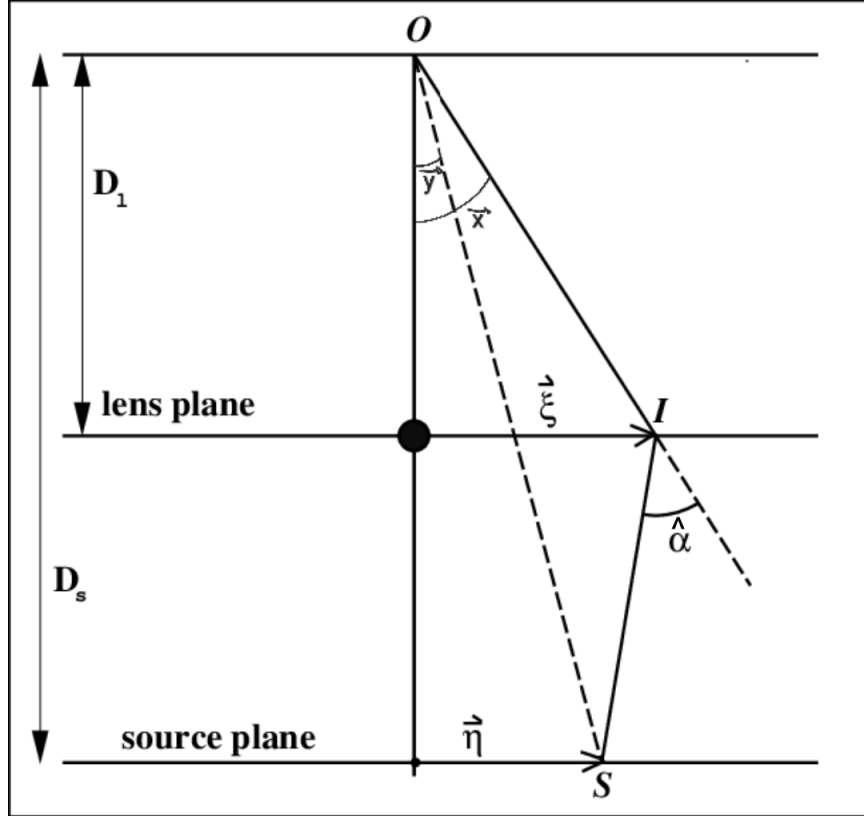
FIGURE 1.1. Schematic representation of a gravitational lensing system. (Figure from S. Frittelli, T. P. Kling and E. T. Newman, Image distortion from optical scalars in nonperturbative gravitational lensing, Phys. Rev. D 63.2: 023007 (2000)).

properties, $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are vectors of the lens and source planes, respectively. $\boldsymbol{\xi}$ indicate the position of light in the lens plane, while, $\boldsymbol{\eta}$ locates the source in the source plane. D_{LS} is the distance between the gravitational lens and the source, D_L is the distance from the observer to the lens, and D_S is the distance between the observer and the source. These are angular-diameter distances which differ from luminosity distances in Cosmology. It is Cosmology that forces us to use three distances instead of two, because $D_S \neq D_L + D_{LS}$. $\boldsymbol{y}$ is the source position angle in the absence of the lens and $\boldsymbol{x}$ is the image position angle (where the image of the source is observed). $\hat{\boldsymbol{\alpha}}$ is the deflection angle. These three angles are angular vectors that locate points in the sky. In a gravitational lensing system, the angles $\boldsymbol{x}$, $\boldsymbol{y}$ and $\hat{\boldsymbol{\alpha}}$ are assumed to be

small. This assumption is supported by the hypotheses made in Sect. 1.1, namely *weak gravitational field*, *thin lens*, and *non-relativistic lens* (velocity of the deflecting mass much smaller than the speed of light in vacuum). This permits to write the following relations between angles and distances:

$$\boldsymbol{\xi} = D_L \boldsymbol{x}, \tag{1.2.1}$$

$$\boldsymbol{\eta} = D_S \boldsymbol{y}. \tag{1.2.2}$$

And, after some geometrical considerations, we can obtain the lens equation (also known as ray-tracing equation):

$$\boldsymbol{y} = \boldsymbol{x} - \frac{D_{LS}}{D_S} \hat{\boldsymbol{\alpha}}(\boldsymbol{x} D_L). \tag{1.2.3}$$

Defining the scaled deflection angle

$$\boldsymbol{\alpha}(\boldsymbol{x}) := \frac{D_{LS}}{D_S} \hat{\boldsymbol{\alpha}}(\boldsymbol{x} D_L), \tag{1.2.4}$$

the equation can be written in a cleaner version:

$$\boldsymbol{y} = \boldsymbol{x} - \boldsymbol{\alpha}(\boldsymbol{x}). \tag{1.2.5}$$

1.3 Lensing potentials and time delay

This section will define and discuss some crucial quantities in gravitational lensing theory.

The dimensionless surface mass density[6] (or *convergence*) is defined as:

$$\kappa(\boldsymbol{x}) := \frac{\Sigma(\boldsymbol{x} D_L)}{\Sigma_{cr}}, \tag{1.3.1}$$

where the *critical surface mass density* Σ_{cr} is

$$\Sigma_{cr} := \frac{c^2}{4\pi G} \frac{D_S}{D_L D_{LS}}. \tag{1.3.2}$$

---

[6]κ is non-negative because it is related to a surface mass density.

The deflection angle can be expressed in terms of $\kappa(\boldsymbol{x})$ as

$$\boldsymbol{\alpha}(\boldsymbol{x}) = \frac{1}{\pi} \int_{\mathbb{R}^2} d^2 x' \kappa(\boldsymbol{x}') \frac{\boldsymbol{x} - \boldsymbol{x}'}{|\boldsymbol{x} - \boldsymbol{x}'|^2}. \tag{1.3.3}$$

Taking a closer look at the preceding Eq. (1.3.3), it can be noticed that $\boldsymbol{\alpha}(\boldsymbol{x})$ is equal to the gradient of the following function:

$$\psi(\boldsymbol{x}) := \frac{1}{\pi} \int_{\mathbb{R}^2} d^2 x' \kappa(\boldsymbol{x}') \ln |\boldsymbol{x} - \boldsymbol{x}'|, \tag{1.3.4}$$

$$\boldsymbol{\alpha}(\boldsymbol{x}) = \nabla_{\boldsymbol{x}} \psi(\boldsymbol{x}). \tag{1.3.5}$$

$\psi(\boldsymbol{x})$ is called *lensing potential*[7] . The relation (1.3.5) suggests defining the *Fermat potential*

$$\phi(\boldsymbol{x}, \boldsymbol{y}) := \frac{1}{2} |\boldsymbol{x} - \boldsymbol{y}|^2 - \psi(\boldsymbol{x}), \tag{1.3.6}$$

in order to obtain the lens equation in the following form:

$$\nabla_{\boldsymbol{x}} \phi(\boldsymbol{x}, \boldsymbol{y}) = 0. \tag{1.3.7}$$

From (1.3.4), another important relation between $\kappa(\boldsymbol{x})$ and $\psi(\boldsymbol{x})$ can be derived[8]

$$\nabla_{\boldsymbol{x}}^2 \psi(\boldsymbol{x}) = 2\kappa(\boldsymbol{x}). \tag{1.3.8}$$

Combing two ingredients, it can be shown that Eq. (1.3.8) is equivalent to Poisson's equation in two dimensions. Firstly, $\Sigma(\boldsymbol{x})$ is the surface mass density projected on the lens plane and, secondly, the Poisson equation is $\nabla^2 \phi_g(\boldsymbol{x}, z) = 4\pi G \rho(\boldsymbol{x}, z)$ (with ϕ_g the Newtonian potential of the mass, ρ the three-dimensional mass density and z-axis orthogonal to the lens plane). These two elements lead to:

$$\Sigma(\boldsymbol{x}) = \frac{1}{4\pi G} \int_{-\infty}^{+\infty} \nabla^2 \phi_g(\boldsymbol{x}, z) \, dz. \tag{1.3.9}$$

Now, this expression, together with Eq. (1.3.1), proves the statement.

---

[7]To obtain Eq. (1.3.5), it has been employed $\nabla_{\boldsymbol{x}} \ln(|\boldsymbol{x}|) = \frac{\boldsymbol{x}}{|\boldsymbol{x}|^2}$.
[8]To find Eq. (1.3.8), it has been used $\nabla_{\boldsymbol{x}}^2 \ln(|\boldsymbol{x}|) = 2\pi \delta^2(\boldsymbol{x})$.

The quantities defined in this section are fundamental to study the gravitational lensing systems. For example, the Fermat potential is exploited in the calculation of the time delay between a deflected light ray and the non-deflected ray (without the lens). In fact, from the Fermat's principle, we can compute the time that a light ray needs to go from a source, through the lens gravitational field, to the observer. This travel time t is the sum of two contributions:

$$t(\boldsymbol{x}, \boldsymbol{y}) = t_{geom}(\boldsymbol{x}, \boldsymbol{y}) + t_{grav}(\boldsymbol{x}), \qquad (1.3.10)$$

where t_{geom} is the time that the light would spend to go from the source at $\boldsymbol{y}$ in the source plane to $\boldsymbol{x}$ in the lens plane and from $\boldsymbol{x}$ to the observer position if it had a speed equal to the speed of light in vacuum. Whereas, t_{grav} (also known as *Shapiro delay*) accounts for the relativistic time dilation due to the gravitational field of the lens. The expression of time delay, namely the difference between the travel time (see Eq. (1.3.10)) of the real light path and the time spent by the unlensed light ray travelling from the source to the observer is equal to:

$$\Delta t(\boldsymbol{x}, \boldsymbol{y}) = \frac{D_L D_S}{D_{LS}} \frac{1 + z_l}{c} [\phi(\boldsymbol{x}, \boldsymbol{y})] + const. \qquad (1.3.11)$$

where z_l is the redshift of the lens plane.

In the next section, we will discuss further applications of the deflection potential ψ, the Fermat potential ϕ, and the convergence κ.

1.4 Magnification, distortion and classification of images

Not only does the direction of a light ray change in a gravitational field, but also the cross-section of a bundle of light rays does. This cross-section is different when the bundle of light rays is on the source plane or on the lens plane. In fact, an observer will see the image of a source having a different area, A_l, from the real intrinsic area of the source, A_s.

Gravitational lensing is not affected by emission or absorption phenomena, so it can be demonstrated that the *surface brightness*[9], I, is preserved along a light ray. Furthermore, the *surface brightness* of the image of a source (the observed I) is the same as the surface brightness of the source in the absence of the deflecting mass.[10]

Now, we can define the *magnification* μ of an image as the ratio between the flux of the image F and the flux of its source without the lens F_0:

$$\mu := \frac{F}{F_0} = \frac{d\omega}{d\omega_0}, \tag{1.4.1}$$

where $d\omega$ is the solid angle subtended by the image and $d\omega_0$ is subtended by the source in the absence of the lens. In the second equivalence, it has been used the fact that the flux of the image of an infinitesimal source is equal to the product of its surface brightness I and its solid angle[11] ($F = Id\omega$).

To study the distortion of images, it is convenient to *locally* linearize the lens map. This can be done if the image is small compared to the length scale on which the physical properties of the lens vary.

$$\boldsymbol{y} - \boldsymbol{y}_0 = A(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) + o(\|\boldsymbol{x} - \boldsymbol{x}_0\|), \tag{1.4.2}$$

$A(\boldsymbol{x}_0)$ is the *Jacobian* of the lens map $\boldsymbol{x} \longmapsto \boldsymbol{y}$ at the point $\boldsymbol{x}_0$. Before examining this matrix, we recall that the Jacobian determinant of a function outlines how a surface is distorted by that map. Thus, the determinant of the inverse[12] of the Jacobian at $\boldsymbol{x}_0$ is equal to the magnification (see Eq. (1.4.1)) at that point:

$$det A^{-1}(\boldsymbol{x}_0) = \frac{1}{det A(\boldsymbol{x}_0)} = \mu(\boldsymbol{x}_0). \tag{1.4.3}$$

---

[9]Surface brightness is the flux density per unit solid angle observed in an image. The unit of measure of the surface brightness is $\frac{L_\odot}{pc^2}$.

[10]In particular, the two surface brightness are identical per unit frequency since gravitational lensing is *achromatic* (see the comment below Eq. (1.1.6) in Sect. 1.1).

[11]A solid angle subtended by a spherical surface is defined as the ratio of that area to the squared radius of the sphere.

[12]The inverse of A describes the distortion due to the inverse of the lens map.

The Jacobian A of the lens mapping is explicitly

$$A_{ij} = \phi_{,ij} = \delta_{ij} - \psi_{,ij} \tag{1.4.4}$$

so that

$$A(\boldsymbol{x}) = \begin{pmatrix} 1 - \kappa(\boldsymbol{x}) - \gamma_1(\boldsymbol{x}) & -\gamma_2(\boldsymbol{x}) \\ -\gamma_2(\boldsymbol{x}) & 1 - \kappa(\boldsymbol{x}) + \gamma_1(\boldsymbol{x}) \end{pmatrix}, \tag{1.4.5}$$

where

$$\gamma_1 := \frac{1}{2}(\psi_{,11} - \psi_{,22}) \tag{1.4.6}$$

and

$$\gamma_2 := \psi_{,12} = \psi_{,21}. \tag{1.4.7}$$

The determinant and the trace can be rapidly computed:

$$det\, A = (1 - \kappa)^2 - \gamma^2, \tag{1.4.8}$$

$$tr\, A = 2(1 - \kappa). \tag{1.4.9}$$

γ is called the *shear* and it is defined as

$$\gamma := \sqrt{\gamma_1^2 + \gamma_2^2}. \tag{1.4.10}$$

A is a real symmetric matrix, so it can be diagonalized ($D = P^T A P$, where P is an orthogonal matrix and D a diagonal matrix). Its eigenvalues are

$$\lambda_{1,2} = 1 - \kappa \mp \gamma. \tag{1.4.11}$$

Accordingly, A is similar to the following diagonal matrix

$$A \sim D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \tag{1.4.12}$$

In order to better understand how the image of a source is distorted by a gravitational lens, it is helpful to write the matrix A in a new form:

$$A(\boldsymbol{x}) = (1 - \kappa)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \gamma \begin{pmatrix} cos(2\varphi) & sin(2\varphi) \\ sin(2\varphi) & -cos(2\varphi) \end{pmatrix}, \tag{1.4.13}$$

where φ is defined by: $\gamma_1 = \gamma cos(2\varphi)$ and $\gamma_2 = \gamma sin(2\varphi)$. From this expression and the locally linearized lens function $d\boldsymbol{y} = A(\boldsymbol{x})d\boldsymbol{x}$ (see Eq. (1.4.2)), it can be seen that the distortion of an image is caused by the concurrent effects of the convergence κ and of the shear γ. The convergence determines an istotropic scaling[13], whereas, the action of the shear is to stretch or compress the image along the line making an angle φ with the abscissa. For instance, the image of a small circular source is an ellipse stretched in the direction identified by φ. In Eq. (1.4.13), the argument of sin and cos is 2φ because an elongation along a certain line is equivalent to an elongation along that line rotated by 180°.

The eigenvectors of A can be calculated from Eq. (1.4.13).

$$\boldsymbol{v}_{1,2} = \begin{pmatrix} sin(\varphi) \\ -cos(\varphi) \end{pmatrix}, \begin{pmatrix} cos(\varphi) \\ sin(\varphi) \end{pmatrix}, \tag{1.4.14}$$

where $\boldsymbol{v}_1$ is the eigenvector related to the eigenvalue λ_1 and $\boldsymbol{v}_2$ is tied to λ_2 (see Eq. (1.4.11)). In the basis of eigenvectors ($\boldsymbol{v}_1$ and $\boldsymbol{v}_2$), the locally linearized lens map ($A(\boldsymbol{x})d\boldsymbol{x} = d\boldsymbol{y}$ in Cartesian coordinates) becomes $D(\boldsymbol{x})d\boldsymbol{x} = d\boldsymbol{y}$.

For a fixed source position $\tilde{\boldsymbol{y}}$, $\phi(\boldsymbol{x}, \tilde{\boldsymbol{y}})$ defines a surface which is strictly related to the arrival time of an image (see Eq. (1.3.11)). The position $\boldsymbol{x}$ of an image, at which the lens equation ($\nabla_{\boldsymbol{x}}\phi(\boldsymbol{x}, \boldsymbol{y}) = 0$, Eq. (1.3.7)) holds, is a stationary point of ϕ. An image is said to be *non-critical* if the Jacobian A is non-degenerate. There are three different types of *non-critical images*. They are classified by what type of stationary point they represent. Before listing these types of images, it is important to underline that A is the *Hessian* of the function $\phi(\boldsymbol{x}, \boldsymbol{y})$ (with fixed $\boldsymbol{y}$).

Type I ϕ is minimum at $\boldsymbol{x}$, which means that $det A(\boldsymbol{x}) > 0$ and $1 - \kappa(\boldsymbol{x}) - \gamma_1(\boldsymbol{x}) > 0$.
This implies

$$\kappa(\boldsymbol{x}) < 1. \tag{1.4.15}$$

Type II ϕ is a saddle at $\boldsymbol{x}$, i.e. $det A(\boldsymbol{x}) < 0$.

---

[13]It is isotropic because the first matrix in the sum of Eq. (1.4.13) is the identity.

Type III ϕ is maximum at $\boldsymbol{x}$, i.e. $detA(\boldsymbol{x}) > 0$ and $1 - \kappa(\boldsymbol{x}) - \gamma_1(\boldsymbol{x}) < 0$. As a consequence

$$\kappa(\boldsymbol{x}) > 1. \tag{1.4.16}$$

With reference to the magnification of an image, $\mu(\boldsymbol{x})$ is positive if $detA(\boldsymbol{x}) > 0$ and it is negative if $detA(\boldsymbol{x}) < 0$. In the former case, the image is said to have a *positive parity* and, in the latter, a *negative parity*. The physical meaning of these definitions is that an image with negative parity is mirrored, while an image with positive parity is not. It is possible to classify the parity of images of the same source by recognizing an image of type I (a minimum point). In that case, an image will have negative parity if its shape is flipped with respect to the image of type I, otherwise it will have a positive parity. The global minimum point (see Sect. 1.5) can be found if the luminosity of the source is variable. In this situation, since this image arrives at the observer first (see relation between ϕ and time delay: Eq. (1.3.11)), it can be identified by measuring the variation over time of luminosity among the images.

An other important aspect about magnification is the de-magnification of images that occur in the innermost regions of a *lens* mass distribution. One can easily prove that a large value of $\kappa(\boldsymbol{x})$ causes a small magnification at $\boldsymbol{x}$. In fact, exploiting Eq. (1.4.8) and Eq. (1.4.3), it is apparent that

$$\mu(\boldsymbol{x}) = \frac{1}{(1 - \kappa(\boldsymbol{x}))^2 - \gamma(\boldsymbol{x})^2}. \tag{1.4.17}$$

In mass distributions where the central density rockets, the images occurring in the central region are too feeble to be observed.

1.5   Number of images, critical curves and caustics

In the lens plane, there are curves on which at least one eigenvalue of A vanishes. These curves, called *critical curves*, divide the plane in regions where eigenvalues have

constant sign. The eigenvalues of A (see Eq. (1.4.12)) depend continuously[14] on $\boldsymbol{x}$, therefore, they can change sign only if they cross a critical curve. Accordingly, in every region, only one type of image can be found (see Sect. 1.4). The ray-traced images[15] of the critical curves in the source plane are called *caustics*.

Having defined what the critical curves and the caustics are, we shall discuss their employment in gravitational lensing theory. If $\kappa(\boldsymbol{x})$ is smooth and decreases stronger than $|\boldsymbol{x}|^{-2}$ at large $|\boldsymbol{x}|$, it can be demonstrated that the Fermat potential, with a fixed $\boldsymbol{y}$, $\phi(\boldsymbol{x}, \boldsymbol{y})$ has at least a minimum point which satisfies the lens equation (see Eq. (1.3.7)). Thus, the lens map $f : \boldsymbol{x} \longmapsto \boldsymbol{y}$ is surjective because for any source position $\boldsymbol{y}$, there exists at least an image at $\boldsymbol{x}$. In particular, when there is more than one solution to the lens equation, multiple images occur.

Moreover, the lens map is *locally invertible*[16] around $\boldsymbol{x}_0$ if $det A(\boldsymbol{x}_0) \neq 0$. The map can be invertible at some images and not at others where $det A(\boldsymbol{x}_0) = 0$. Instead, the lens function is invertible at every image of a source that does not lie on a caustic. For this reason, varying the source position $\boldsymbol{y}$, the number of images can change only if $\boldsymbol{y}$ crosses a caustics. Otherwise, the lens map remains locally invertible at every images of $\boldsymbol{y}$ and, consequently, the images cannot be created and they cannot disappear. It can be shown that crossing a caustic leads to a variation in the number of images of ± 2.

Now, an important theorem about the number of images will be stated (Schneider, 1992). The hypotheses are: $\kappa(\boldsymbol{x})$ is smooth and decreases faster than $|\boldsymbol{x}|^{-2}$ for $|\boldsymbol{x}| \rightarrow +\infty$ and $\boldsymbol{y}$ does not lie on a caustic.

Theorem 1: Under those hypotheses, the total number of images n is equal to $2n_{II} + 1$, where n_{II} is the number of images of the second type (saddle points).

---

[14]The lens function is assumed to be differentiable.

[15]Here, images stand for the images of the lens function, not to be confused with the images of a source. The former belong to the source plane, the latter to the lens plane.

[16]Taking f differentiable, its local invertibility can be proved by means of the *implicit function theorem*.

A direct corollary is that the total number of images is odd. However, there are many real cases where this does not happen. Two are the main reasons. Firstly, the image might form in a very dense region of the lens plane (usually the center of the lens) and so be highly de-magnified (see the end of Sect. 1.4). Secondly, the lens could be non-*transparent*. This means that the lens blocks the light rays of the source with an impact parameter $|\boldsymbol{\xi}|$ smaller than the lens radius. A non-transparent lens can prevent the occurrence of the central image.

The focus of this chapter is the *strong* gravitational lensing. Accordingly, it is useful to mention the following two conditions (Schneider, 1992) for multiple imaging.

1. A transparent lens produces multiple images if, and only if, there is a point $\boldsymbol{x}_0$ such that $det A(\boldsymbol{x}_0) < 0$. Proof: If there exists $\boldsymbol{x}_0$ such that $det A(\boldsymbol{x}_0) < 0$, the image at $\boldsymbol{x}_0$ is a saddle point. Therefore, there must be at least three images because of Theorem 1. The necessary condition is proved noticing that if $det A(\boldsymbol{x}) > 0$ everywhere, the lens map is globally invertible and only one image is allowed.

2. If there is a point $\boldsymbol{x}_0$ such that $\kappa(\boldsymbol{x}_0) > 1$, multiple images occur. Proof: At $\boldsymbol{x}_0$ there is an image of the third type because $\kappa(\boldsymbol{x}_0) > 1$ (see Eq. (1.4.16)). Accordingly, the statement is demonstrated because there is always at least another image of the first type (minimum point).

With regard to caustics, they can be classified in *folds* and *cusps*[17]. When $\boldsymbol{y}$ crosses a caustic, going form a region where n images occur to a region of $n-2$ images, the images behave differently whether the caustic is a fold or a cusp. As $\boldsymbol{y}$ gets closer to a *fold*, two images approach the critical curve. Then, the images merge when $\boldsymbol{y}$ is on the caustic and they disapper as the caustic is crossed. On the other hand, as $\boldsymbol{y}$ moves closer to a *cusp*, three images approach the critical curve. Then,

---

[17]Other types exist. However, according to Whitney's theorem (Schneider, 1992), a lens map with other types of caustics can be slightly changed in order to produce only folds and cusps.

the images fuse together when $\boldsymbol{y}$ is on the caustic. After crossing the cusp, only one image remains. Cusps have also the following property: the vector tangent to the critical curve (related to the cusp) is an eigenvector of A with vanishing eigenvalue.

To conclude this Section, we emphasize that when an image is located near a critical curve, is highly distorted and magnified ($\frac{1}{detA(\boldsymbol{x})} = \mu(\boldsymbol{x})$). On critical curves, the magnification diverges to infinite. In reality, this does not happen because the geometrical-optics approximation is not valid near caustics (Ohanian, 1983).

1.6 Simple lens models

This Section will illustrate a simple category of lens models that can be treated analytically: the axially symmetric models. Their foremost advantage is that the ray-trace equation reduces to a unidimensional one. In these models, the surface mass density of the lens must be circularly-symmetric around the center of the lens plane, i.e. $\Sigma(\boldsymbol{\xi}) = \Sigma(|\boldsymbol{\xi}|)$. In order to show the reduction of the lens equation to a one dimensional form, we write the scaled deflection angle (see Eq. (1.3.3)) restricting $\boldsymbol{x}$ to the positive x_1-axis of the lens plane ($x \geq 0$). This restriction is allowed because of the circular symmetry. For the same reason, it is convenient to write $\boldsymbol{x}$ in polar coordinates: $\boldsymbol{x} = x(cos\varphi, sin\varphi)$ and $\kappa(\boldsymbol{x}) = \kappa(x)$. Using these elements, the coordinates of the scaled deflection angle are

$$\alpha_1(x) = \frac{1}{\pi} \int_0^{+\infty} x' dx' \kappa(x') \int_0^{2\pi} d\varphi \frac{x - x' cos\varphi}{x^2 + x'^2 - 2xx' cos\varphi}, \qquad (1.6.1)$$

$$\alpha_2(x) = \frac{1}{\pi} \int_0^{+\infty} x' dx' \kappa(x') \int_0^{2\pi} d\varphi \frac{-x' sin\varphi}{x^2 + x'^2 - 2xx' cos\varphi}. \qquad (1.6.2)$$

$\alpha_2(x)$ vanishes because the integral in φ of Eq. (1.6.2) is equal to zero[18]. Accordingly, if $\boldsymbol{x}$ is taken on the x_1-axis, $\boldsymbol{\alpha}$ points to the same direction. In addition, it can be seen from the lens equation (see Eq. (1.2.5)) that also $\boldsymbol{y}$ is parallel to $\boldsymbol{x}$ and $\boldsymbol{\alpha}$ ($\boldsymbol{y} /\!/ \boldsymbol{x} /\!/ \boldsymbol{\alpha}$). Therefore, the lens equation can be reduced to a unidimensional form. Turning to

---

[18]It is an integral over $(0, 2\pi)$ of an odd function with respect to π (i.e. $g(\pi + \varphi) = -g(\pi - \varphi)$).

the first component of α, the integral in φ vanishes for $x' > x$ and equals $2\pi/x$ for $x' < x$. Accordingly, $\alpha_1(x)$ becomes

$$\alpha(x) := \alpha_1(x) = \frac{2}{x} \int_0^x x' dx' \kappa(x') . \tag{1.6.3}$$

From this point forward, $\alpha_1(x)$ will be written without the subscript since the lens equation is one-dimensional. The Eq. (1.6.3) shows that the scaled deflection angle $\alpha(x)$ is determined by the mass inside the circle of radius x, centered at the center of the lens, as if all the mass were concentrated in a single point at the center. In contrast, the mass outside the circle of radius x does not influence the deflection angle. If we consider the (not scaled) deflection angle $\hat{\alpha}$, it can be noticed that the $\hat{\alpha}$ of an axially-symmetric lens has the same form of the $\hat{\alpha}$ of a point-mass (see Eq. (1.1.14)). The only difference is that the mass M in the relation of the axially-symmetric lens corresponds to the mass inside the circle of radius ξ (impact factor, i.e. the distance between the center of the lens and the light ray in the lens plane).

$$\hat{\alpha}(x) = \frac{1}{\xi}\frac{4G}{c^2} \int_0^\xi 2\pi\xi' d\xi' \Sigma(\xi') = \frac{4GM(\xi)}{c^2\xi}. \tag{1.6.4}$$

Now, we can define the dimensionless mass inside a circle of radius x, $m(x)$, as

$$\frac{m(x)}{x} := \alpha(x) = \frac{2}{x} \int_0^x x' dx' \kappa(x') . \tag{1.6.5}$$

While, $\overline{\kappa}(x)$ is the *dimensionless mean surface mass density* inside a circle with radius x:

$$\overline{\kappa}(x) := \frac{m(x)}{x^2}. \tag{1.6.6}$$

The lens equation with $x \in \mathbb{R}$ (not longer restricted to the positive axis) is

$$y = x - \alpha(x) = x - \frac{m(x)}{x}, \tag{1.6.7}$$

with $\alpha(-x) = -\alpha(x)$ and $m(|x|) = m(x)$. From the ray-trace equation, the matrix A can be computed analytically. Accordingly, $det A$ and the eigenvalues can be obtained as well.

$$det A(x) = \left(1 - \frac{m(x)}{x^2}\right)\left(1 - \frac{d}{dx}\left(\frac{m(x)}{x}\right)\right) = \tag{1.6.8}$$

$$= \left(1 - \frac{m(x)}{x^2}\right)\left(1 + \frac{m(x)}{x^2} - 2\kappa(x)\right), \tag{1.6.9}$$

where, in the second equivalence, it has been used the following relation:

$$\frac{dm(x)}{dx} = 2x\kappa(x), \tag{1.6.10}$$

which is a direct consequence of (1.6.5). The two factors in Eqs. (1.6.8) and (1.6.9) are the eigenvalues of A:

$$\lambda_t(x) = 1 - \frac{m(x)}{x^2} = 1 - \overline{\kappa}(x), \tag{1.6.11}$$

$$\lambda_r(x) = 1 + \frac{m(x)}{x^2} - 2\kappa(x), \tag{1.6.12}$$

where the reason of subscripts r and t will be clear by the end of this section. In Eq. (1.6.11), the definition of $\overline{\kappa}(x)$ (see Eq. (1.6.6)) has been exploited. Moreover, it can be proved through the analytical computation of the convergence κ and the shear γ that λ_t is equal to $1 - \kappa - \gamma$ and λ_r to $1 - \kappa + \gamma$.

Requiring that $\lambda_t(x) = 0$ and $\lambda_r(x) = 0$, we obtain the *tangential* and the *radial* critical curves, respectively. These curves have two important properties. Firstly, they are circles with radii determined by the following equations: $\overline{\kappa}(x) = 1$ for tangential critical curves and $\frac{d}{dx}\left(\frac{m(x)}{x}\right) = 1$ for radial critical curves. Secondly, the vector *tangent* to the tangential critical curve, at a point $\tilde{x}$ on the curve, is an eigenvector of $A(\tilde{x})$ with zero eigenvalue. In the case of radial critical curves, the vector *normal* to the critical curve (circle) is an eigenvector with vanishing eigenvalue. From these arguments, it can be shown that an image at $\hat{x}$ close to a tangential critical curve is stretched in the tangential direction to the curve by a factor $\lambda_t^{-1}(\hat{x})$, while, an image near a radial critical curve is elongated in the radial direction by a factor $\lambda_r^{-1}(\hat{x})$.

It is useful to visualize the lens equation for axially-symmetric deflecting masses in the following form:

$$y = x[1 - \overline{\kappa}(|x|)], \tag{1.6.13}$$

where x and y belong to real numbers. This expression is particularly interesting because it shows that the tangential critical curve ($\lambda_t(x) = 0$, i.e. $\overline{\kappa}(|x|) = 1$) is mapped to the origin of the source plane. The equation $\overline{\kappa}(|x|) = 1$ suggests that the mean surface mass density $\langle \Sigma(xD_L) \rangle$ within a circle of radius $R = xD_L$ (with the center at the origin of the lens plane) is equal to Σ_{cr}. Accordingly, this statement leads to the following famous relation:

$$M(R \leq R_E) = \Sigma_{cr} \pi R_E^2, \tag{1.6.14}$$

where $M(R \leq R_E)$ is the total mass (projected on the lens plane) inside the circle of radius R_E. R_E is known as *Einstein radius* and corresponds to the radius of the tangential critical curve (Fig. 1.2). Thanks to this simple expression, only the measurements of the Einstein radius and the cosmological distances (D_L, D_S and D_{LS}) are needed to calculate the projected mass of the lens within the Einstein radius. In addition, Eq. (1.6.14) is valid for all circular lens models, regardless of the density profile.

For the purposes of this thesis, it is convenient to study in further detail a special type of axially-symmetric lens model, that is the SIS (Singular Isothermal Sphere). The SIS model describes an isotropic distribution of self-gravitating matter whose particles follow the Maxwell-Boltzmann velocity distribution. Also, the matter distribution is in hydrostatic and thermal equilibrium. The combination of the Maxwell-Boltzmann distribution and the thermal equilibrium implies that the velocity dispersion of the particles σ is constant with respect to variations of the radius. The three dimensional mass density of a SIS model is

$$\rho_{\text{SIS}}(r) = \frac{\sigma_{los}^2}{2\pi G r^2}, \tag{1.6.15}$$

where r is the distance between the center of the lens and the point where the density is evaluated. σ_{los} is the line-of-sight velocity dispersion of the particles in the matter distribution (e.g. velocity dispersion of galaxies in a galaxy cluster). Knowing the
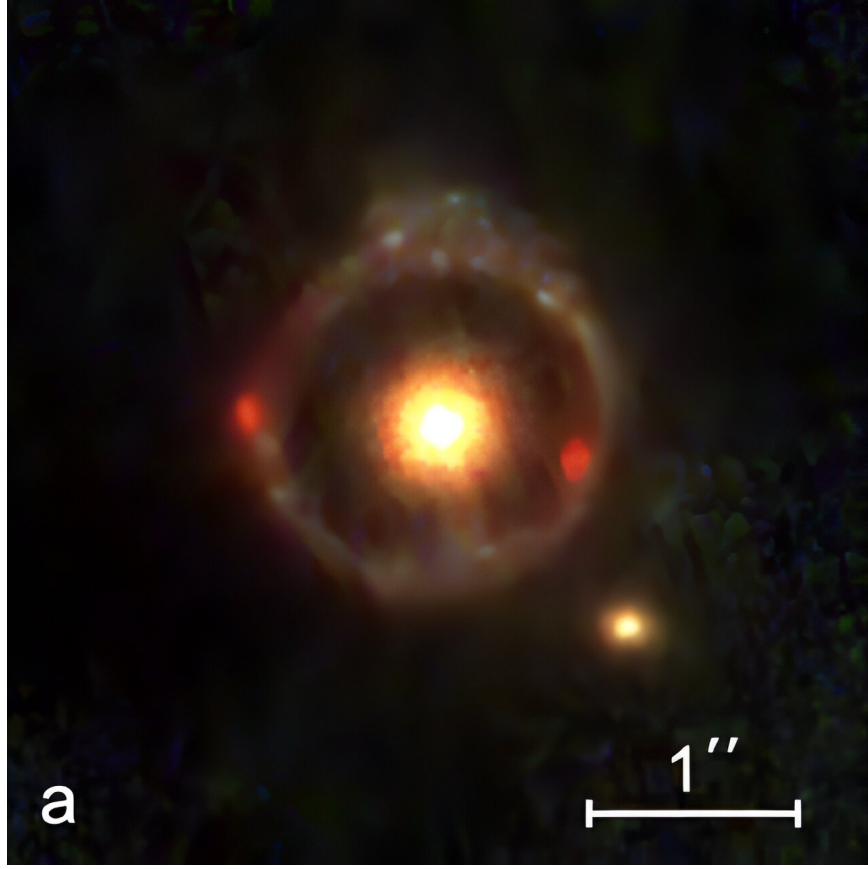
FIGURE 1.2. A JWST image of a complete Einstein ring. It is a color image produced from the NIRCam F115W, F150W, and F277W data. The ring, which is the image of a galaxy in the background ($z = 2.98^{+42}_{-47}$), coincides with the tangential critical curve of the lens. Its radius is the Einstein radius. (van Dokkum, P., Brammer, G., Wang, B. et al. A massive compact quiescent galaxy at z = 2 with a complete Einstein ring in JWST imaging. Nat Astron (2024).)

spherical mass density $\rho(r)$, it is easy to project it on the lens plane and obtain the surface mass density

$$\Sigma_{\text{SIS}}(\xi) = 2 \int_{\xi}^{+\infty} dr \, \rho_{\text{SIS}}(r) \frac{r}{\sqrt{r^2 - \xi^2}} = \frac{\sigma_{los}^2}{2G} \frac{1}{\xi}. \qquad (1.6.16)$$

Then, the total mass within a circle (in the lens plane) of radius ξ is

$$M_{\text{SIS}}(\xi) = \frac{\pi \sigma_{los}^2}{G} \xi, \qquad (1.6.17)$$

and the scaled deflection angle for $x \geq 0$ is equal to

$$\alpha_{\text{SIS}}(x) = \frac{1}{x} \int_0^x dx' 2\kappa_{\text{SIS}}(x')x' = \frac{1}{x} \int_0^x dx' 2\frac{\Sigma_{\text{SIS}}(x'D_L)}{\Sigma_{cr}} x' = \left(\frac{\sigma_{los}}{c}\right)^2 \frac{4\pi D_{LS}}{D_S}.$$
(1.6.18)

Knowing that $\alpha(-x) = -\alpha(x)$ for $x \in \mathbb{R}$, it can be found the scaled deflection angle is: $\alpha_{\text{SIS}}(x) = \frac{x}{|x|} \left(\frac{\sigma_{los}}{c}\right)^2 \frac{4\pi D_{LS}}{D_S}$, $x \in \mathbb{R}$. From the scaled deflection angle, the lens equation is

$$y = x - \theta_E \frac{x}{|x|},$$
(1.6.19)

where, it has been noticed that $|\alpha_{\text{SIS}}(x)|$ is equal to the *Einstein angle* (the Einstein angle is tied to the Einstein radius by $R_E = \theta_E D_L$). This can be proved by verifying that $\lambda_t(\alpha_{\text{SIS}}(x)) = 0$. The shear and the convergence for an image at x are

$$\kappa(x) = \gamma(x) = \frac{\theta_E}{2|x|}.$$
(1.6.20)

If y is taken positive[19], there is only one image for $y > \theta_E$ at $x = y + \theta_E$, while, two images occur for $y < \theta_E$ at $x_1 = y + \theta_E$ and at $x_2 = y - \theta_E$. x_1 is an image of type I (a minimum point) with positive magnification $\mu(y) = 1 + \frac{\theta_E}{y}$ and x_2 is an image of second type (a saddle point) with negative magnification $\mu(y) = 1 - \frac{\theta_E}{y}$. The distance between the two images $|x_1 - x_2|$ is always equal to $2\theta_E$. For $y < \theta_E$, the Theorem 1 (see Sect. 1.5) appears to fail because there are an even number of images (two). This happens because the Singular Isothermal Sphere has a non-physical mass distribution. Firstly, it has an infinite total mass. Nevertheless, this does not affect the imaging considerably and a simple solution to this problem is the introduction of a cut radius that is much larger than the Einstein radius. Secondly, the SIS model has infinite density at the center of the lens. This is what prevent the central image (the third image) to occur. In order to recover the third image, more realistic models have been developed. These have a small core radius so that no divergence in the mass density is present. The effect of the introduction of a core radius on the scaled deflection

---

[19]The case of $y < 0$ is equivalent because of the symmetry. The choice of considering a positive y is made for reasons of clarity.

angle is displayed in Fig. 1.3. In particular, the figure shows the comparison between the discontinuous $\alpha_{\mathrm{SIS}}(x) = \theta_E \frac{x}{|x|}$ and the smooth $\alpha(x)$ of a model with a core radius r_{core} (in particular a NIS, i.e. Non-singular Isothermal Sphere, with a mass density of $\rho_{\mathrm{NIS}}(r) = \frac{\sigma_{los}^2}{2\pi G} \frac{1}{r^2 + r_{core}^2})$. Given a certain y, if the function $f(x) = x - y$ is drawn in the same graph of $\alpha(x)$ ($\alpha(x)$ on the y-axis and x on the x-axis), the points of intersection between $\alpha(x)$ and $f(x)$ identify the positions of the multiple images x of the source at y. The reason for this is that the lens equation is satisfied at these points of intersection, where $x - y = \alpha(x)$. Applying this method, it is straightforward to count the number of observed multiple images, given the scaled deflection angle $\alpha(x)$ and the source position y. We note that, in the NIS model, the third central image (not predicted by SIS) is restored. To conclude the discussion of the core radius, we emphasize that if the core radius is very small, the deflection angle is almost identical to the one of the SIS model.

(a) SIS

(b) NIS-src
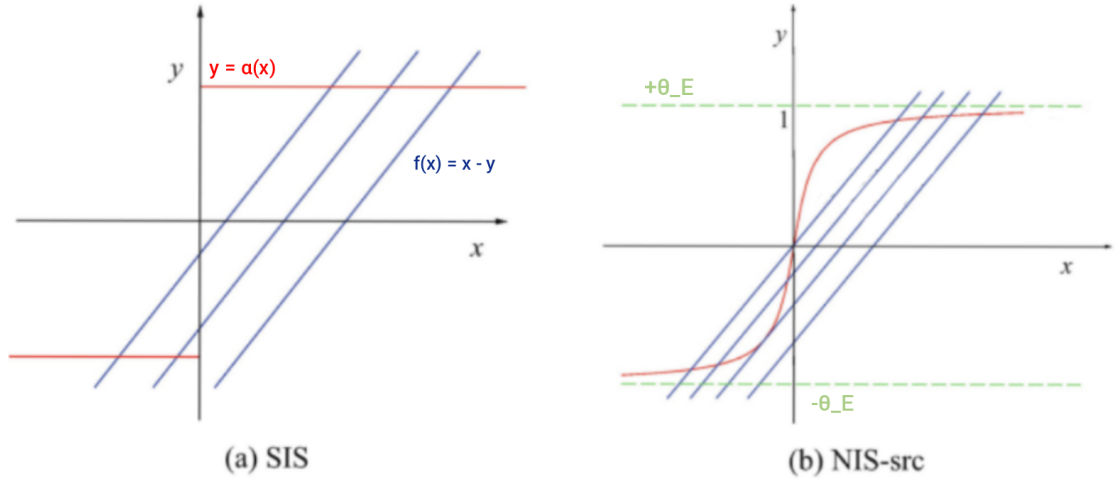
FIGURE 1.3. Comparison between the deflection angle of a SIS model and of a NIS model (with a core radius). Left: it illustrates how $\alpha_{\mathrm{SIS}}(x) = \theta_E \frac{x}{|x|}$ (red line) changes in function of the position angle x (on the x-axis). The blue lines represent the functions $f(x) = x - y$ for different values of y. Right: the deflection angle $\alpha_{\mathrm{NIS}}(x)$ is in red and $f(x)$ in blue. (Figure from Principles of Gravitational Lensing by Arthur B. Congdon and Charles R. Keeton, 2018).

Another improvement of the SIS model is obtained by taking into account the

ellipticity of the lens. A popular lens model that has a core radius, a cut radius, and takes into consideration the ellipticity is the dual Pseudo Isothermal Elliptical model (dPIE). This specific model will be widely discussed and used in the next chapters.

To conclude our discussion on the SIS model, we calculate the time delay between two images of the same source for a SIS model:

$$\Delta t_{delay,\text{SIS}}(y) = \frac{1+z_l}{c} \frac{D_L D_S}{D_{LS}} 2y\theta_E. \tag{1.6.21}$$

If $y \longrightarrow 0$, $\Delta t \longrightarrow 0$. While, if $y \longrightarrow \theta_E$, $\Delta t \longrightarrow \frac{1+z_l}{c} \frac{D_L D_S}{D_{LS}} 2\theta_E^2$. Assuming that the lens is a galaxy, the typical values of time delay vary from a few days for y close to zero to hundreds of days for y near θ_E. For $y > \theta_E$, we remark that there is only one image (or type I). The measurement of time delay and of the image positions provides a method (Grillo et al. 2020, Suyu et al. 2012) to estimate the Hubble constant H_0 since the angular-diameter distances depend on H_0 ($D \propto H_0^{-1}$).

1.7 The Schwarzschild lens model

This section is focused on the Schwarzschild lens model. In this model, the lens consists of a point mass M. Its deflection angle was obtained in Sect. 1.1 (see Eq. (1.1.14)). Consequently, the lens equation of a point mass, located at the origin of the lens plane, is:

$$y = x - \frac{4GM}{c^2 x D_L} \frac{D_{LS}}{D_S}, \tag{1.7.1}$$

where x and y are the one-dimensional position angles, since the Schwarzschild lens is axially-symmetric (see Sect. 1.6). Given a certain source position y, two images can be found at $x_{1,2}$:

$$x_{1,2} = \frac{1}{2} \left(y \pm \sqrt{y^2 + 4\left(\frac{4GM}{c^2} \frac{D_{LS}}{D_S D_L}\right)} \right). \tag{1.7.2}$$

It can be noticed that the two images occur on each side of the lens. Moreover, the number of images is even because the central density of the point mass is a Dirac

delta function. Accordingly, the central image cannot occur since the mass density of the lens and the corresponding deflection angle are singular in the centre of the lens (see Sect. 1.6, above Fig. 1.3). As described in Sect. 1.6, in axially-symmetric lenses, if the source is located in the origin of the source plane ($y = 0$), the value of the image position x, obtained from Eq. (1.7.2), is equal to the Einstein angle $\theta_E = \sqrt{\frac{4GM}{c^2}\frac{D_{LS}}{D_L D_S}}$. We recall that θ_E is the angular radius of the tangential critical curve.

From Eqs. (1.6.5) and (1.6.8), we obtain:

$$det A(x) = \left(1 - \frac{\alpha(x)}{x}\right)(1 - \alpha'(x)) = \left(1 - \frac{\theta_E^2}{x^2}\right)\left(1 + \frac{\theta_E^2}{x^2}\right) = 1 - \left(\frac{\theta_E}{x}\right)^4, \quad (1.7.3)$$

thus, the magnification of an image at x is

$$\mu(x) = \frac{1}{det A(x)} = \left(1 - \left(\frac{\theta_E}{x}\right)^4\right)^{-1}. \quad (1.7.4)$$

Combining Eq. (1.7.4) and Eq. (1.7.2), we can obtain the magnification of each image as a function of the source position:

$$\mu_{1,2}(y) = \frac{1}{4}\left(\frac{y}{\sqrt{y^2 + 4\theta_E^2}} + \frac{\sqrt{y^2 + 4\theta_E^2}}{y} \pm 2\right), \quad (1.7.5)$$

and the ratio between the magnifications of the two images:

$$\frac{\mu_1(y)}{\mu_2(y)} = \left(\frac{\sqrt{y^2 + 4\theta_E^2} + y}{\sqrt{y^2 + 4\theta_E^2} - y}\right)^2. \quad (1.7.6)$$

From the last equation, it can be noted that the two images have comparable magnifications and, thus comparable fluxes, only if y is close to zero. Finally, we note that if the values of the ratio $\frac{\mu_1(y)}{\mu_2(y)}$ and the separation of the two image positions are measured, one can estimate the Einstein angle θ_E and the source position y.

1.8 Dual Pseudo Isothermal Elliptical mass distribution

The dPIE model takes into account three elements that the SIS model neglects. Firstly, the dPIE has a core radius that is used to remove the central singularity of

the mass density of the SIS model. Secondly, a cut radius ensures that the total mass is finite (the mass enclosed within an infinite radius converges). Moreover, the dPIE distribution is elliptical. Therefore, the dPIE is completely defined by 7 parameters: the coordinates (x, y) of its centre, the cut radius r_t, the core radius r_c, the one-dimensional line-of-sight velocity dispersion σ_0, the ellipticity ϵ, and the position angle θ, which determines the orientation of the dPIE. This distribution is convenient because it is possible to calculate analytically the gravitational potential, its first and second partial derivatives and, consequently, the deflection angle, convergence, shear, and time delays. The convergence κ_{dPIE} of a general dPIE is:

$$\kappa_{\mathrm{dPIE}}(R) = \frac{\sigma_0^2 r_t}{2G\Sigma_{cr}\left(r_t - r_c^2\right)}\left(\frac{1}{\sqrt{R^2 + r_c^2}} - \frac{1}{\sqrt{R^2 + r_t^2}}\right), \qquad (1.8.1)$$

where R is the distance from the center of the mass distribution projected on the lens plane given by $R = \frac{x^2}{(1+\epsilon)^2} + \frac{y^2}{(1-\epsilon)^2}$, where ϵ is the ellipticity ($\epsilon = \frac{1-q}{1+q}$, where q is the ratio between minor axis and major axis).

The three-dimensional mass density of the "spherical" dPIE (Limousin et al. 2005) is

$$\rho_{\mathrm{dPIE}}(r) = \frac{\rho_0}{(1 + r^2/r_c^2)(1 + r^2/r_t^2)}, r_t > r_c, \qquad (1.8.2)$$

where r is the distance from the center of the mass distribution and ρ_0 is the central density. This mass density has different behaviours depending on the value of r: for $0 < r < r_c$, $\rho \propto \rho_0$; for $r_c < r < r_t$, $\rho \propto r^{-2}$ and for $r > r_t$, $\rho \propto r^{-4}$. ρ_0 is related to the line-of-sight velocity dispersion σ_0 by:

$$\rho_0 = \frac{\sigma_0^2}{2\pi G}\frac{r_t + r_c}{r_c^2 r_t}. \qquad (1.8.3)$$

The total projected mass within a circle of radius R can be derived using Eq. (1.8.2):

$$
\begin{aligned}
M(R) &= 2\pi \int_0^R \Sigma(R') R' \, dR' \\
&= 2\pi \int_0^R \left(\int_{R'}^{+\infty} \frac{2\rho(r)r}{\sqrt{r^2 - R^2}} \, dr \right) R' \, dR' \\
&= \frac{\pi \sigma_0^2}{G} \frac{r_t}{r_t - r_c} \left(\sqrt{r_c^2 + R^2} - r_c - \sqrt{r_t^2 + R^2} + r_t \right)
\end{aligned}
\tag{1.8.4}
$$

Taking the limit of Eq. (1.8.4) for R approching infinity ($R \to +\infty$), the total mass of a spherical dPIE (the mass within an infinite radius) is

$$
M_{tot} = \frac{\pi \sigma_0^2 r_t}{G}.
\tag{1.8.5}
$$

Chapter 2

Galaxy Clusters

In the Universe, galaxies tend to gather together in groups and clusters. These two categories differ in the number N of cluster members and in size: groups have $N \lesssim 50$ in a region with a diameter D of $D \lesssim 1.5\text{Mpc}$, whereas clusters have $N \gtrsim 50$ and $D \gtrsim 1.5\text{Mpc}$. However, there is not a precise transition between groups and galaxy clusters.

Galaxy clusters are the largest gravitationally-bound systems in the Universe. Initially, astronomers used to identify them as overdensities of galaxies in the sky. Today, from the observations of the X-ray emission, the kinematics of galaxies and the gravitational lensing effect, it is known that the total mass of a galaxy cluster (typically around $10^{14} - 10^{15} M_\odot$) is composed of dark matter for $\sim 80 - 85\%$, IntraCluster Medium (hot gas) for $\sim 15\%$ and stars for $\sim 5\%$.

In 1958, George Abell compiled a catalog of 2712 clusters. He classified as galaxy clusters those satisfying the following two properties: 1) they had to contain at least 50 galaxies with a magnitude m ranging from m_3 to m_3+2, where m_3 is the magnitude of the third brightest cluster member; 2) these galaxies had to be located within the Abell angular radius $\theta_{Abell} = 1.7/z$ arcmin, where z is the estimated redshift of the cluster. In 1989, 1361 clusters were added to the catalog by Harold G. Corwin and Ronald P. Olowin. The Abell catalog is affected by projection effects. In fact, galaxy clusters are three-dimensional objects, but their observed images are projected onto the sky. Accordingly, an identified overdensity of galaxies in the sky can be caused by galaxies that do not belong to the same structure, but that lie close to the observer line of sight. In addition to this catalog, which is based on optical observations, surveys based on X-ray obervations (e.g. Röntgensatellit, ROSAT) have been carried

out. These are much less affected by line-of-sight projection effects.

In the literature, there are several classification schemes for galaxy clusters. The Abell classification divides clusters into *regular* and *irregular*. The former are nearly circularly symmetric with a high central concentration of galaxies, manly elliptical or lenticular (see Sect. 2.3). The latter are less symmetrical. They are characterized by the presence of substructures and their central density of galaxies is smaller. Another notable classification, by Rood and Sastry (1971), is based on the spatial distribution of the ten brightest galaxies of the cluster. The result is a scheme with six types of galaxy clusters (see Fig. 2.1):

1. *cD* (*supergiant*): galaxy clusters dominated by a single cD galaxy (see Sect. 2.3). Abell S1063 (the cluster studied in this thesis) is a cluster of type *cD*;

2. *B* (*binary*): they contain two bright galaxies in the centre. Abell 1656, also known as the Coma cluster, belongs to the type *B*;

3. *L* (*line*): at least three of the brightest galaxies lie on a straight line. Abell 426 (Perseus cluster) belongs to the type *L*;

4. *C* (*core-halo*): there is a central core with at least four of the ten brightest galaxies. Abell 2065, a cluster in the constellation of Corona Borealis, belongs to the type *C*;

5. *F* (*flat*): the brightest galaxies form a flattened distribution. Abell 2152 is a flat (*F*) cluster;

6. *I* (*irregular*): the brightest galaxies form an irregular distribution. Abell 400, in the constellation of Cetus, belongs to the type *I*.

Finally, Augustus Oemler (1974) classified galaxy clusters according to the ratio between elliptical, lenticular ($S0$) and spiral galaxies. This classification defines three classes: *spiral-rich* clusters, in which spiral galaxies prevail, *spiral-poor* clusters, in
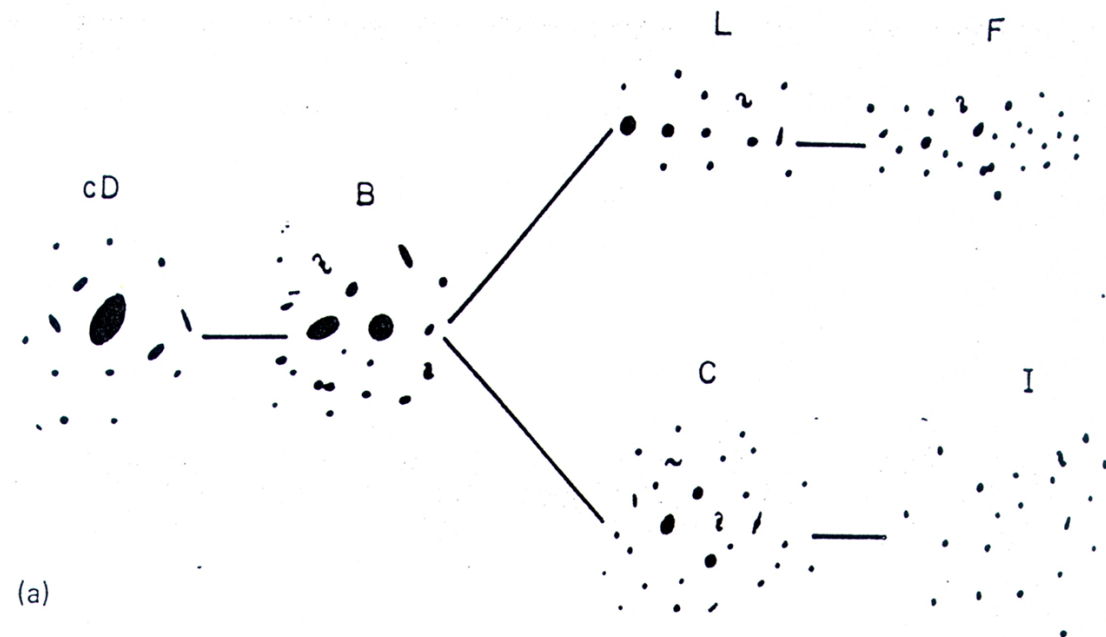
FIGURE 2.1. Diagram illustarting the classification of galaxy clusters by Rood and Sastry (1971). (sites.astro.caltech.edu).

which $S0$ are the most frequent galaxy type, and cD clusters which are characterized by a central cD galaxy and by the predominance of elliptical and lenticular galaxies.

The next sections will examine the primary mass components of galaxy clusters, namely dark matter (Sect. 2.1), the IntraCluster Medium (hot gas, in Sect. 2.2) and galaxies (Sect. 2.3, 2.4, 2.5). The last two sections (Sect. 2.6 and 2.7) will discuss another aspect: the presence of supermassive black holes at the centers of galaxies.

2.1 Dark matter

In 1933, Fritz Zwicky, studying the Coma cluster and, in particular the radial velocities of the cluster members, realized that the total mass of Coma had to be much larger than the total mass in stars of the cluster galaxies.

The typical radius R_{cl} of a galaxy cluster is $R_{cl} \approx 1\text{Mpc}$ and the typical velocity v of the galaxies in the cluster is $v \approx 1000\text{km/s}$. Accordingly, the estimated time t_{cross}

that a cluster member needs to cross the cluster is:

$$t_{cross} = \frac{R_{cl}}{v} \approx \frac{1\text{Mpc}}{1000\text{km/s}} \approx 10^9 \text{years}. \qquad (2.1.1)$$

It can be noticed that t_{cross} is significantly shorter than the typical age of a cluster, which can be approximated with the age of the Universe (≈ 14Gyr). Thus, cluster members have had enough time to possibly leave the cluster. The cluster has not dissolved because it is a gravitationally-bound system, in quasi-equilibrium. For this reason, one can use the Virial theorem[1] to estimate the total mass of the cluster. Given the typical velocity dispersion σ_{los} of the cluster galaxies ($\sigma_{los} \approx 1000$km/s), it can be demonstrated that the total mass of the cluster M_{cl} is:

$$M_{cl} = k_{vir}\frac{\sigma_{los}^2}{G}R_{cl}, \qquad (2.1.2)$$

where k_{vir} is a constant of order $\approx 3-5$ that, among other things, takes into account the internal structure of the cluster. It is crucial to understand that M_{cl} is the total mass (not only the stellar mass) because, in Eq. (2.1.2), M_{cl} is the mass contributing to the total gravitational potential of the cluster. Now, one can compare the estimate of the total mass of the cluster with that of the visible stellar mass of the cluster members. The latter estimate can be obtained from a measurement of the stellar mass-to-light ratio of galaxies (see Sect. 2.3). The result is that the mass of the stars in galaxies contributes for less than 5% to the total mass of the cluster. Even considering the hot-gas mass component, from X-ray observations (see Sect. 2.2), a significant fraction of the total mass is missing. This missing mass is known as dark matter, because it does not emit electromagnetic radiation. The nature of dark matter is still unknown. Recent works suggest that dark matter is mainly formed of some sort of non-baryonic matter.

Finally, cosmological N-body simulations[2] have found that the mass distribution of

---

[1] In the case of a self-gravitating system of galaxies in equilibrium, the Virial theorem yields: $2K + U = 0$, where K is the kinetic energy and U the potential energy of the system.

[2] Given a cosmological model, N-body simulations are employed to investigate non-linear structure evolution led by dark matter.

cosmological structures consists of large and extended halos and small-scale sub-halos, mainly corresponding to cluster galaxies. Moreover, according to these simulations, the mass profile of the dark-matter extended halos is characterized by a decrease at large radii steeper than that of an isothermal distribution and a flatter profile in the central region.

2.2   IntraCluster Medium

Galaxy clusters are very bright sources of X-ray emission. Their characteristic luminosity in the X-ray band ranges from 10^{43}erg/s to 10^{45}erg/s. This emission comes from an extended region (with a radius of $\gtrsim 1$Mpc) and does not arise from single galaxies. Also, it does not appear to vary significantly over time.

The observed X-ray radiation of galaxy clusters is primarly caused by thermal Bremsstrahlung from a hot gas (which consists mainly of electrons and ionized hydrogenum and helium). Bremsstrahlung is a continuum radiation emitted by the acceleration of electrons in the electric field of protons and atomic nuclei.

Now, assuming that the hot-gas particles move as fast as the galaxies in the cluster and assuming that the particles have a Maxwellian velocity distribution at temperature T and that the hot-gas distribution is isotropic, one can estimate the typical temperature T of the gas and its thermal energy ϵ:

$$T = \frac{m_p \sigma_{los}^2}{2k_B} \approx 6 \times 10^7 \text{K}, \tag{2.2.1}$$

where σ_{los} is the velocity dispersion of the gas in the cluster, m_p is the mass of the proton and k_B is the Boltzmann constant. Galaxy clusters have hot-gas temperatures of $T = 10^7\text{K} - 10^8\text{K}$ and corresponding thermal energies of $\epsilon = k_B T = 1\text{keV} - 10\text{keV}$ (X-ray energy range). In addition, it can be proven that the total power per unit volume emitted by thermal bremsstrahlung is proportional to the square root of the gas temperature T and to the square of the number density n_e of electrons in the gas:

$\propto T^{0.5} n_e^2$. The quadratic dependence on n_e is due to the fact that bremsstrahlung is a two-body process.

Looking at the morphology of the X-ray emission, two types of clusters can be classified: regular and irregular. Regular clusters have a surface brightness distribution approximately centered in their optical center and decreasing smoothly in their outer regions. Moreover, regular clusters have a high luminosity in the X-rays and high temperatures. In the presence of a central dominant (cD, see Sect. 2.3) galaxy, the cluster shows a strong central X-ray emission. In contrast, irregular clusters show several peaks in their X-ray emission, often corresponding to some cluster members.

Regarding the dynamics of the hot gas, the mean free path l of the gas particles ($l \sim 10\text{kpc}$) is much shorter than the typical dimension of a cluster ($R_{cl} \sim 1\text{Mpc}$). Therefore, the hot gas can be considered collisional. Furthermore, since the crossing time of the gas is much shorter than the age of the cluster (which is comparable to the age of the Universe), the hot gas is usually considered to be in hydrostatic equilibrium.

Finally, thanks to its X-ray emission, it can be estimated that the hot gas contributes for the $\sim 15\%$ to the total mass of a galaxy cluster.

2.3   Galaxies

Approximately $10 - 20\%$ of all galaxies live in clusters and at least 50% of them in clusters or groups. Galaxies can be classified according to their morphology. One of the most famous classification (based on optical photometry) is due to Hubble (1926): the Hubble tuning fork diagram (Fig. 2.2). This diagram divides the galaxies in three main types: elliptical, lenticular and spiral galaxies. Looking form the left to the right of the diagram, one can see the elliptical (E) galaxies, then the lenticular ($S0$) galaxies and finally the spiral (S) galaxies. The galaxies on the left (E and $S0$)

are called early-type[3] galaxies, while, the spiral galaxies, on the right, are usually referred to as late-type galaxies.

Elliptical galaxies (E) These galaxies are characterized by elliptical isophotes[4] and by the lack of a clear internal structure. They are further subdivided according to the value of their observed ellipticity ϵ. An elliptical galaxy classified as $E0$ is almost round, while, an $E7$ galaxy is very elliptical. Precisely, the number after the letter E is equal to ten times the value of the ellipiticity ϵ of the galaxy ($\epsilon = 1-b/a$, where a and b are the semimajor and semiminor axes, respectively). Observed elliptical galaxies have ellipticity values between 0 and around 0.7.

Spiral galaxies (S-SB) The structure of these galaxies consists of a central bulge and a disk with spiral arms. They are subdivided according to the shape of their arms. There are *normal spiral* (S) and *barred spiral* (SB) galaxies. In addition, a letter from a to c is assigned to each galaxy. Spiral galaxies, denoted with a, have a high ratio between the luminosity of the bulge L_{Bulge} and the luminosity of the disk L_{Disk} and have tightly wound spiral arms, whereas, spiral galaxies classified as c have a smaller ratio L_{Bulge}/L_{Disk} and more loosely wound arms. Spiral galaxies of type b have intermediate features.

Lenticular galaxies ($S0$) This type of galaxies represents the transition between elliptical and spiral galaxies. These lentil-shaped galaxies have a bulge and a disk similar to spiral galaxies, but do not exhibit any signs of spiral arms. Lenticular galaxies that show a bar are called $SB0$. Otherwise, they are denoted as $S0$.

Irregular galaxies (Irr) These are galaxies with a slightly regular ($IrrI$) or a completely irregular ($IrrII$) structure.

---

[3]These names, although frequently used, are only hystorical. They do not describe any properties of the galaxies.

[4]Isophotes are the contours of constant surface brightness.
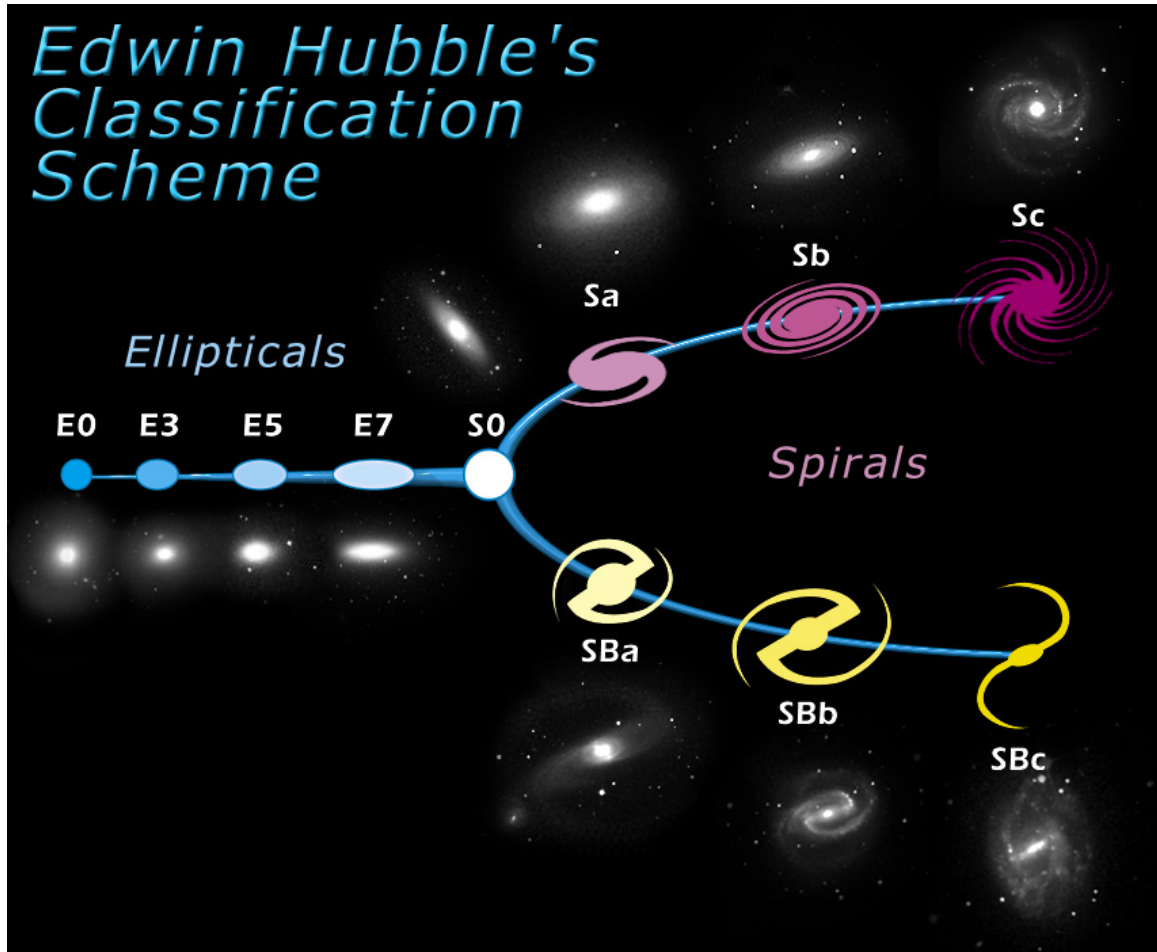
FIGURE 2.2. The Hubble classification of galaxies: Hubble tuning fork diagram. (esahubble.org).

The Hubble tuning fork diagram is affected by possible projection effects. In fact, the ellipticity of a galaxy depends on its inclination. Although this scheme is not an evolutionary sequence, as Hubble thought, galaxies that belong to the same class share several common properties and scaling relations. Besides the Hubble diagram, there are several other classification schemes for galaxies, based not only on optical photometry, but also on their spectral emission.

To conclude this section, some properties about the galaxy population in clusters are summarized. A correlation exists between the galaxy type and the number density

of galaxies of the surrounding environment. Elliptical and lenticular galaxies tend to be found predominantly in dense regions (Dressler 1980) of the Universe and late-type galaxies represent a larger portion of the galaxy population in regions of lower density. Thus, the majority of cluster galaxies are early-type galaxies. Finally, it has been observed that cluster galaxies at similar distances from the Earth have similar colors. This suggests that their evolution has been very similar.

2.4 Elliptical galaxies

This section is devoted to the study of elliptical galaxies since they are predominant in high-density environments, such as clusters.

The typical mass of an elliptical galaxy ranges from $10^8 M_\odot$ to $10^{14} M_\odot$. The diameter D_{25} at which the surface brightness is lower than 25mag in the B-band is $0.1 - 1000$kpc. The dominant color of elliptical galaxies is red, which implies an old stellar population, and their typical values of stellar mass-to-light ratio M/L_B[5] are approximately equal to $5-10 M_\odot/L_\odot$ (Sargent et al. 1977). The highest values of these three quantities are observed in the cD galaxies. These incredibly luminous elliptical galaxies can only be observed at the centers of galaxy clusters. Very often, regular galaxy clusters have at least one cD galaxy, which is also called BCG (Brightest Cluster Member). The BCG can account for up to about 30% of the total luminosity in the visible of a cluster.

The surface brightness profile of elliptical galaxies, can be well approximated by a De Vaucouleur profile:

$$I(R) = I_e e^{-7.669\left(1-\left(\frac{R}{R_e}\right)^{\frac{1}{4}}\right)}, \tag{2.4.1}$$

where $I(R)$ is the surface brightness at distance R, which is the projected distance from the center of the galaxy. I_e is the value of the surface brightness at the half-

---

[5]The ratio M/L_B is between the mass M of the stars in the galaxy and the luminosity L_B of the galaxy in the band B.

light radius: $I_e = I(R_e)$. R_e is the half-light radius, i.e. the radius of the circle that encloses half of the galaxy total luminosity. In other words, R_e satisfies the following equation:

$$\int_0^{R_e} I(R) R \, dR = \frac{1}{2} \int_0^{+\infty} I(R) R \, dR. \tag{2.4.2}$$

Regarding the stellar kinematics, elliptical galaxies possess three-dimensional ellipsoid shapes, where stars follow mostly random orbits around the center. Moreover, given an elliptical galaxy characterized by a radius r_e and a velocity dispersion σ_e, the estimated time t_{cross} that a star needs to cross the elliptical galaxy is

$$t_{cross} = \frac{r_e}{\sigma_e}. \tag{2.4.3}$$

For a typical elliptical galaxy, with $r_e = 10 \text{Kpc}$ and $\sigma_e = 200 \text{km/s}$, one obtains a crossing time of $t_{cross} \approx 10^8 \text{years}$. Accordingly, since the time t_{cross} is much smaller than the typical age of a galaxy ($\approx 10^{10} \text{years}$), the elliptical galaxies can be regarded as relaxed systems in dynamic equilibrium under gravity (Schneider, 2006).

2.5 Scaling relations for elliptical galaxies

In this section, the main scaling relations for elliptical galaxies will be presented. These relations describe empirically the regularity of some physical properties of elliptical galaxies.

Faber-Jackson relation: this relation (Faber and Jackson 1976) shows a correlation between the total luminosity L of an elliptical galaxy and the central stellar velocity dispersion σ_0 (i.e., corrected to a circular aperture of radius $R_e/8$):

$$L \propto \sigma_0^\delta. \tag{2.5.1}$$

The value of δ is approximately equal to 4.

Fundamental Plane: the Fundamental Plane (Djorgovski and Davis 1987; Dressler et al. 1987) relates three quantities of elliptical galaxies, namely the half-light

radius R_e (see Eq. (2.4.2)), the mean surface brightness $< \mu_e >$ within R_e, measured in mag/arcsec^2, and the central stellar velocity dispersion σ_0:

$$\log R_e = a \log \sigma_0 + b < \mu_e > +c. \qquad (2.5.2)$$

Bender et al. (1992) obtained $a = 1.4$ and $b = -0.85$ for a sample of elliptical galaxies in the Virgo cluster. Combining the Fundamental Plane with the virial theorem, one can derive an important correlation between the total mass M_{tot} of an elliptical galaxy (within R_e) and its luminosity L:

$$\frac{M_{tot}}{L} \propto L^{\gamma}, \qquad (2.5.3)$$

where $\gamma \approx 0.2$, if $a = 1.4$ and $b = -0.85$.

2.6 Black holes at the center of galaxies?

Black holes are a direct consequence of Einstein's General Relativity, whose surrounding space-time is described by the Schwarzschild metric. This metric identifies a region, namely a sphere of radius R_S[6], within which no particle or photon can escape the gravitational pull of the black hole. There are two main types of black holes: Stellar Black Holes and SuperMassive Black Holes (SMBHs). The former have a mass ranging from $5M_\odot - 30M_\odot$ and originate from the last phases of the life of massive stars. The latter, with a mass of $10^5 M_\odot - 10^9 M_\odot$, are normally found in the galactic nuclei and are thought to have a key-role in the evolution of galaxies. Astrophysicists believe that there is a supermassive black hole at the center of almost every galaxy. This section will briefly go through the sequence of events that have led to this idea.

In 1943, Seyfert observed six galaxies with the following properties: 1) very luminous nucleus in visible light ($L = 10^{43} - 10^{45}$erg/s); 2) strong optical emissions, classified in *narrow* emission lines (Doppler broadening: 10^3km/s) and *broad* emission

---

[6]The Schwarzschild radius of an object with mass M is $R_S = \frac{2GM}{c^2} = 3\frac{M}{M_\odot}$km.

lines (Doppler broadening: 10^4km/s); 3) strong X-ray emission. Today, it is known that this type of galaxies (called Seyfert galaxies) belongs to a wider group: galaxies with Active Galactic Nuclei. These galaxies are characterized by a extremely bright central region and, sometimes, by relativistic jets.

In 1969, Donald Lynden-Bell proposed that AGN are super massive black holes that are increasing their mass through the absorption of surrounding material. Now, astrophysicists believe that this mass accretion is the cause of the enormous luminosity in the center of these galaxies. Super massive black holes can also be quiescent (for example because they run out of surrounding material). In this case, some of these black holes can be detected by studying the dynamics of the stars[7] of the host galaxy. For instance, the mass of the black hole at the center the Milky Way has been estimated by modeling the orbits of approximately 20 stars in the proximity to the black hole. Other mass measurement methods that deserve to be mentioned are based on: 1) MASER emission; 2) reverberation mapping (Peterson, B. M., Ferrarese, L., Gilbert, K. M., et al. 2004).

Only at the end of the 1990s, after the studies which led to the scaling relations discussed in the next section, enough evidence had been collected to confirm that almost every (sufficiently big) galaxy had a super massive black hole in its nucleus.

2.7 $M_\bullet$-σ scaling relation

Thanks to the kinematic data of several galaxies, the measurements of the masses of their central black holes have been collected. In 2000, (Gebhardt et al 2000, Ferrarese & Merritt 2000), an empirical correlation between the mass of the central black hole $M_\bullet$ and the velocity dispersion of the bulge of the host galaxy σ (i.e., integrated out

---

[7]Stars are better kinematics indicators than gas, because gas can be influenced by other elements aside from gravity, such as viscosity or magnetic fields.

to one effective radius, R_e) was discovered:

$$M_\bullet = 10^\alpha \left(\frac{\sigma}{200\text{km/s}} \right)^\beta M_\odot, \tag{2.7.1}$$

where 200km/s is the typical stellar velocity dispersion in a galaxy bluge. The values of the coefficients α and β that will be used in this thesis are obtained by McConnell & Ma (2013). Their work studied a sample of 72 galaxies (fig. 2.4) and found $\alpha = 8.32 \pm 0.05$ and $\beta = 5.64 \pm 0.32$. It can be noticed that, in the case of a typical bulge with $\sigma = 200\text{km/s}$, the mass of the black all is $M_\bullet = 10^{8.32} M_\odot \approx 2 \times 10^8 M_\odot$.

Other correlations between black holes and their host galaxies can be identified. It is worth mentioning the relations $M_\bullet - L_V$ (Kormendy & Richstone 1995) between the black hole mass and the V-band bulge luminosity and $M_\bullet - M_{Bulge}$ (Magorrian et al. 1998) between the black hole mass and the stellar mass of the bulge. Here, these two relations are displayed:

1. $M_\bullet - L_V$

$$\log_{10} \left(\frac{M_\bullet}{M_\odot} \right) = \alpha' + \beta' \log_{10} \left(\frac{L_V}{10^{11} L_\odot} \right). \tag{2.7.2}$$

   From a sample of 44 galaxies (McConnell & Ma 2013), the best-fitting values for the coefficients are: $\alpha' = 9.23 \pm 0.10$ and $\beta' = 1.11 \pm 0.13$.

2. $M_\bullet - M_{Bulge}$

$$\log_{10} \left(\frac{M_\bullet}{M_\odot} \right) = \alpha'' + \beta'' \log_{10} \left(\frac{M_{Bulge}}{10^{11} M_\odot} \right). \tag{2.7.3}$$

   From a sample of 35 galaxies (McConnell & Ma 2013), the best-fitting values for the coefficients are: $\alpha'' = 8.46 \pm 0.08$ and $\beta'' = 1.05 \pm 0.11$.

Recently, a work by Pacucci et al. (2023) analyzed 21 galaxies, with redshift $z = 4 - 7$, that violate the scaling relations of this section. The black holes at the center of these galaxies are found to be more massive than expected. In fact, the $M_\bullet - M_{Bulge}$ relation of this section implies a ratio $M_\bullet/M_{Bulge}$ of around 1/1000, while, the values of this ratio in the galaxies studied by Pacucci et al. (2023) are approximately $1/100 - 1/10$.
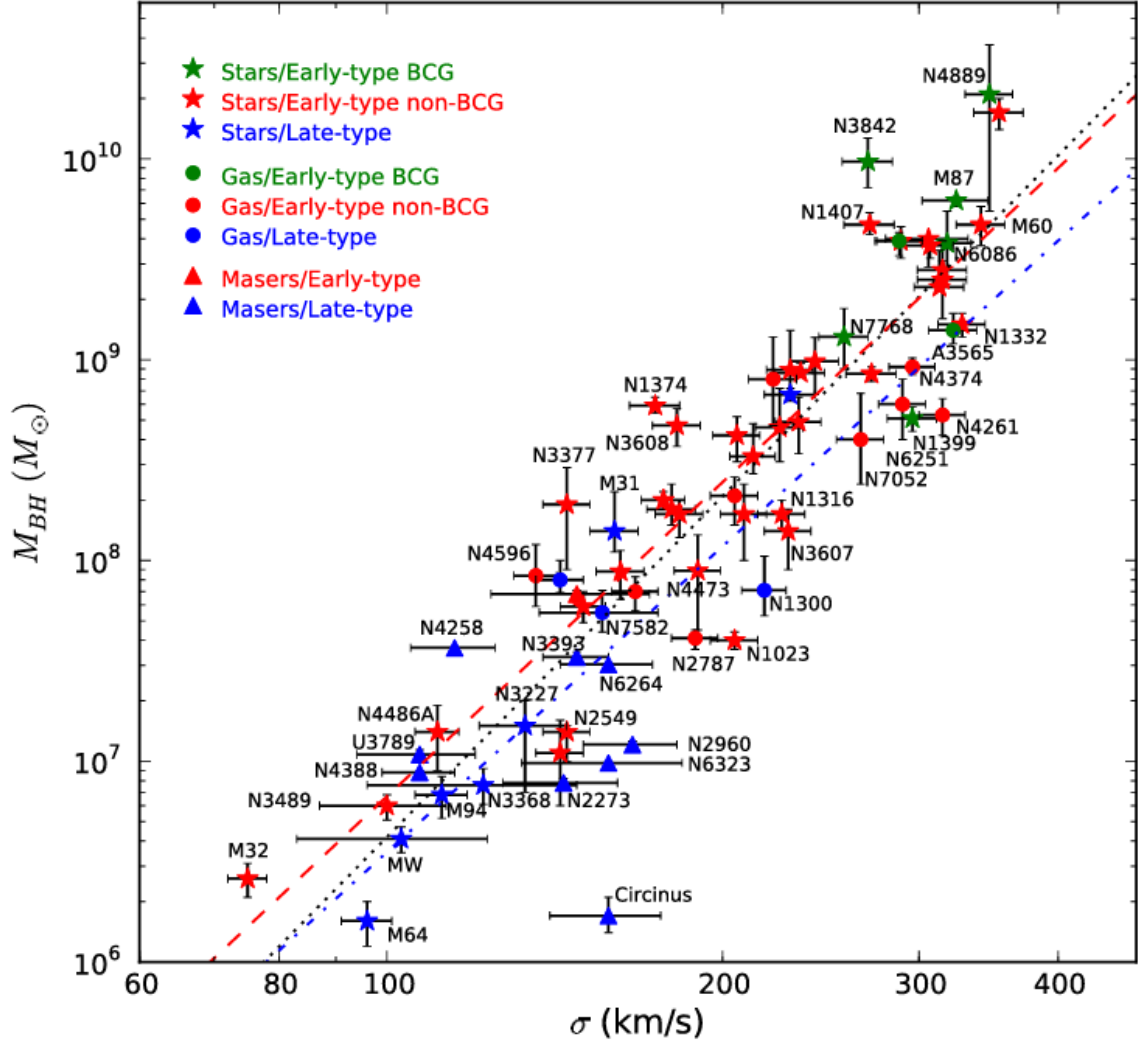
FIGURE 2.3. $M_\bullet - \sigma$ relation of 72 galaxies (listed at http://blackhole.berkeley.edu). The black hole mass measurements are based on the dynamics of MASERs (triangles), gas (circles) and stars (stars). (McConnell & Ma 2013).

Chapter 3

Lensing Model of a Galaxy Cluster

In this chapter, we present the main ingredients of state-of-the-art parametric strong lensing models of galaxy clusters and the steps required for their creation.

A parametric strong lensing model describes the total mass distribution of the cluster with a series of mass components defined by a set of parameters. These parametric mass components can be divided, according to their extent, in two categories: large-scale halos (dark-matter halos and hot-gas halos) and small-scale halos (subhalos associated with the cluster galaxies). Given a set of observed multiple images, produced by the lens cluster, the aim of a strong lensing modeling is to find the values of the model free parameters that best reproduce the positions of these images. From these parameters, one can then estimate the total mass distribution of the galaxy cluster.

The chapter is organized as follows. In Sect. 3.1, we describe the technique used to identify the systems of multiple images. In Sect. 3.2, we present the mass components used to model the mass distribution of a cluster. In Sect. 3.3, we examine a novel method to model the cluster galaxies. In Sect. 3.4, we give an overview of the software used to model the galaxy clusters and the optimization methods. Finally, at the end of this chapter, the strong lensing model of the cluster Abell S1063 will be studied. In particular, in Sect. 3.5, we briefly describe the data used to create the model of Abell S1063 and, in Sect. 3.6, we discuss the strong lensing model of Abell S1063 developed by Bergamini et al. (2019).

3.1 Identification of multiple images

The first step to create a strong lensing model of a galaxy cluster is the identification of the families of multiple images produced by the cluster.

Each family consists of a series of images that originate from the same background source. In order to identify these families, one has to examine the available spectro-photometric data of the cluster obtained for example by the HST (see Fig. 3.1), JWST[1] (see Sect. 3.5), and MUSE (see Sect. 3.5) and group the multiple images belonging to the same family by analyzing their spectrum, color and shape. In fact, since gravitational lensing is an *achromatic* phenomenon (see Sect. 1.1), images of the same source have the same spectrum and, consequently, the same colour. With reference to the morphology of multiple images, although the shape can change among the images of a family (due to lensing distortion), it remains a useful tool to identify the systems of multiple images especially in regions of low magnification. Furthermore, the expected parity from the gravitational lensing theory can be helpful in this process. A redshift value is assigned to each image family based on the measured spectroscopic redshift of their multiple images.

The precise position of the multiple images and the corresponding errors are generally derived from the distribution of their surface brightness, for example, by identifying the peak of their light emission and assuming an uncertainty proportional to the image resolution. The angular coordinates (x, y) are usually measured with respect to the position of the BCG (see Sect. 2.4) of the cluster. We note that, sometimes a single extended lensed source can be further divided in two or more subsystems. For instance, a source with two luminosity peaks can be considered as two separate sources. This method can improve the constraints on the critical lines (Monna et al. 2014).

---

[1]HST stands for Hubble Space Telescope and JWST for James Webb Space Telescope.

FIGURE 3.1. HST color composite image of Abell S1063. The crosses identify the 55 multiple images included in the lens model. The coordinates of the BCG are: R.A. = 22:48:43.970, DEC = -44:31:51.16. North is top and East is left.

3.2   Mass model components

The total mass distribution of a galaxy cluster can be divided in three main mass components (see Chap. 2): the extended dark-matter halos, the diffuse hot-gas halos, and the galaxy-scale sub-halos. These mass components are described with parametric mass profiles.

3.2.1 Cluster members modeling

To model the galactic sub-halos, it is crucial to identify all the galaxies belonging to the cluster. In order to obtain a pure and complete catalog of cluster members, the method described by Grillo et al. (2015) is usually used.

As a first step, the cluster members are selected according to their redshift values. In particular, given the redshift z_{cl} of the cluster, the galaxies with a redshift ranging from $z_{cl} - \Delta z$ to $z_{cl} + \Delta z$ are considered as members. Δz^2 is usually assumed equal to three times the typical velocity dispersion of galaxies in a cluster. These cluster galaxies are called spectroscopic members.

Secondly, Based on the spectroscopic members and their colors, the spectroscopic sample of cluster galaxy is complemented by the photometric members.

Once the cluster galaxies are identified, their mass distributions are modeled. Each cluster galaxy is described by a spherical ($\epsilon = 0$) and coreless ($r_c = 0$) dPIE (see Sect. 1.8). We note that, although the majority of the cluster galaxies are elliptical or lenticular (see Sect. 2.3), it would be incorrect to parametrize them as dPIEs with the ellipticity of their isophotes, because the shape of the total mass (mainly dark matter) of galaxies can be sensibly different from the shape of their surface brightness. Accordingly, one could leave the ellipticity as a free parameter, drastically increasing the number of free parameters (see Sect. 3.3) and the computational time or decide, more wisely, to use spherical dPIEs.

The three-dimensional mass density ρ_{gal} of the spherical and coreless dPIEs is proportional to:

$$\rho_{gal}(r) \propto \frac{1}{r^2(r_t^2 + r^2)}, \tag{3.2.1}$$

where r_t is the cut radius and r the distance from the center of the mass distribution.

---

[2]From some spectroscopic considerations, one can derive the following equation that correlates the redshift dispersion Δz and the velocity dispersion σ_{cl} of galaxies in a cluster: $\sigma_{cl} = \frac{c}{1+z_{cl}}\Delta z$. We recall that the typical velocity dispersion σ_{cl} of a cluster is equal to $\sigma_{cl} \approx 1000 km/s$, rest frame.
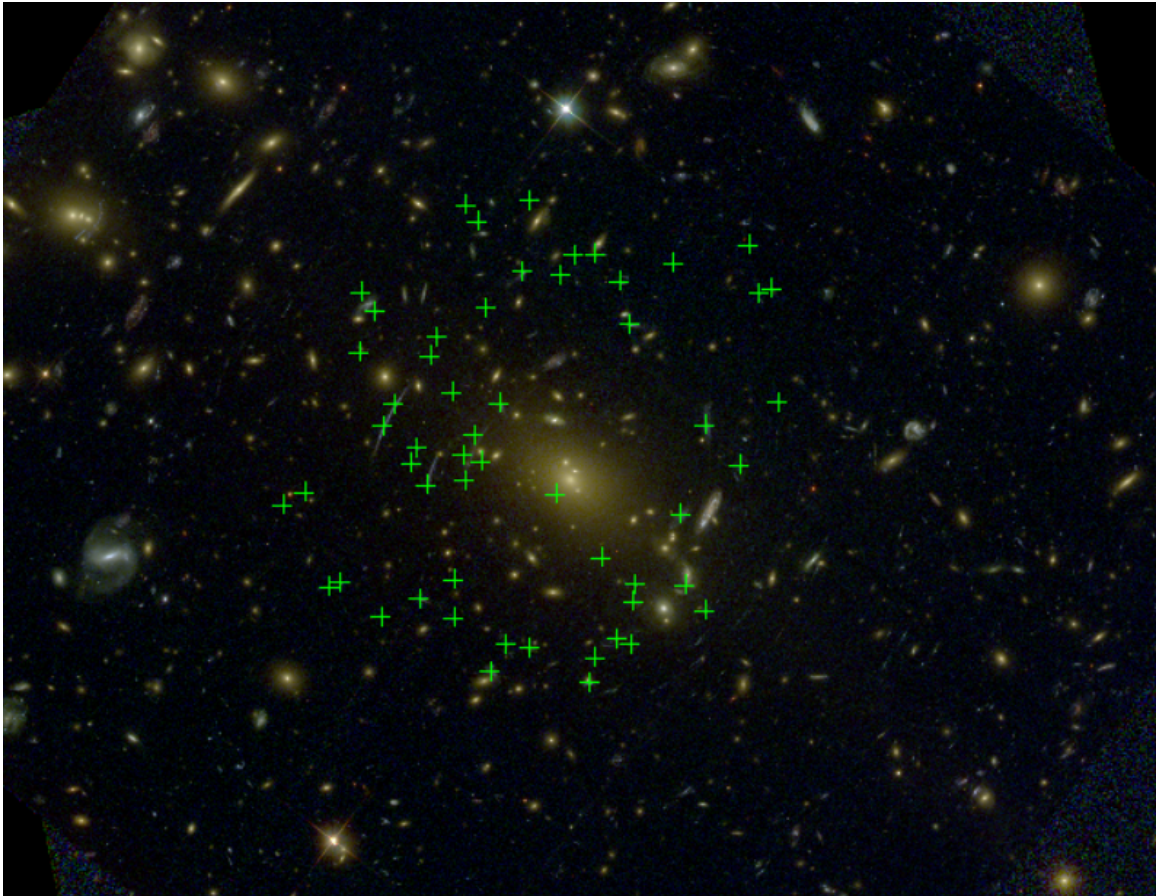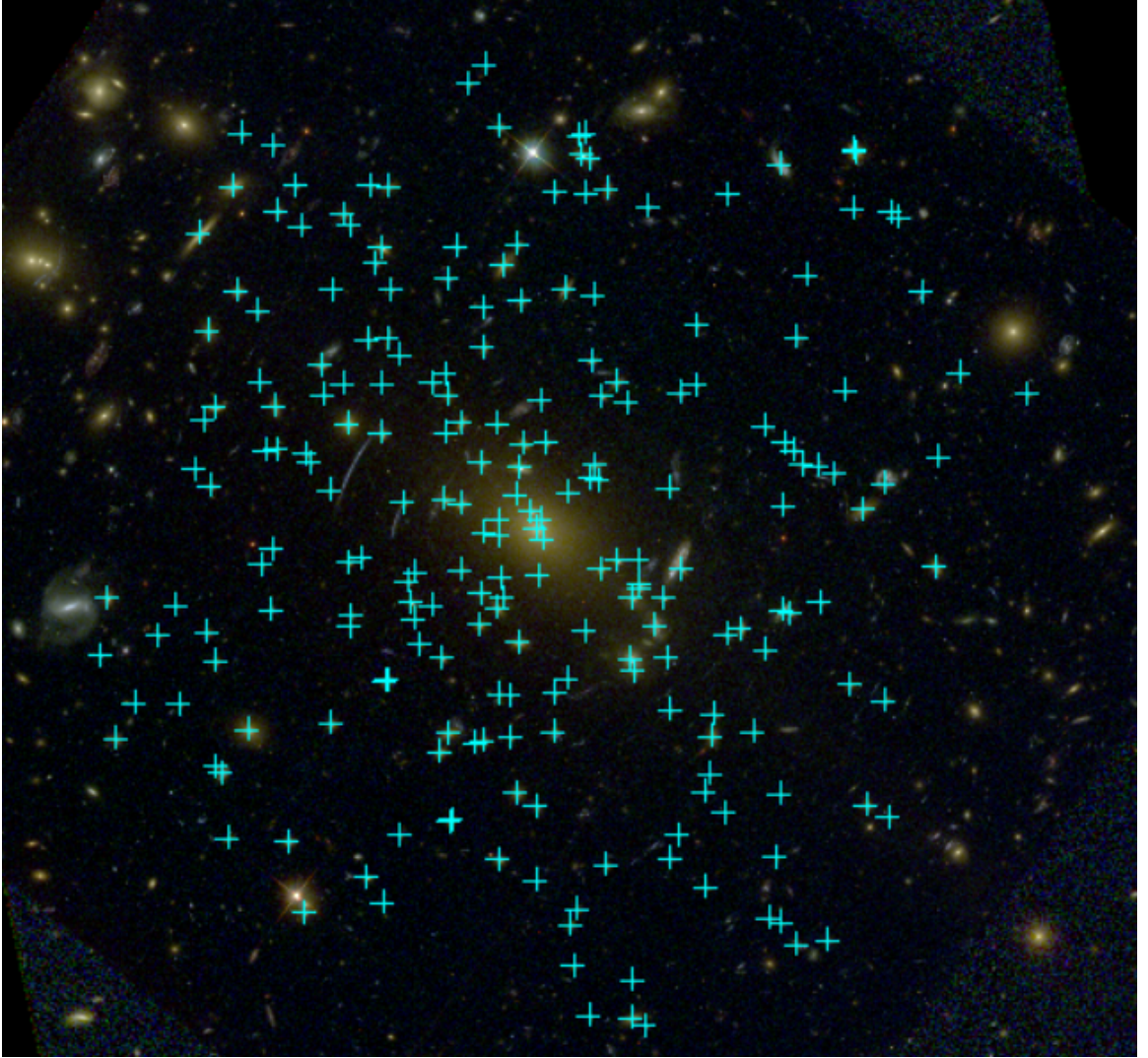
FIGURE 3.2. HST color composite image of Abell S1063. The crosses identify the 222 cluster members included in the lens model. The coordinates of the BCG are: RA = 22:48:43.970, DEC = -44:31:51.16. North is top and East is left.

Depending on the value of r, the behaviour of $\rho_{gal}(r)$ varies: for $r \leq r_t$, $\rho \propto r^{-2}$ (isothermal behaviour) and for $r \geq r_t$, the mass density decreases more rapidly, i.e. $\rho \propto r^{-4}$. The dPIEs of cluster members have 5 fixed parameters: the coordinates (x, y) of the center of the galaxies are known, the core radius r_c and the ellipticity ϵ are both equal to zero, and the position angle θ, for a spherical distribution is needless.

Therefore, each i-th cluster member has two free parameters $\sigma_{0,i}$ and $r_{t,i}$. Since the number of multiple images (model constraints) does not permit constraining the free parameters of each single cluster galaxy, the following two scaling relations are usually adopted to reduce the number of free parameters of the sub-halo component to just two free parameters.

$$\sigma_{0,i} = \sigma_0^{ref} \left(\frac{L_i}{L_0} \right)^{\alpha}, \tag{3.2.2}$$

$$r_{t,i} = r_t^{ref} \left(\frac{L_i}{L_0} \right)^{\beta}. \tag{3.2.3}$$

In this equation, L_i is the luminosity of the i-th cluster member, and the two parameters σ_0^{ref} and r_t^{ref} and the luminosity L_0 are reference values and, normally, they are assumed corresponding to the parameters of the BCG. The luminosities are usually measured in the HST F160W band to minimize the contamination of the blue galaxies near the cluster members. Moreover, they are good tracers of the total mass of the galaxies (see Grillo et al. 2015). Finally, the coefficients α and β are determined from the Faber-Jackson and the fundamental plane relations for elliptical galaxies (see Sect. 2.5).

To summarize, the cluster members are modeled by spherical dPIEs with vanishing core radii. The number of free parameters are reduced to two free parameters thanks to the scaling relations $\sigma_{0,i} - L_i$ and $r_{t,i} - L_i$. We remark that this decrease in the number of free parameters is fundamental because it reduces the model degeneracies (see Sect. 3.3) and prevent the lens model from the overfitting condition.

3.2.2 dark-matter halos modeling

The diffuse dark-matter halos, are modeled by elliptical dPIE mass profiles.

The dPIEs used to describe the dark-matter halos have six free parameters because the cut radius r_t is assumed infinite. The value of r_t is infinite because the radius of

the large-scale dark-matter halos are much larger than the distance from the cluster central region where multiple images occur.

3.2.3 Hot-gas modeling

In the model considered in this thesis, following the work by Bonamigo et al. (2017), the hot-gas mass component of the cluster is modeled separately from the dark-matter component.

In particular, the mass density distribution of the hot gas can be obtained from its X-ray emission (see Sect. 2.2). There are two main techniques to model the hot-gas halo(s): 1) by analyzing the system's geometry, one deprojects the X-ray surface brightness and obtains the hot-gas mass distribution; 2) one models the hot-gas mass distribution (usually with a sum of dPIEs) and projects it onto the lens plane. Then, this projected model is fitted to the observed X-ray surface brightness in order to determine the parameters of the hot-gas mass distribution.

The dPIE parameters associated with the hot-gas halos, obtained from the fit, are fixed (they cannot vary) in the lensing model. This is justified by the few assumptions made to derive the mass distribution from the X-ray surface brightness (assumption on the mechanism of photon emission) and the smaller statistical errors in the hot-gas mass profile compared to the statistical errors of the other components.

3.3 Enhanced modeling of cluster galaxies with galaxy kinematics

In this section, an improved technique, developed by Bergamini et al. 2019, to model the cluster members will be discussed.

This technique make use of the integral field spectroscopic data obtained by MUSE to extract the spectra of a large sample of cluster galaxies within apertures, centered at the centers of the galaxies, with radius R_{ap}. The value of this radius is selected in

order to have spectra with both a good signal-to-noise ratio and a weak contamination from other sources. The spectra of the members that are considerably contaminated by the light of a close brighter galaxy are discarded because the evaluation of their velocity dispersion would be biased.

Then, one has to compute the velocity dispersions of the cluster members. The measurement is performed using the public software pPXF (penalized pixel-fitting, Cappellari & Emsellem 2004). It is important to notice that this measurement returns the average line-of-sight velocity dispersion σ_{ap} of the cluster members within the aperture with radius R_{ap}. Accordingly, in the lensing model, the velocity dispersions σ_{ap} have to be converted into the central velocity dispersions σ_0 of the dPIE profile.

From the kinematic and photometric measurements of the cluster members, one can obtain the coefficient α and the the reference velocity dispersion σ_0^{ref} of the scaling relation (see Eq. (3.2.2)) between the velocity dispersion $\sigma_{0,i}$ of the i-th galaxy and its luminosity L_i. In particular, by fitting this scaling relation to the measured velocity dispersion and luminosity of the cluster members, one finds the value of α and σ_0^{ref} that best reproduce the data. In addition, the fit provides an estimation of the scatter $\Delta\sigma_0$ of the galaxy velocity dispersions with respect to the best-fit relation.

Following this approach, a Gaussian prior on the reference velocity dispersion can be derived. Also, the value of α, obtained from the fit, is considered a fixed parameter in the lens model. With reference to the scaling relation (see Eq. (3.2.3)) between the cut radius $r_{t,i}$ of the i-th galaxy and its luminosity L_i, the reference cut radius r_t^{ref} is free to vary within a large flat prior (as in the method of Sect. 3.2.1), while, the coefficient β is fixed using the following equation:

$$\beta = \gamma - 2\alpha + 1, \tag{3.3.1}$$

where the coefficient γ of the scaling relation between the total mass $M_{tot,i}$ of the i-the galaxy and its luminosity L_i (see Eq. (2.5.4)) is fixed consistently with the fundamental plane relation. The Eq. (3.3.1) can be proven combining the scaling

relations $\sigma_{0,i} - L_i$, $r_{t,i} - L_i$ and $M_{tot,i} - L_i$ with the expression of the total mass of a dPIE (see Eq. (1.8.5)).

Summarizing, the coefficients α and β are fixed, the reference cut radius r_t^{ref} is a free parameter and the reference velocity dispersion σ_0^{ref} has a Gaussian prior.

This approach presents some advantages compared to lens models (see Sect. 3.2.1) that do not employ galaxy kinematics. The most important one is the strong reduction of the degeneracy between the reference velocity dispersion σ_0^{ref} and the reference cut radius r_t^{ref} of the sub-halos. In addition, the transfer of mass between the diffuse mass components and the sub-halos is significantly reduced.

To conclude this section, the $\sigma_0 - r_t$ degeneracy will be discussed further. We recall that the total projected mass (see Eq. (1.8.4)) of a spherical dPIE model, with vanishing r_c, within a projected radius equal to R_E is:

$$M_{tot}(R < R_E) = \frac{\pi \sigma_0}{G} \left(R_E + r_t - \sqrt{r_t^2 + R_E^2} \right). \qquad (3.3.2)$$

In a strong lensing model of a galaxy cluster, the total mass $M_{tot}(R < R_E)$ of a galaxy inside the Einstein radius R_E is well constrained by the multiple images around the galaxy. Therefore, since many combinations of σ_0 and r_t can provide the same value of the total mass (as it can be seen in Eq. (3.3.2)), a strong lensing model is affected by the degeneracy between σ_0 and r_t.

3.4   Lenstool

This section will briefly discuss the public software Lenstool (Kneib et al. 1996; Jullo et al. 2007) used to perform strong lensing modeling of galaxy clusters.

In the parametric software Lenstool, the mass distribution of the lens is described by parametrized mass profiles. As discussed in the previous sections, in the case of galaxy clusters, the mass distribution is decomposed in different mass components, each one defined by a set of parameters. Once the mass distribution of the lens

is modeled, Lenstool constrains the model parameters $\boldsymbol{\eta}$ using the positions of the observed multiple images (see Sect. 3.1). The best-fit model parameters are obtained by minimizing the distance between the observed image positions and the model-predicted image positions. In particular, Lenstool uses a Bayesian Markov Chain Monte Carlo (MCMC) approach in order to find the set of parameters that maximizes the the posterior probability distribution function $P(\boldsymbol{\eta}|\boldsymbol{d_{pos}})$ of the model parameters $\boldsymbol{\eta}$, given the observed multiple image positions $\boldsymbol{d_{pos}}$:

$$P(\boldsymbol{\eta}|\boldsymbol{d_{pos}}) \propto P(\boldsymbol{d_{pos}}|\boldsymbol{\eta})P(\boldsymbol{\eta}), \tag{3.4.1}$$

where[3] $P(\boldsymbol{\eta})$ is the *prior* probability distribution function (hereafter only *prior*) and $P(\boldsymbol{d_{pos}}|\boldsymbol{\eta})$ is the *likelihood* function and is defined as:

$$P(\boldsymbol{d_{pos}}|\boldsymbol{\eta}) = \prod_{j=1}^{N_{fam}} \frac{1}{\prod_{i=1}^{N_{images}^j} \sqrt{2\pi}\sigma_{ij}^{obs}} e^{-\frac{1}{2}\chi_j^2(\boldsymbol{\eta})}, \tag{3.4.2}$$

where

$$\chi_j^2(\boldsymbol{\eta}) := \sum_{i=1}^{N_{images}^j} \frac{|\boldsymbol{R}_{ij}^{obs} - \boldsymbol{R}_{ij}^{pred}(\boldsymbol{\eta})|^2}{(\sigma_{ij}^{obs})^2}; \tag{3.4.3}$$

where N_{fam} is the number of families of multiple images, N_{images}^j is the number of images belonging to the j-th family, $\boldsymbol{R}_{ij}^{obs} = (x_{ij}^{obs}, y_{ij}^{obs})$ is the observed position of the i-th image of the j-th family, $\boldsymbol{R}_{ij}^{pred}$ is the model-predicted image position and σ_{ij}^{obs} is the uncertainty of the observed image position.

In Lenstool, there are several optimization methods. One of these is based on the *BayeSys* algorithm by John Skilling (2004) and uses 10 interlinked MCMC chains to sample the parameters space and to obtain the posterior probability distribution of the model parameters.

Normally, a strong lensing model of a galaxy cluster is optimized twice. In the first optimization, the uncertainty σ^{obs} of the observed image positions is set equal to

---

[3]In Eq. (3.4.1), the Bayes' theorem has been used.

an arbitrary value (e.g. the instrumental error of the telescope). While, in the second optimization, the uncertainty σ^{obs} is multiplied by a constant factor in order to obtain a value of the best-fit chi-squared approximately equal to the number of degrees of freedom dof[4] of the lensing model. By doing so, one accounts for some simplifying assumptions of the model, such as neglected line-of-sight mass structures or small dark-matter clumps. Also, since the size of the marginalized posterior distribution of the model parameters are correlated with the uncertainty σ^{obs}, one obtains realistic uncertainties from a MCMC analysis.

Another useful quantity to quantify the "goodness" of a model is:

$$\Delta_{rms} = \sqrt{\frac{1}{N_{pos}} \sum_{i=1}^{N_{pos}} |\boldsymbol{R}_i^{obs} - \boldsymbol{R}_i^{best}|^2} \qquad (3.4.4)$$

where N_{pos} is the number of multiple images, $\boldsymbol{R}_i^{obs}$ is the position of the observed i-th image and $\boldsymbol{R}_i^{best}$ is the position predicted by the best-fit model. Δ_{rms} is used to compare different models because it does not depend on the uncertainty of the observed image positions σ_i^{obs}.

Finally, it is important to notice a technical aspect about Lenstool: the velocity dispersion used in Lenstool σ_{LT} is related to the line-of-sight central velocity dispersion σ_0 through the following equation:

$$\sigma_0 = \sqrt{\frac{3}{2}} \sigma_{LT}. \qquad (3.4.5)$$

3.5 Data and Instruments

The strong lensing model of Abell S1063 is based on high-quality spectroscopic and photometric data.

The imaging data are provided by the Hubble Space Telescope (HST). Abell S1063 was observed in 16 broadband filters, from the UV to the near-IR, with the cameras

---

[4]The number of degrees of freedom is given by the number of images minus the number of model parameters. We emphasize that the source positions count as model parameters.

ACS and WFC3 during the Cluster Lensing And Supernova survey with Hubble (CLASH; P.I.:M. Postman; Postman et al. 2012a). The camera ACS has two working channels, namely the Solar Blind Camera (SBC) and the Wide Field Camera (WFC with pixel scale: $0.049''$, wavelength range: $3700\text{Å} - 11000\text{Å}$, FoV: $202'' \times 202''$). The camera WFC3 has two working channels: one for ultraviolet and visible light (UVIS, pixel scale: $0.04''$, wavelength range: $2000\text{Å} - 10000\text{Å}$, FoV: $160'' \times 160''$) and another for near infrared (NIR, pixel scale: $0.13''$, wavelength range: $8500\text{Å} - 17000\text{Å}$, FoV: $123'' \times 137''$). Moreover, another HST program, namely the Hubble Frontier Fields (HFF; P.I.: J. Lotz 2017) campaign, provided data in 7 broadband filters.

With reference to spectroscopic data, the majority was obtained with the ESO's Very Large Telescope (VLT). In particular, the VLT's VIsible wide field imager and MultiObject Spectrograph (VIMOS; Le Fèvre et al. 2003) and Multi-Unit Spectroscopic Explorer (MUSE; Bacon et al. 2012) collected high-quality spectroscopic data of Abell S1063. MUSE is an integral field spectrograph with two working modes: 1) the Wide Field Mode (WFM, a spectral range of $4750\text{Å} - 9350\text{Å}$, a maximum FoV of about $1 \times 1\text{arcmin}^2$ and a pixel size of around $0.2''$); 2) the Narrow Field Mode (NFM, a spectral range of $4750\text{Å} - 9350\text{Å}$, a maximum FoV of about $7.5 \times 7.5\text{arcsec}^2$ and a pixel size of around $0.025''$). VIMOS (spectral range of $3600\text{Å} - 10000\text{Å}$) was used during the campaign called CLASH-VLT (P.I.: P.Rosati; Rosati et al. 2014), which combined photometric and spectroscopic data.

Finally, X-ray data are crucial in modeling the hot-gas halos. In the case of Abell S1063, most of the X-ray data were collected with the space telescope Chandra[5] X-ray Observatory (spectral range: $0.1\text{keV} - 10\text{keV}$).

---

[5]https://cxc.harvard.edu/index.html.

3.6 Strong lensing model of Abell S1063

This section will describe the main characteristics of the strong lensing model of the galaxy cluster Abell S1063 developed by Bergamini et al. in 2019.

According to the classification by Rood and Sastry (see Chap. 2), Abell S1063 is a cD galaxy cluster at redshift $z_{cl} = 0.348$. Its optical and X-ray surface brightness are quite symmetric around the central BCG (RA = 342.18321, DEC = -44.530878). Abell S1063 has a mass M_{200_c} within a sphere of radius R_{200_c}[6] equal to $M_{200_c} = (2.03 \pm 0.67) \times 10^{15} M_{\odot}$ (Umetsu et al. 2014).

Following the methodology of Sect. 3.1, 55 multiple images, divided in 20 families, have been identified from the HST images of the cluster (Fig. 3.1). Moreover, from the spectra of the images, the redshift values of the 20 sources have been measured. These redshift values range from 0.73 to 6.11.

In this strong lensing model, 222 cluster members (spectroscopic[7] and photometric members) have been selected. For each galaxy, its luminosity in the F160W band have been measured. The average line-of-sight velocity dispersion σ_{ap} (see Sect. 3.3) of 37 galaxies within an aperture with radius equal to $R_{ap} = 0.8''$ have also been measured.

The 222 galaxies have been modeled with 222 spherical dPIEs with vanishing core radius r_c. The cut radius $r_{t,i}$ and the central velocity dispersion $\sigma_{0,i}$ of every cluster galaxy follow the scaling relations of Sect. 3.2.1, where the reference parameters r_t^{ref}, σ_0^{ref}, and L_0 correspond to the parameters of the BCG of Abell S1063. Then, applying the method of Sect. 3.3, the $\sigma_{0,i} - L_i$ scaling relation parameters have been determined from the measurements of the luminosity and the velocity dispersion of the cluster galaxies. The obtained best-fit values (see Eq. (3.4.6)) are: $\alpha = 0.27$, $\sigma_{LT} = 310\text{km/s}$, $\Delta\sigma_{LT} = 15\text{km/s}$, where $\Delta\sigma_{LT}$ is the scatter associated with the

---

[6] R_{200_c} is defined as the radius of the sphere inside which the mean density is equal to 200 times the Universe critical density at the redshift z_{cl}. In the case of Abell S1063, this radius is equal to $R_{200_c} = (2.32 \pm 0.26)\text{Mpc}$ (Umetsu et al. 2014).

[7] In the case of Abell S1063 ($z_{cl} = 0.348$), the spectroscopic members are the galaxies with a redshift value belonging to the following range: $(0.348 - 0.0135, 0.348 + 0.0135)$.

reference velocity dispersion σ_{LT}. The coefficient γ of the $M_{tot,i} - L_i$ scaling relation (see Eq. (2.5.3)) has been fixed consistently with the fundamental plane relation of the cluster galaxies: $\gamma = 0.2$. Consequently, the coefficient β of the $r_{t,i} - L_i$ relation (see Eq. (3.2.3)) has been obtained from Eq. (3.3.1): $\beta = \gamma - 2\alpha + 1 = 0.66$.

Summarizing, the 222 cluster members of Abell S1063 are modeled with 222 spherical dPIEs with vanishing core radius r_c. The number of free parameters of these dPIEs are strongly reduced using the scaling relations $\sigma_{0,i} - L_i$ and $r_{t,i} - L_i$. Furthermore, thanks to the method described in Sect. 3.3, the coefficients of these scaling relations ($\alpha = 0.27$ and $\beta = 0.66$) and the mean and the standard deviation of the reference velocity dispersion ($\sigma_{LT} = (310 \pm 15)$km/s) are found. Therefore, there are only 2 parameters allowed to vary: 1) the reference cut radius r_t^{ref} with a flat prior probability distribution; 2) the reference velocity dispersion σ_0^{ref} with a Gaussian prior characterized by the mean and the standard deviation obtained from the fit.

Besides the cluster members, the total mass distribution of Abell S1063 is described by two large-scale dark-matter halos, and three diffuse hot-gas halos. The first dark-matter halo consists of an elliptical dPIE with infinite cut radius and the three hot-gas halos are modeled by elliptical dPIEs with fixed parameters (see Sect. 3.2.3). The second large-scale halo is a spherical dPIE with vanishing core radius and infinite cut radius. The second dark-matter halo has been introduced to reduce the offset between the observed and model-predicted image positions in the North-East region of Abell S1063. The first large-scale dark-matter halo has 6 free parameters, the hot-gas halos have zero free parameters and the second dark-matter halo has 3 free parameters. Therefore, this strong lensing model is characterized by 10 parameters with flat priors and 1 parameter with a Gaussian prior.

Once the model was defined, it was optimized with Lenstool. In the first optimization, the uncertainty of the image positions observed with HST was fixed to $0.5''$, while, those observed with MUSE (see Sect. 3.5) was fixed to $1.0''$. In a second step, these values were modified in order to obtain a best-fit chi-squared value comparable

to the number of degrees of freedom of the lensing model. In this strong lensing model, the number of degrees of freedom is equal to $55 \times 2 - (20 \times 2 + 11) = 59$.

In conclusion, the second optimization, obtained from a MCMC chain of 10^5 samples, provided a best-fit chi-squared of 69.90 and a $\Delta_{rms} = 0.55''$.

Chapter 4

Procedure

The aim of this thesis is to investigate the impact, in strong lensing models of galaxy clusters, of the presence of supermassive black holes at the center of cluster galaxies. In order to reach this goal, we use the public software Lenstool (see Sect. 3.4) to examine the differences between a strong lensing model that ignores the existence of supermassive black holes and a model that takes them into account. We study the possible variations, between these models, in the predicted multiple image positions and in the values, obtained through a strong lensing analysis, of the model parameters that describe the total mass distribution of a cluster. In particular, this study has been carried out in the specific case of the galaxy cluster Abell S1063 (see Sect. 3.6).

Throughout this thesis, we assume a flat ΛCDM cosmology with $\Omega_m = 0.3$ and $H_0 = 70$ km/s Mpc^{-1}. In this cosmology, $1'' = 4.92$ kpc at the redshift, z_{cl}, of Abell S1063 ($z_{cl} = 0.348$). The magnitudes are given in the AB system. Moreover, we assume that the origin of the lens plane corresponds to the center of the BCG of Abell S1063 (RA = 342.18321, DEC = -44.530878). Finally, we emphasize that the velocity dispersion values in this chapter represent the physical values, not the values of Lenstool (see Sect. 3.4).

This chapter will describe the steps taken that led to the results of this thesis, which will be shown in Chap. 5. Sect. 4.1 will discuss how supermassive black holes have been modeled. Then, in the following sections (Sect. 4.2, Sect. 4.3, Sect. 4.4), the analyses made in this thesis will be explained.

4.1 Black holes modeling

This section will explain how we model the supermassive black holes at the center of the cluster galaxies.

First of all, we estimate the value of the mass $M_\bullet$ of the black holes using the $M_\bullet - \sigma$ relation (see Sect. 2.7), where σ is the value of the velocity dispersion of the bulge of their host galaxies. The values of the velocity dispersion of the cluster galaxies have been obtained from the best-fit[1] model of Abell S1063 (Bergamini et al., 2019, see Sect. 3.6). We notice that these velocity dispersions are not the actual physical quantities, instead, they are the Lenstool parameters σ_{LT} (see Sect. 3.4). Accordingly, one has to transform the values of the Lenstool velocity dispersions σ_{LT} into their physical quantities σ_0. Furthermore, we emphasize that we assume that the central velocity dispersions, σ_0 (see fundamental plane, Sect. 2.5), of the galaxies are equal to the velocity dispersions σ (see $M_\bullet - \sigma$ relation, Sect. 2.7) of the bulge of the galaxies.

We model the black holes with point-mass components. Practically, the black holes have been modeled with tiny spherical dPIE profiles centered at the center of the cluster members, because the point-mass profile does not seem to work properly in Lenstool. After several tests, we have found the suitable values of the cut radius r_t^{BH} and core radius r_c^{BH}. Namely, $r_t^{BH} = 4.921 \times 10^{-4}$ kpc $= 1.000 \times 10^{-4}$ arcsec and $r_c^{BH} = 4.921 \times 10^{-6}$ kpc $= 1.000 \times 10^{-6}$ arcsec for all the dPIEs. These values are the smallest values that have been verified to work well in Lenstool.

Then, in order to completely determine the dPIE profiles that describe the black holes, we have to compute one last parameter value for each black hole, namely that of the velocity dispersion σ_{BH}. Obviously, σ_{BH} has no physical meaning. It is only a parameter used to define the total mass of the dPIE profile. For each black hole, the parameter σ_{BH} can be computed from Eq. (1.8.5), where the cut radius r_t is

---

[1]The data of the best-fit model of Abell S1063, by Bergamini et al. (2019), are available at https://www.fe.infn.it/astro/lensing/.

$r_t = r_t^{BH} = 4.921 \times 10^{-4}$ kpc and the total mass M_{tot} of the dPIE (i.e. the mass $M_\bullet$ of the black hole) has been obtained through the $M_\bullet - \sigma$ scaling relation.

It is worth noting that the total mass M_{tot} of a dPIE profile is the mass enclosed within a sphere with infinite radius, not by a sphere of radius r_t. However, it can be shown that, in this case, where $r_t^{BH} = 100 r_c^{BH}$, the projected mass $M(R)$ (see Eq. (1.8.4)) of a dPIE within a circle with a projected radius R of $5 r_t^{BH}$ is approximately equal to 90% of the total mass M_{tot}. Within $10 r_t^{BH}$, there is $\approx 95\%$ of the total mass M_{tot}. Therefore, $\approx 95\%$ of the mass $M_\bullet$ of every black hole is contained in a circle with a projected radius of $10 r_t^{BH} = 4.921 \times 10^{-3}$ kpc.

To summarize, the supermassive black holes have been modeled as small spherical dPIEs centered at the center of their host galaxies with fixed values of cut and core radii. Finally, for every black hole, the dPIE parameter σ_{BH} has been obtained through the $M_\bullet - \sigma$ scaling relation in order to set its total mass.

4.2   First methodology

First of all, we collect all the information of the strong lensing model by Bergamini et al. (2019), which is discussed in Sect. 3.6. We recall that this model (hereafter B19) describes the total mass distribution of Abell S1063 with 10 free parameters, with flat priors, and one parameter (namely the velocity dispersion of the reference galaxy, σ_{gal}), with a Gaussian prior. The model B19 considers 222 spherical sub-halos corresponding to the 222 cluster members, 3 elliptical hot-gas halos, one elliptical dark-matter halo and an additional small-scale spherical dark-matter halo. Moreover, 55 multiple images, originating from 20 sources, are detected.

Secondly, we produce another model using the model B19 as starting point. In addition to the mass components of the model B19, we introduce the supermassive black holes. To be specific, we add a black hole at the center of every cluster member using the $M_\bullet - \sigma$ scaling relation, as explained in Sect. 4.1. Hereafter, this model

will be denoted as BH.

Thirdly, we decide to develop a third strong lensing model. This is identical to the model BH, apart from the mass values of the black holes. In this third model, the masses of black holes are ten times larger than those of the model BH. We build this model (hereafter BH10) to study the effects of a more massive population of black holes in a strong lensing model of Abell S1063.

It can be noticed that each of the three models has the same number of free parameters (11 describing the total mass profile of the cluster and 40 corresponding to the coordinates of the 20 sources) and, thus, of degrees of freedom (59), since the black hole parameters are determined by the $M_\bullet - \sigma$ scaling relation. In contrast, the difference between the model B19 and BH (or BH10) is that, the former models the total mass distribution of each cluster galaxy (which is made mainly of dark matter, stars and a central black hole) with a spherical sub-halo, while, the latter with a spherical sub-halo and a black hole. Therefore, the sub-halos of the model B19 describe the total mass of the cluster members, whereas, the sub-halos of the model BH (or BH10) describe the total mass of the galaxies except for the mass of their supermassive black hole, which is considered separately.

Once we have prepared the three strong lensing models of the cluster (B19, BH and BH10), we optimize them, based on the 55 observed multiple image positions (see Sect. 3.1). In Sect. 3.6, the optimization process of the model B19 is described. In the same way, we use Lenstool to run the optimizations of the models BH and BH10, which were obtained from MCMC chains of 10^5 samples, besides the burn-in phase (see Sect. Lenstool). Also, we retain the same positional uncertainty σ^{obs} of the observed images of the model B19. This has been justified a posteriori verifying that the minimum values of the χ^2 were similar between the three optimized models: $\chi^2_{B19} = 69.90$, $\chi^2_{BH} = 69.12$ and $\chi^2_{BH10} = 69.41$.

During the optimization of a model, Lenstool samples the posterior probability distribution of the parameters that describe the total mass distribution of a cluster.

In particular, we focus on the parameters that define the total mass distribution of the cluster galaxies, which should be more affected by the presence of the black holes than the mass distribution of the extended halos. Moreover, the optimization provides the multiple image positions predicted by the three best-fit[2] models.

After comparing the model-predicted image positions and the posterior probability distribution of the parameters of the three optimized models, we use Lenstool (specifically the MARCHINGSQUARE algorithm) to compute their critical curves and caustics. In particular, we obtain the critical curves and caustics of each best-fit model (hereafter fB19, fBH and fBH10) for a source at redshift $z = 0.7301$. One could have fixed another redshift, nevertheless, we have verified that another redshift would have not significantly modified our results.

Lastly, analyzing the critical curves in the three models, we notice that almost every cluster galaxy produces a tangential and a radial critical curve around it. Consequently, driven by the idea that the effects of the presence of black holes should be stronger on the small spatial scale of a cluster, we decide to create a list of point-like sources (14 sources) located close to the caustics corresponding to some cluster members, at redshift $z = 0.7301$. The sources were positioned near the caustics of the cluster galaxies whose critical curves and caustics differ significantly between the three models.

We recall that, if a source lies close to a caustic on the source plane, its images will occur near the corresponding critical curve on the lens plane. With this list of sources, we compute their multiple image positions produced by the total mass distributions of the three best-fit models of B19, BH, and BH10. In particular, Lenstool inverts the lens map (see Sect. 1.2) and calculates numerically the image positions $\boldsymbol{x}$ for each source position $\boldsymbol{y}$, given a mass distribution. Finally, we study these image positions predicted by the three models.

---

[2]As explained in Sect. 3.4, the best-fit model is determined by the set of parameters $\boldsymbol{\eta}$ that minimizes the distance between the observed and model-predicted multiple image positions.

4.3 Second methodology

This section will discuss a second set of analyses.

We recall that the optimization of the model B19, in Sect. 4.2, determined the values of the model parameters that best reproduce the observed multiple image positions and, consequently, best describe the mass distribution of the galaxy cluster. Starting from the best-fit model of B19, we develop a new model of the cluster mass distribution: we introduce the 222 supermassive black holes (modeled in Sect. 4.1) into the best-fit model of B19. We emphasize that this mass model is totally characterized by fixed parameters: the values of the parameters that describe the extended halos and the cluster galaxies are defined by the optimization of the model B19 and those of the parameters of the black holes are set through the $M_\bullet - \sigma$ scaling relation. Similarly to the previous section, we decide to produce a third model, whose black holes are ten times more massive. For clarity, we will refer to the best-fit model of B19 as fB19[3], to the mass model with black holes (fB19 + black holes) as fB19+bh and to the model with the heavier black holes (fB19 + heavier black holes) as fB19+10bh.

Afterwards, given the mass distribution of the model fB19 (not the actual galaxy cluster) and the 55 observed images, we obtain the positions of the sources of these images (as if they were produced by fB19). In order to find these source positions $\boldsymbol{y}$, Lenstool exploits the lens map $\boldsymbol{y} = \boldsymbol{x} - \boldsymbol{\alpha}(\boldsymbol{x})$ (see Sect. 1.2), where $\boldsymbol{x}$ are the observed image positions and the deflection angle $\boldsymbol{\alpha}(\boldsymbol{x})$ at $\boldsymbol{x}$ is determined by the total mass distribution of the cluster (which is described with the model fB19). Practically, Lenstool computes a source position for each image, not a source for each family of multiple images. Accordingly, we obtain the final list of source positions by calculating the average between the source positions for each family.

Once we have the source positions and the three different mass distributions de-

---

[3]The letter f in this notation stands for *fixed*, since fB19 refers to a mass model with no parameters that are allowed to vary.

scribed by the mass models fB19, fB19+bh and fB19+bh10, we calculate the positions of the images, produced by each mass distribution, of these sources. As explained in the previous section, Lenstool inverts the lens map and calculates the predicted image positions[4] for each source, given the mass distribution of the cluster.

After comparing the predictions of the image positions among the three models, similarly to what we have done in Sect. 4.2, we compute the caustics and the critical curves, at redshift $z = 0.7301$, produced by the mass models fB19, fB19+bh and fB19+bh10. Then, for each model, we calculate the predicted image positions of the list[5] of point-like sources that are located close to the caustics generated by the galaxies. Once again, we compare the data obtained with the three models.

To conclude, it is important to clarify the idea behind this methodology. As explained in the previous section, the galactic sub-halos of the best-fit model fB19 describe the total mass of the cluster galaxies (dark matter, stars and supermassive black hole). Accordingly, the model fB19+bh (and fB19+bh10) introduces the population of supermassive black holes to a model (fB19) that already contains its mass inside the sub-halos. In fact, the aim of this section is not to accurately reproduce the mass distribution of Abell S1063, instead, it is to study how the extra masses (central black holes) in the simulated cluster galaxies of fB19+bh (and in fB19+bh10) modify the predictions of fB19.

4.4 Third methodology

In the previous section, after obtaining the positions of the sources (whose images, produced by the mass distribution of the model fB19, are the 55 observed images), we found their predicted image positions with the mass models fB19, fB19+bh and fB19+bh10. In this third part of analyses, conversely to what has been done in

---

[4]We found 62 images in total.

[5]This list of sources is the same list of Sect. 4.2, since these sources remains in the proximity of the caustics computed in the current section.

Sect. 4.2, we optimized the model B19, not on the observed images, but on the aforementioned predicted images of fB19+bh and, then on the predicted multiple images of fB19+bh10. For clarity, we will denote the best-fit model obtained from the optimization of the model B19 on the predicted images of fB19+bh as fB19(bh), while, the best-fit model obtained from the optimization of B19 on the predicted images of fB19+bh10 as fB19(bh10).

The difference between the optimization of Sect. 4.2 and this one is the following: the former has been performed on observed multiple images produced by the actual Abell S1063, whose total mass distribution is unknown, on the contrary, the latter has been obtained using the images predicted by the known total mass distribution fB19+bh (or fB19+bh). Accordingly, in this case, we have studied in which way a model, which neglects the presence of black holes (i.e. B19), manages to recreate the images produced by a cluster total mass distribution (fB19+bh or fB19+bh10) that contains black holes.

The two optimizations were performed with Lenstool similarly to the ones of Sect. 4.2. The only differences are: 1) the adopted value of the positional error σ^{obs} of the images. In this case, we set a smaller error ($\sigma^{obs} = 0.01''$), because we are modeling a simulated galaxy cluster whose total mass distribution is known. Accordingly, the B19 model describes the cluster total mass distribution of fB19+bh (or fB19+bh10), which produces the multiple images, much more accurately than how B19 describes the actual galaxy cluster Abell S1063. 2) The number of degrees of freedom dof has changed. In the models of Sect. 4.2, $dof = 59$ (55 images from 20 families and 11 free parameters that describe the cluster total mass distribution), while, here $dof = 73$ (62 images, 20 families and 11 parameters that describe the cluster total mass distribution). As discussed in Sect. 4.2, after optimizing the model B19, Lenstool provides the image positions predicted by the best-fit model. Furthermore, Lenstool samples the posterior probability distribution of the model parameters.

Therefore, we compare the predicted image positions of the best-fit models fB19(bh) and fB19(bh10) with the image positions predicted by the mass models of the previous section, namely fB19+bh and fB19+bh10. Finally, we study the posterior probability distributions of the parameters of fB19(bh) and fB19(bh10) and we confronted them with the values of the parameters of fB19+bh and fB19+bh10.

Chapter 5

Results and Discussion

In this chapter, we will present and discuss the relevant results of the analyses described in Chap. 4. In particular, Sect. 5.1, Sect. 5.2, and Sect. 5.3 show and examine the results obtained from the studies in Sect. 4.2, Sect. 4.3, and Sect. 4.4, respectively.

5.1 First methodology

We recall that we optimized the three models B19, BH, and BH10 assuming the positions of 55 observed multiple images as model constraints (see Sect. 4.2). For each best-fit model, which is determined by a set of parameters $\boldsymbol{\eta_{best}}$, we obtained its chi-squared value $\chi^2(\boldsymbol{\eta_{best}})$. Since the three best-fit models (fB19, fBH, fBH10) have the same number of degrees of freedom (i.e. 59), their reduced[1] chi-squared values $\tilde{\chi}^2(\boldsymbol{\eta_{best}})$ are: $\tilde{\chi}^2_{B19}(\boldsymbol{\eta_{best}}) = 1.185$, $\tilde{\chi}^2_{BH}(\boldsymbol{\eta_{best}}) = 1.172$ and $\tilde{\chi}^2_{BH10}(\boldsymbol{\eta_{best}}) = 1.176$. It can be noticed that, although these values are similar, the following inequality holds: $\tilde{\chi}^2_{BH}(\boldsymbol{\eta_{best}}) < \tilde{\chi}^2_{BH10}(\boldsymbol{\eta_{best}}) < \tilde{\chi}^2_{B19}(\boldsymbol{\eta_{best}})$. This means that the best-fit model fBH, which takes into account the presence of the supermassive black holes at the center of the cluster galaxies, reproduces the observed multiple image positions slightly better than the best-fit model fB19, which ignores their presence. Even the best-fit model fBH10, which considers the black holes with larger masses, is a little better in recreating the observed multiple images than fB19.

As reported in Sect. 4.2, Lenstool provides the values, expressed in arcseconds and rounded up to the second decimal digit, of the multiple image positions predicted

---

[1]The reduced chi-squared, $\tilde{\chi}^2$, is equal to χ^2/dof; where χ^2 is the chi-squared and dof is the number of degrees of freedom.

by the optimized models. From these data, we computed the root-mean-square displacement, Δ_{rms}, (see Eq. (3.4.5)) between the positions of the images reproduced by the best-fit model and the observed image positions, for each of the three best-fit models fB19, fBH and fBH10. We obtained the same root-mean-square value for each model: $\Delta_{rms}^{fB19} = \Delta_{rms}^{fBH} = \Delta_{rms}^{fBH10} = 0.55''$. Thus, the presence of the supermassive black holes in the lens models BH and BH10 does not affect the first two decimal digits of the Δ_{rms}.

In Sect. 4.2, we computed the caustics and the critical curves of the three best-fit models (fB19, fBH, and fBH10), for a source at redshift $z = 0.7301$ (see Sect. 4.2). Analyzing these curves, we observe non-negligible differences between the caustics of the three best-fit models. In particular, we observe that the caustics produced by the cluster galaxies of the three best-fit models appear at slightly different positions. Moreover, the radial caustics (critical curves) produced by the cluster galaxies are often absent in the model fBH10 and, sometimes, even in the model fBH. As an example, the caustics produced by the cluster galaxy 7484^2 of fB19, fBH, and fBH10 are shown in Fig. 5.1.

The disappearance of the radial caustics (critical curves) in the models fBH10 and fBH, is caused by the presence of the supermassive black holes at the center of the cluster galaxies. In fact, these black holes, which are taken into account in the models BH10 and BH (but not in the model fB19), increase the central mass density of their host galaxy and, consequently, can prevent the formation of the radial caustic (critical curve) around the galaxy. As explained in Sect. 1.6 and 1.7, a divergent central mass density can prevent the occurrence of central images and of the radial caustic. Moreover, since the best-fit model fBH10 contains black holes ten times more massive than those of fBH, it is more likely that the former model shows a lower number of radial caustics.

Regarding the displacement between the caustics of the three best-fit models, we

---

[2]Every cluster member is denoted with a number in our catalog.

studied further its implications as follows. We generated a series of sources (sources SG, see Sect. 4.2) close to the caustics produced by the cluster galaxies. Then, we computed the corresponding images predicted by the three best-fit models fB19, fBH and fBH10. Here, we present the results for the sources injected close to the caustics of the galaxy 7484. In this exemplary case, we strategically located the sources as shown in Fig. 5.1:

1. Source 1 (the blue source in Fig. 5.1): it is located inside the tangential caustics of fB19 and fBH, but it is outside the caustic of fBH10.

2. Source 2 (the yellow source in Fig. 5.1): it is inside the radial caustic of fB19 and outside its tangential caustic, while it is outside both the caustics of fBH and fBH10.

3. Source 3 (the pink source in Fig. 5.1): it is located outside the caustics of the models fBH and fBH10, but just inside the radial caustics of fB19.

For each source, we obtained its multiple image positions predicted by the fB19, fBH, and fBH10 (Fig. 5.2):

1. Source 1: as expected from the theory of strong gravitational lensing (see Chap. 1), we obtained five multiple images predicted by fB19 and fBH, and one image predicted by fBH10.

2. Source 2: we found one image produced by fBH, one image by fBH10, and two images by fB19. In this case, we expected three images from fB19, however, one image is probably highly de-magnified and, consequently, not found by Lenstool.

3. Source 3: we obtained one image for each model. In theory, the images of fB19 should be three. Instead, since the source is very close to the radial caustic, the other two images occur at the center of the galaxy and happen to be strongly de-magnified.
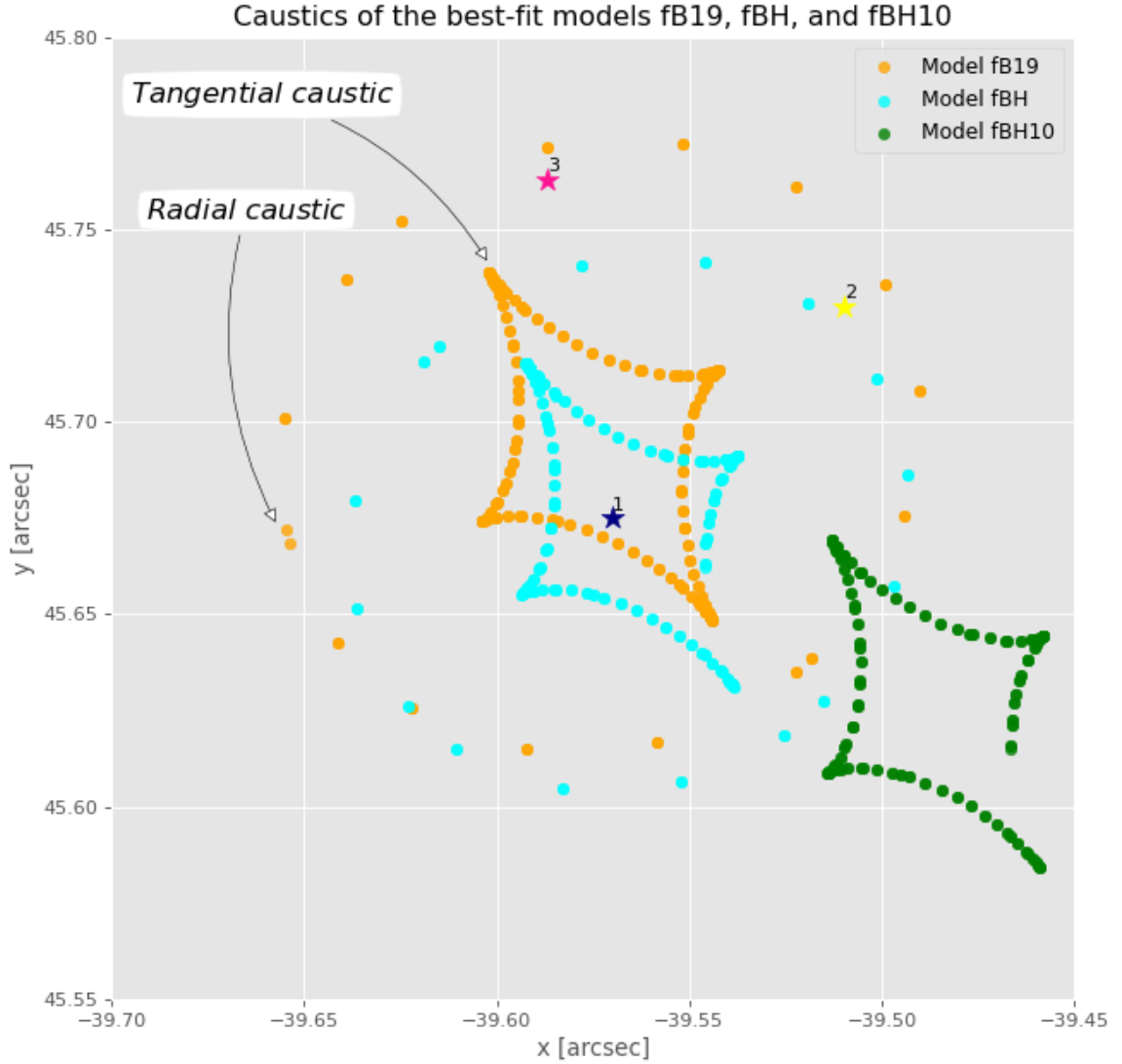
FIGURE 5.1. Caustics predicted by the three best-fit models fB19, fBH, and fBH10. These caustics are produced by the cluster galaxy 7484 (R.A.: 22:48:48.591, DEC.: −44:30:44.935). The orange, light-blue, and green dots correspond to the caustics produced by fB19, fBH, and fBH10, respectively. The stars represent the positions of the simulated sources. Both the caustics and the sources are at redshift $z = 0.73$. The origin of the reference frame corresponds to the center of the brightest cluster galaxy (R.A. = 22:48:43.970, DEC. = −44:31:51.132) of Abell S1063.

Our analysis demonstrates that the introduction of the supermassive black holes

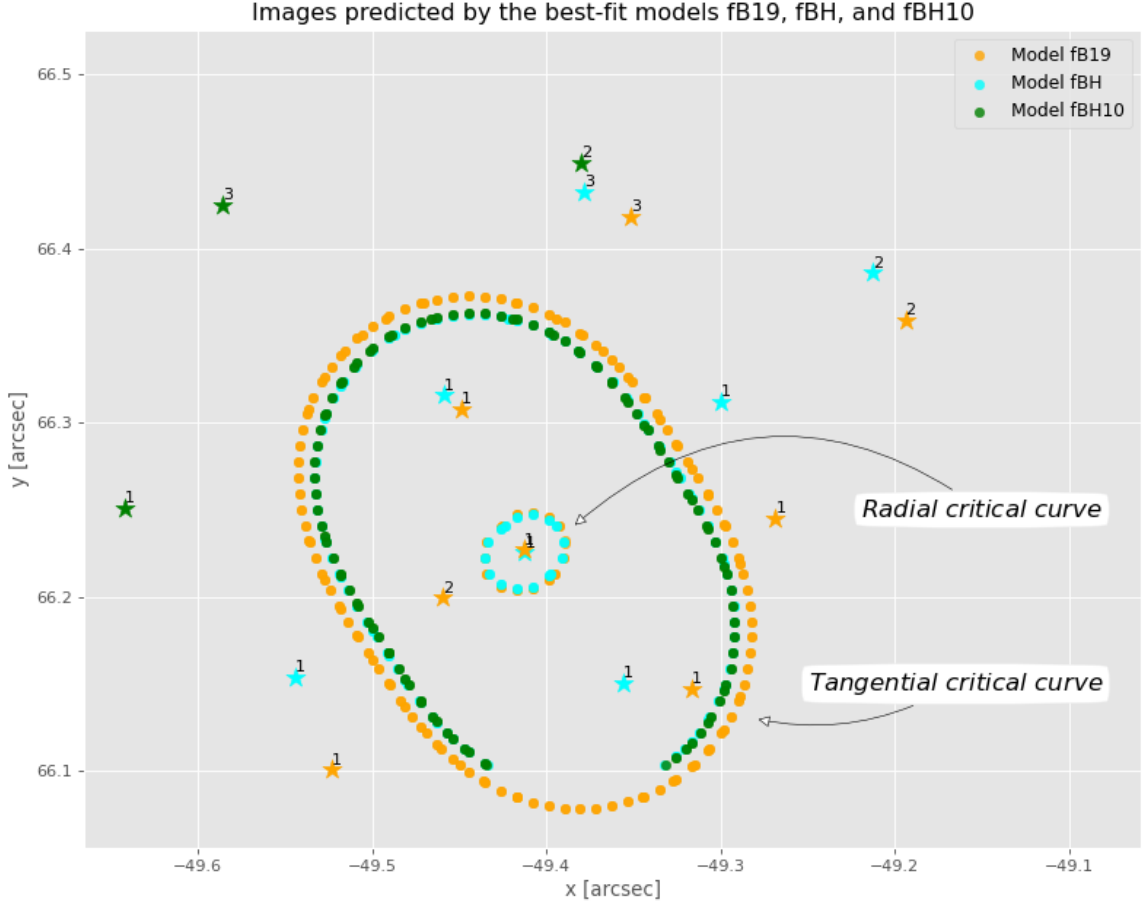at the center of the cluster galaxies in strong lensing models can modify the number

FIGURE 5.2. Images predicted by the three best-fit models fB19, fBH, and fBH10. The images predicted by the model fB19, fBH, and fBH10 are displayed as orange, light-blue stars, and green stars. The images with the label "i" originate from the source with the label "i" in Fig. 5.1; where $i = 1, 2, 3$. In addition, the critical curves produced by the three models are drawn, following the same colours of Fig. 5.1. Both critical curves and the images are displayed in the lens plane at redshift $z = 0.3480$.

of model-predicted multiple images from the sources that lie in the proximity of the caustics produced by the cluster galaxies.

5.2 Second methodology

In this section, we will present and discuss the results of the analyses described in Sect. 4.3.

As explained in Sect. 4.3, we computed the predicted images of the sources S1 (see Sect. 4.3) for the lens models fB19, fB19+bh, and fB19+bh10. Then, we calculated the differences between the image positions predicted by fB19+bh and those predicted by fB19 (Fig. 5.3) and the differences between the predicted image positions of fB19+bh10 and those of fB19 (Fig. 5.4). In the former case, the root-mean-square displacement $\Delta_{rms}^{\mathrm{bh,S1}}$ between the image positions predicted by the two models fB19+bh and fB19 is: $\Delta_{rms}^{\mathrm{bh,S1}} = 0.0028''$. While, in the latter case, the root-mean-square displacement $\Delta_{rms}^{\mathrm{bh10,S1}}$ between the image positions predicted by the two models fB19+bh10 and fB19 is: $\Delta_{rms}^{\mathrm{bh10,S1}} = 0.028''$.

From these results, we notice that the lens model fB19+bh (fB19+bh10), which contains the masses of the supermassive black holes at the centers of cluster galaxies, modifies by $\sim 10^{-3}$ arcsec ($\sim 10^{-2}$ arcsec) the predicted positions of the images with respect to the model fB19, which ignores the presence of the black holes. Accordingly, since the typical Δ_{rms} of a strong lensing model of a galaxy cluster is of the order of $\Delta_{rms} = 0.3'' - 0.5''$, the incorporation of the supermassive black holes in a strong lensing model is unlikely to alter significantly the value of Δ_{rms}.

Furthermore, we observe that the lens model fB19+bh10, which contains supermassive black holes ten times more massive than those of the lens model fB19+bh, produces a $\Delta_{rms}^{\mathrm{bh10,S1}}$ ten times larger than the $\Delta_{rms}^{\mathrm{bh,S1}}$ of fB19+bh. Thus, as expected, the effect of the population of black holes of the model fB19+bh10 on the position of the images is stronger than the effect of the less massive black holes of the model fB19+bh.

After presenting the impact of the supermassive black hole masses on the positions of the whole sample of multiple images, we now focus our attention on the images produced by the sources SG (the sources near the caustics of the cluster galaxies). Similarly to the previous case, we computed the displacement between the image positions predicted by fB19+bh and those predicted by fB19 (Fig. 5.5) and the displacement between the predicted image positions of fB19+bh10 and those of fB19
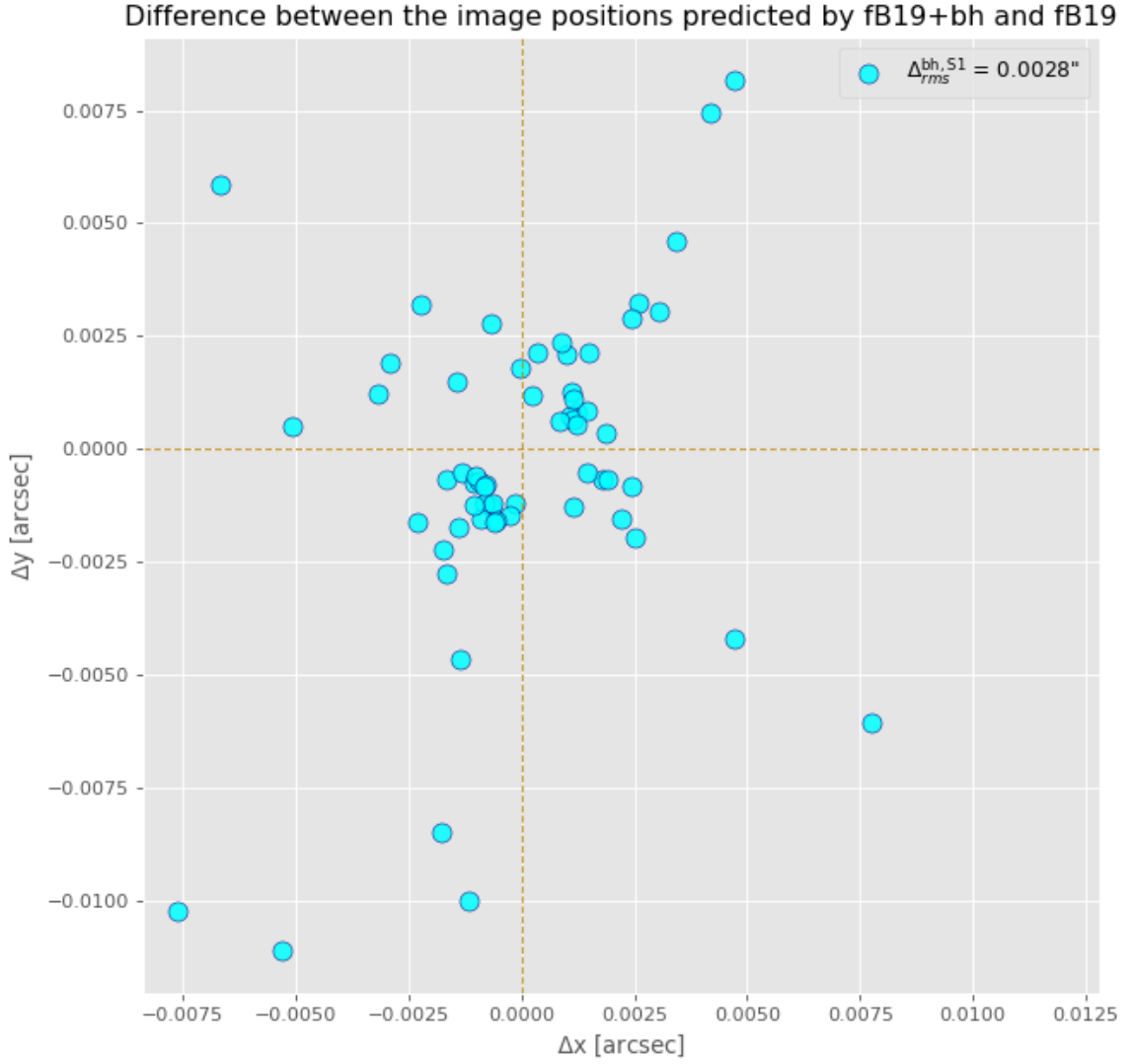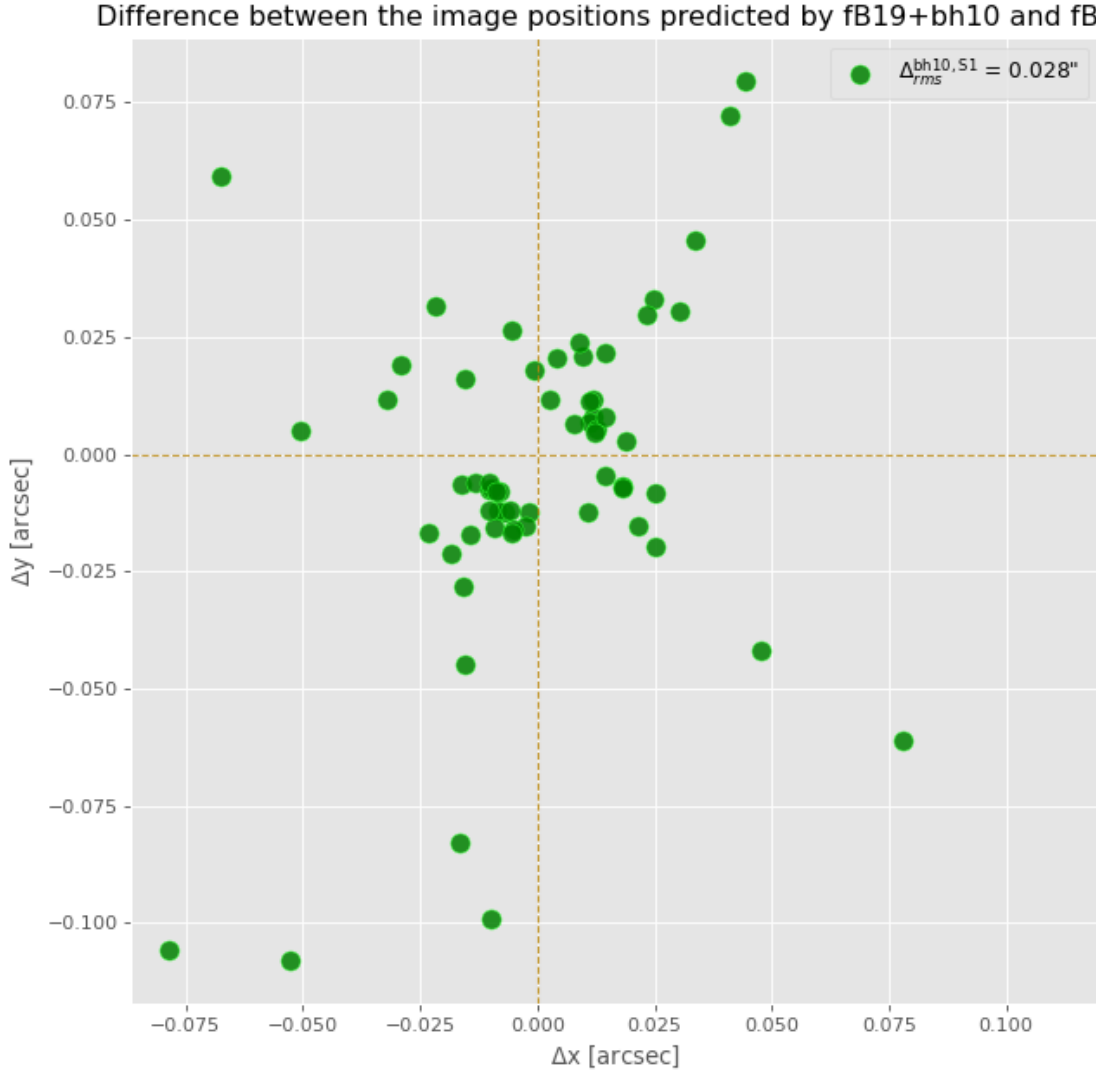
FIGURE 5.3. Difference between the image positions predicted by fB19+bh and those predicted by fB19. The images originate from the sources S1 (see Sect. 4.3). Each point marks the displacement (along the horizontal and vertical directions) between the position of the corresponding multiple image predicted by the two models. The total root-mean-square displacement is $\Delta_{rms}^{\mathrm{bh,S1}} = 0.0028''$.

(Fig. 5.6). In the former case, the $\Delta_{rms}^{\mathrm{bh,SG}}$ value corresponding to the image positions predicted by the two models fB19+bh and fB19 is: $\Delta_{rms}^{\mathrm{bh,SG}} = 0.0040''$. While, in the latter case, the $\Delta_{rms}^{\mathrm{bh10,SG}}$ value corresponding to the image positions predicted by the two models fB19+bh10 and fB19 is: $\Delta_{rms}^{\mathrm{bh10,SG}} = 0.038''$.

FIGURE 5.4. Difference between the image positions predicted by fB19+bh10 and those predicted by fB19. The images originate from the sources S1 (see Sect. 4.3). Each point marks the displacement (along the horizontal and vertical directions) between the position of the corresponding multiple image predicted by the two models. The total root-mean-square displacement is $\Delta_{rms}^{bh10,S1} = 0.028''$.

The previous observations regarding the case of the predicted images associated with the sources S1, still hold for the predicted images of the sources SG.

However, we note that the positions of the predicted images of the sources SG are more sensitive to the presence of the supermassive black holes compared to the
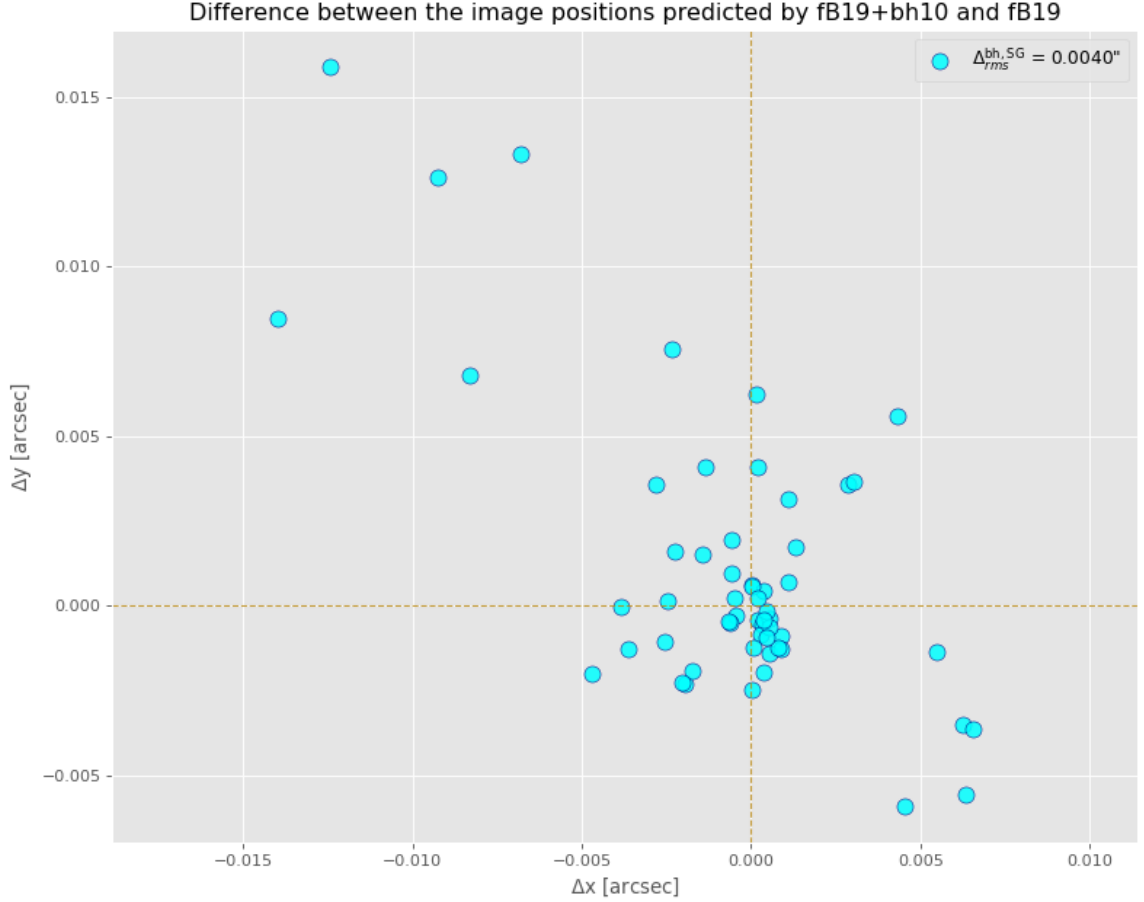
FIGURE 5.5. Difference between the image positions predicted by fB19+bh and those predicted by fB19. The images originate from the sources SG (see Sect. 4.2). Each point marks the displacement (along the horizontal and vertical directions) between the position of the corresponding multiple image predicted by the two models. The total root-mean-square displacement is $\Delta_{rms}^{\mathrm{bh,SG}} = 0.0040''$.

positions of the predicted images of the sources S1. In fact, considering the sources SG, we obtain values of $\Delta_{rms}^{\mathrm{bh,SG}}$ $(0.0040'')$ and $\Delta_{rms}^{\mathrm{bh10,SG}}$ $(0.038'')$ around 1.4 times larger than the values of $\Delta_{rms}^{\mathrm{bh,S1}}$ $(0.0028'')$ and $\Delta_{rms}^{\mathrm{bh10,S1}}$ $(0.028'')$ for the sources S1. This suggests that, if the sources are close to the caustics produced by the cluster galaxies, the positions of their images are more influenced by the gravitational field of the supermassive black holes at the center of the cluster galaxies, compared to when the sources are far from these caustics.
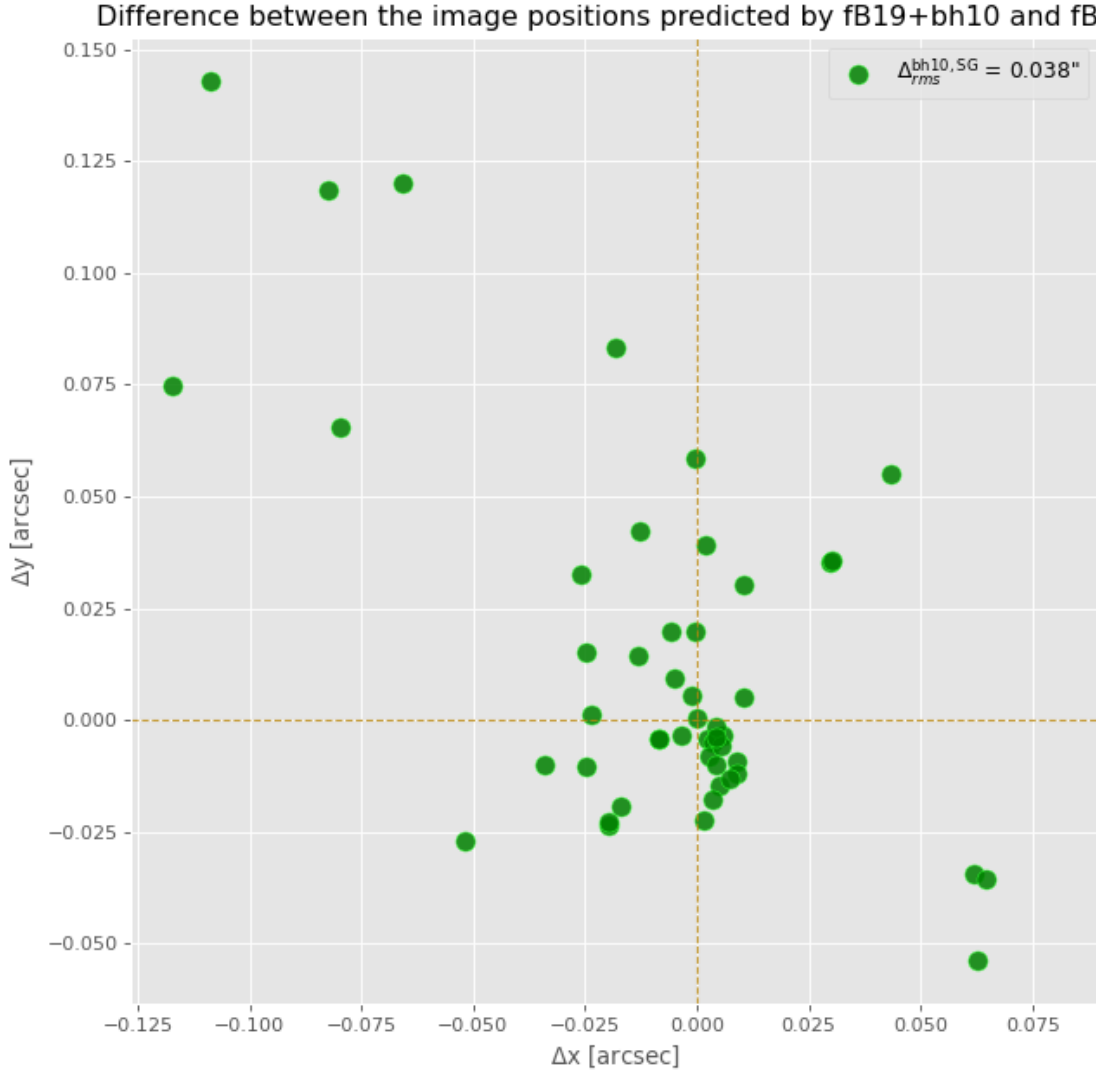
FIGURE 5.6. Difference between the image positions predicted by fB19+bh10 and those predicted by fB19. The images originate from the sources SG (see Sect. 4.2). Each point marks the displacement (along the horizontal and vertical directions) between the position of the corresponding multiple image predicted by the two models. The total root-mean-square displacement is $\Delta_{rms}^{\mathrm{bh10,SG}} = 0.038''$.

5.3   Third methodology

As explained in Sect. 4.4, we performed two optimizations of the strong lensing model B19: the first one on the images of the sources S1 (see Sect. 4.3) predicted by

the lens model fB19+bh and the second one on the images predicted by fB19+bh10. From the former optimization, we obtained the best-fit model fB19(bh) and, from the latter, the best-fit model fB19(bh10). The χ^2 values of the two best-fit models (with $dof = 73$) are: $\chi^2_{\text{fB19(bh)}} = 0.75$ and $\chi^2_{\text{fB19(bh10)}} = 4.14$.

We observe that the chi-squared value $\chi^2_{\text{fB19(bh10)}}$ of the best-fit model fB19(bh10) is approximately 5.5 times larger than the chi-squared $\chi^2_{\text{fB19(bh)}}$ of fB19(bh). Thus, the best-fit model fB19(bh) reproduces the image positions predicted by the model fB19+bh better than how the best-fit model fB19(bh10) reconstructs the image positions predicted by fB19+bh10. This is not surprising since the model B19, which does not take into account the supermassive black holes, describes a mass distribution of the galaxy cluster more similar to the mass distribution of fB19+bh than that of fB19+bh10. In fact, B19 and fB19+bh10 differ more significantly than B19 and fB19+bh due to the presence of the ten times more massive black holes in the fB19+bh10 model.

Furthermore, the root-mean-square displacement (rounded up to the second decimal digit), Δ_{rms}, between the image positions predicted by the best-fit model fB19(bh) (fB19(bh10)) and the image positions produced by the model fB19+bh (fB19+bh10) is $\Delta_{rms}^{\text{fB19(bh)}} = 0.00''$ ($\Delta_{rms}^{\text{fB19(bh10)}} = 0.00''$). Therefore, the positions of the images predicted by the model fB19+bh (and fB19+bh10) can be well reconstructed by the best-fit model fB19(bh) (fB19(bh10)).

As discussed in Sect. 4.4, during the optimization process, Lenstool samples the posterior probability distribution of the model parameters. In Fig. 5.7, 5.8, and 5.9, we show the marginalized posterior distributions for the parameters that describe the mass distribution of the cluster galaxies of the models fB19(bh) and fB19(bh10). In particular, in Fig. 5.7, we display the marginalized posterior probability distribution of the cut radius, r_t^{ref}, of the reference galaxy (see Sect. 3.2); in Fig. 5.8, we show the marginalized posterior probability distribution of the central velocity dispersion, σ_0^{ref}, of the reference cluster galaxy; in Fig. 5.9, we display the probability distribution

marginalized over all the model parameters except r_t^{ref} and σ_0^{ref}. In these plots, we display, using vertical dashed lines, the values of the parameters r_t^{ref} and σ_0^{ref} of the lens model fB19+bh (or fB19+bh10). We recall that the models fB19+bh and fB19+bh10 (see Sect. 4.3) are completely determined by fixed parameters and differ only by the mass values of the black holes. Therefore, the parameters, r_t^{ref} and σ_0^{ref}, of the sub-halos are fixed and identical in the two models. Finally, by combining the marginalized posterior probability distributions of r_t^{ref} and σ_0^{ref} (see Eq. (1.8.5)), we also computed the posterior probability distribution of the total mass M_{tot}^{ref} (see Sect. 3.3) of the reference galaxy (Fig. 5.10).

Examining these results, we notice that the galactic sub-halos obtained by the model fB19(bh10) are more compact than the sub-halos of fB19(bh). In fact, the former model has a smaller value of the cut radius r_t^{ref} of the reference galaxy and a larger value of the reference velocity dispersion σ_0^{ref}. Also, the sub-halos in the lens models fB19(bh) and fB19(bh10) are more compact than those in fB19+bh and fB19+bh10.

In order to explain this result, we highlight some characteristics of the lens models used in this section. The fixed models fB19+bh and fB19+bh10 are simulated[3] galaxy clusters. The former describes each cluster galaxy with a sub-halo and a central black hole, whose mass is estimated using the $M_\bullet - \sigma$ scaling relation (see Sect. 4.1). The latter models each galaxy with the same sub-halo and a central black hole, whose mass is ten times larger than that of the former model. When we optimize the strong lensing model B19, which ignores the presence of supermassive black holes, assuming the image positions predicted by fB19+bh (fB19+bh10) as observables, the sub-halos of B19 have to compensate the high central mass density of the galaxies (due to the presence of the central black holes) of the model fB19+bh (fB19+bh10). Therefore, in this context, the sub-halos of the best-fit models of B19 (fB19(bh) and fB19(bh10)) appear more compact than the sub-halos of the models fB19+bh and fB19+bh10.

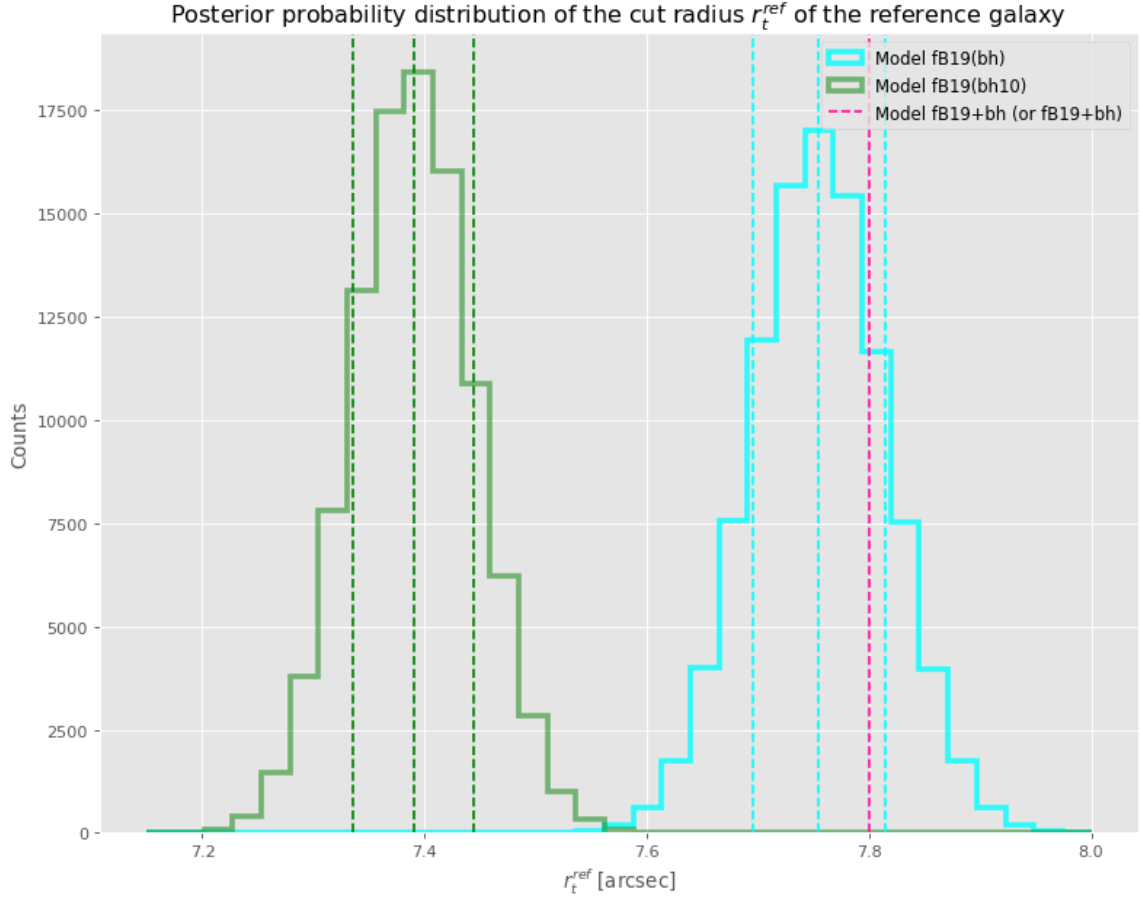[3]The models fB19+bh and fB19+bh10 are not realistic models of Abell S1063 (see Sect. 4.3).

FIGURE 5.7. Marginalized posterior probability distributions of the cut radius, r_t^{ref}, of the reference galaxy of the models fB19(bh) and fB19(bh10). The marginalized posterior probability distributions of fB19(bh) and fB19(bh10) are plotted in light-blue and green, respectively. The light-blue and green vertical lines show the 16th, 50th, and 84th percentiles of the two distributions. The pink vertical line corresponds to the cut radius, r_t^{ref}, of the reference galaxy of the model fB19+bh (or fB19+bh10). In the model fB19(bh), the cut radius $r_t^{ref,b} = 7.75_{-0.06}^{+0.06}$ arcsec; in the model fB19(bh10), $r_t^{ref,b10} = 7.39_{-0.05}^{+0.05}$ arcsec; in the model fB19+bh (or fB19+bh10), $r_t^{ref} = 7.80$ arcsec.

For the same reason, in the best-fit model fB19(bh10), they are even more compact compared to those of the best-fit model fB19(bh).

Turning to Fig. 5.10, we can observe that the total masses of the sub-halos are smaller in the lens models where the sub-halos are more compact. This result is unexpected since the most compact sub-halos reproduce galaxies containing the most
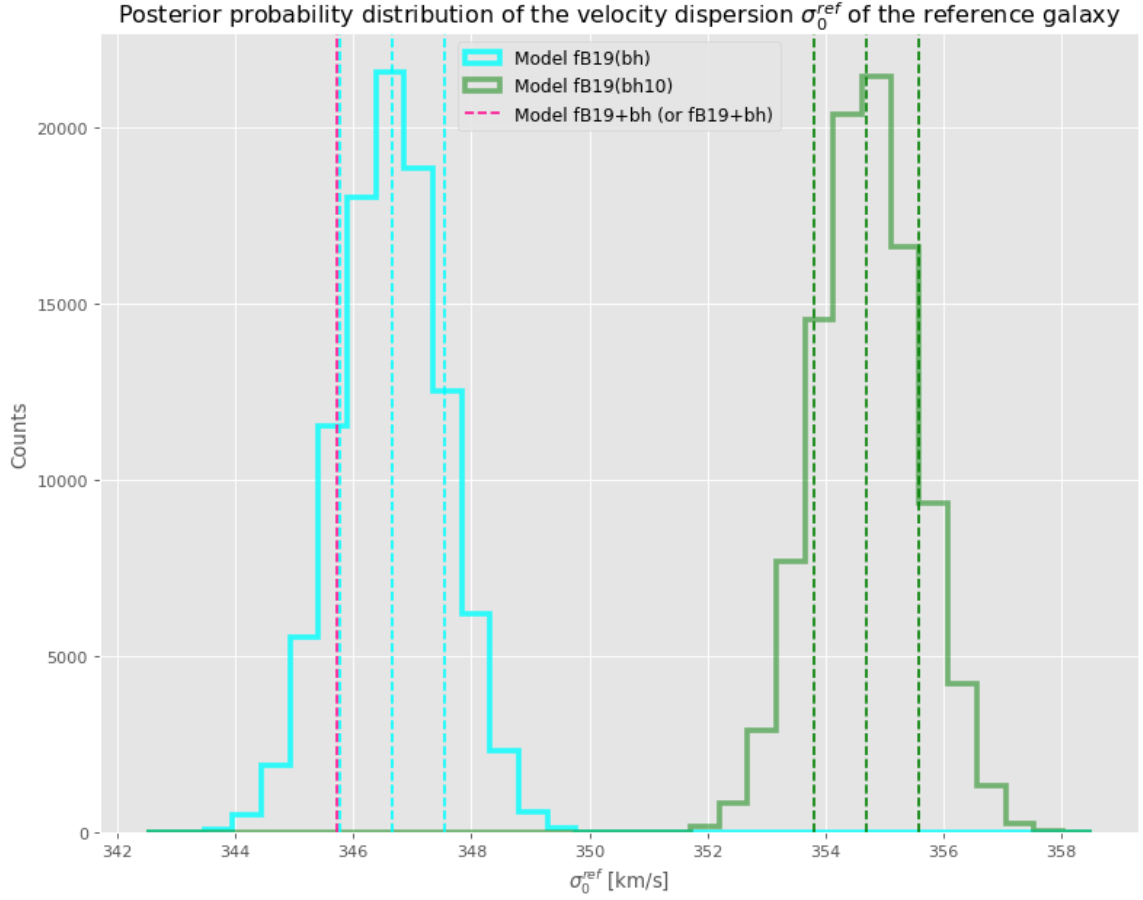
FIGURE 5.8. Marginalized posterior probability distribution of the velocity dispersion, σ_0^{ref}, of the reference galaxy of the models fB19(bh) and fB19(bh10). The marginalized posterior probability distributions of fB19(bh) and fB19(bh10) are plotted in light-blue and green, respectively. The light-blue and green vertical lines show the 16th, 50th, and 84th percentiles of the two distributions. The pink vertical line corresponds to the cut radius, r_t^{ref}, of the reference galaxy of the model fB19+bh (or fB19+bh10). In the model fB19(bh), $\sigma_0^{ref,b} = 346.66_{-0.89}^{+0.89}$ km/s; in the model fB19(bh10), $\sigma_0^{ref,b10} = 354.70_{-0.88}^{+0.87}$ km/s; in the model fB19+bh (or fB19+bh10), $\sigma_0^{ref} = 345.73$ km/s.

massive black holes. However, this is likely to be caused by an "exchange"[4] of mass between the sub-halos and the extended halos in the strong lensing model B19.

---

[4]A transfer of mass between the small-scale sub-halos and the extended halos always affects a strong lensing model of galaxy clusters.
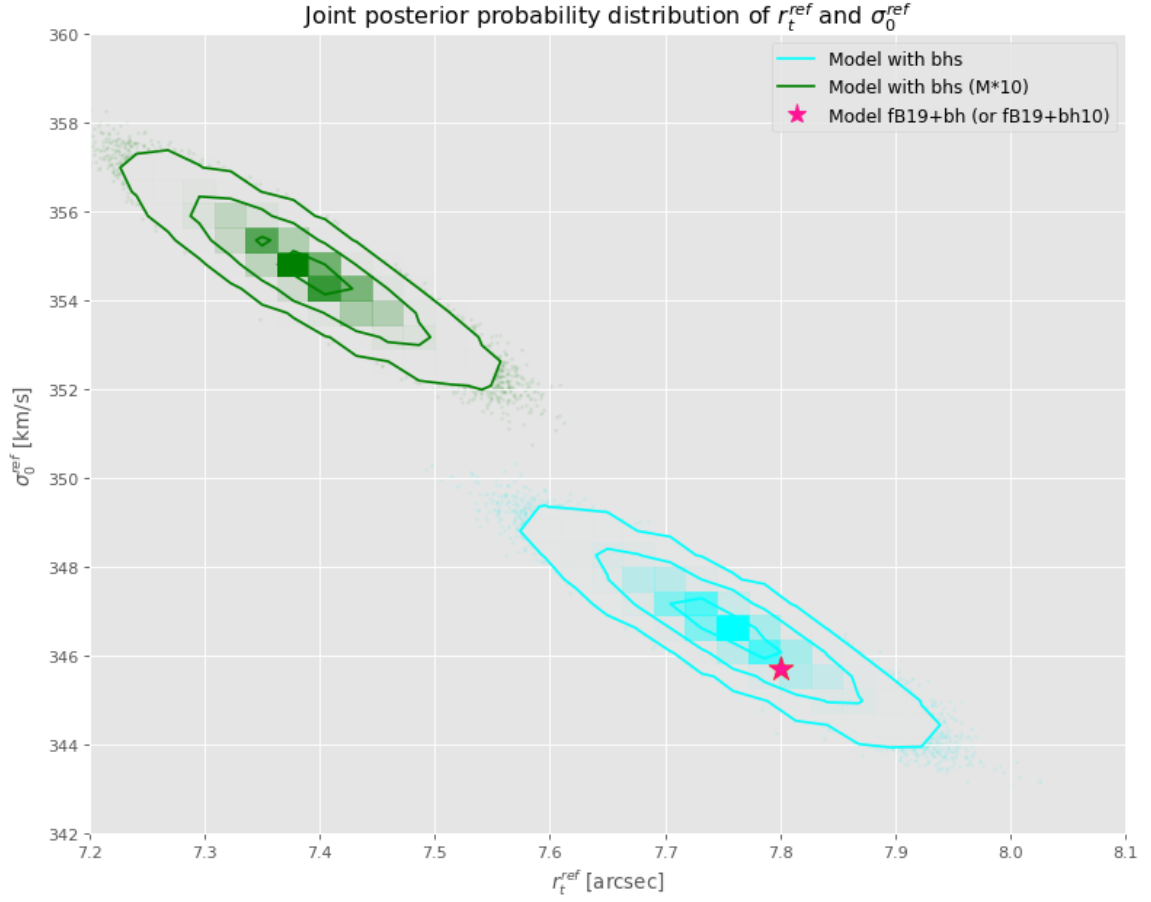
FIGURE 5.9. Joint posterior probability distributions of the parameters r_t^{ref} and σ_0^{ref}. The joint posterior probability distributions of fB19(bh) and fB19(bh10) are plotted in light-blue and green, respectively. The light-blue and green contours represent the 68%, 95.4% and 99.7% confidence levels of the two two-dimensional probability distributions. The pink star corresponds to values of the parameters σ_0^{ref} and r_t^{ref} of the reference galaxy in the model fB19+bh (or fB19+bh10).
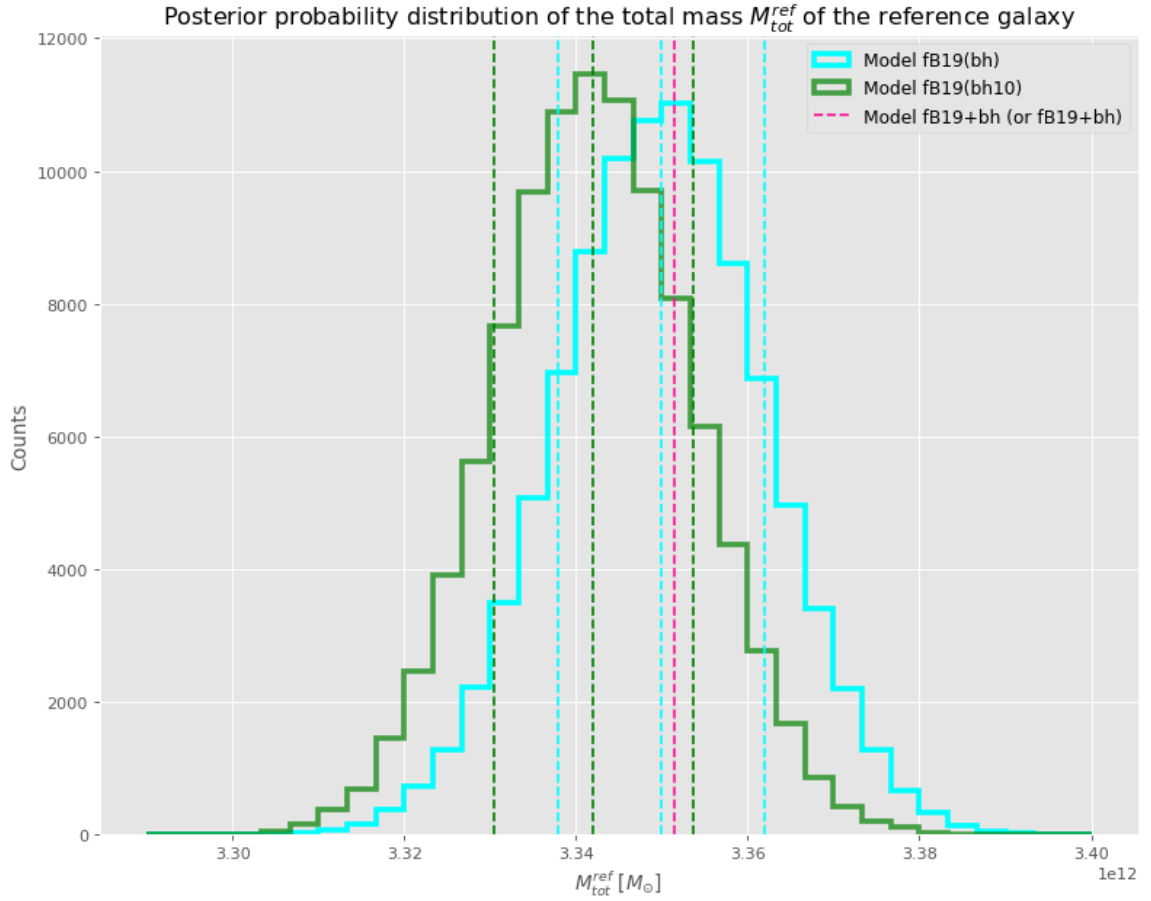
FIGURE 5.10. Posterior probability distributions of the total mass, M_{tot}^{ref}, of the reference galaxy of the models fB19(bh) and fB19(bh10). The probability distributions of fB19(bh) and fB19(bh10) are plotted in light-blue and green, respectively. The light-blue and green vertical lines show the 16th, 50th and 84th percentiles of the two distributions. The pink vertical line corresponds to the total mass M_{tot}^{ref} of the reference galaxy in the model fB19+bh (or fB19+bh10). In the model fB19(bh), $M_{tot}^{ref,b} = 3.35_{-1.20}^{+1.20} \times 10^{12} M_\odot$; in the model fB19(bh10), $M_{tot}^{ref,b10} = 3.34_{-1.15}^{+1.17} \times 10^{12} M_\odot$; in the model fB19+bh (or fB19+bh10), $M_{tot}^{ref} = 3.35 \times 10^{12} M_\odot$.

Chapter 6

CONCLUSIONS AND FUTURE PERSPECTIVES

In the strong gravitational lensing model of a galaxy cluster, the total mass distribution of the cluster normally includes three main mass components: extended dark-matter halos, extended hot-gas halos, and sub-halos corresponding to the cluster galaxies. In this thesis, we have also considered the supermassive black holes located at the centers of the cluster galaxies. As a representative example, we have used the galaxy cluster RXC J2248.7$-$4431 (also known as Abell S1063), at a redshift of $z = 0.348$, studied by Bergamini et al. (2019).

We have used the public software Lenstool (Kneib et al. 1996, Jullo et al. 2007) to quantify the differences between mass models that include or ignore the presence of supermassive black holes. In particular, we have focused on the positions of the model-predicted multiple images and on the values, estimated through a strong lensing analysis, of the parameters describing the total mass distribution of the lens.

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

1. We have reconstructed the observed image positions of 55 multiple images from 20 sources using three strong lensing models of Abell S1063: a model (B19) that ignores the presence of supermassive black holes, another model (BH) that takes into account the supermassive black holes, whose masses have been estimated using the $M_\bullet - \sigma$ relation, and a third model (BH10) with black holes ten times more massive than those of BH. We have obtained that the Δ_{rms} value between the observed and model-predicted images is equal to $0.55''$ for all three best-fit models. Accordingly, we have concluded that the effect on the overall model precision of the presence of supermassive black holes in a strong lensing model is negligible compared to other simplifying assumptions of the model, such as

neglected dark-matter clumps and unconsidered mass structures along the line of sight.

2. By computing the caustics (and the corresponding critical curves) of the afore-mentioned best-fit models, we have found that the presence of supermassive black holes in the strong lensing model of Abell S1063 can cause the disappearance of the radial caustics (and the radial critical curves) produced by the cluster galaxies, and it can modify the number of model-predicted multiple images from specific positions of the sources located near the caustics produced by the cluster galaxies.

3. Given a set of sources (S1), we have computed their multiple image positions produced by three fixed mass models. The first model (fB19) describes a simulated galaxy cluster without supermassive black holes at the centers of the cluster galaxies; the second mass model (fB19+bh) adds the masses of the supermassive black holes (estimated using the $M_\bullet - \sigma$ relation) at the centers of the cluster galaxies of fB19; the third one (fB19+bh10) contains black holes ten times more massive than those of fB19+bh. By examining the root-mean-square displacement, Δ_{rms}, between the image positions predicted by the two models fB19+bh and fB19 and that between the image positions predicted by the models fB19+bh10 and fB19, we have concluded that the incorporation of the supermassive black holes in a strong lensing model of galaxy clusters cannot alter significantly the value of Δ_{rms}.

4. By following the steps of the previous point for the sources SG (the sources close to the caustics produced by cluster galaxies), we have found that: if the sources are located in the proximity of the caustics generated by the cluster galaxies, the positions of their images are more influenced by the gravitational field of the supermassive black holes at the centers of the cluster galaxies, compared to

when the sources are farther away from these caustics.

5. We have optimized the strong lensing model B19, which ignores the presence of supermassive black holes, assuming the image positions predicted by fB19+bh (and fB19+bh10) as observables. For clarity, we refer to the two best-fit models of B19 as fB19(bh) and fB19(bh10), respectively. We have found that the sub-halos of these best-fit models, in order to replicate the central mass concentration of the galaxies (due to the presence of the central black holes) of fB19+bh and fB19+bh10, appear more compact (higher values of velocity dispersion and smaller size) compared to the galaxy halos of fB19+bh and fB19+bh10. For the same reason, in the best-fit model fB19(bh10), they are even more compact compared to those of the best-fit model fB19(bh).

6. We have observed that the total masses of the sub-halos are smaller in the lens models where the sub-halos are more compact. This result is unexpected since the most compact sub-halos reproduce galaxies containing the most massive black holes. However, this is likely to be caused by an "exchange" of mass between the sub-halos and the extended halos in the strong lensing model B19.

We defer the exploration of other possible analyses regarding the influence of supermassive black holes in strong gravitational lensing models of galaxy clusters to future studies. In particular, a recent work (Ricarte et al. 2021) has suggested the existence of a population of supermassive black holes (referred to as *wandering*) in the galaxy clusters, located outside the cluster galaxies. Therefore, it could be worthwhile to extend the analyses conducted in this thesis by considering also this type of supermassive black holes. Finally, thanks to new James Webb Space Telescope observations, Pacucci et al. (2023) have suggested that distant, young galaxies might harbor supermassive black holes violating the $M_\bullet - \sigma$ relation used in this thesis. These black holes could be up to $100 - 1000$ times more massive than thought so far.

Thus, it could be interesting to investigate the impact of these black holes in strong lensing models of galaxy clusters.

Bibliography

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

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

Bender, R., Burstein, D., & Faber, S. M. 1992, ApJ, 399, 462

Bergamini, P., Rosati, P., Mercurio, A., et al. 2019, A&A, 631, A130

Bonamigo, M., Grillo, C., Ettori, S., et al. 2018, ApJ, 864, 98

Bonamigo, M., et al 2017, ApJ, 842, 132

Bovi, J., 2017, lectures notes of "Galactic Structure and Dynamics"

Caminha, G. B., Karman, W., Rosati, P., et al. 2015, A&A, submitted [arXiv:1512.05655]

Caminha, G. B., Grillo, C., Rosati, P., et al. 2016, A&A, 587, A80

Cappellari, M. & Emsellem, E. 2004, Publications of the Astronomical Society of the Paci c, 116, 138

Congdon, A. B., Keeton, C. R. 2018, Springer Cham

Djorgovski, S. & Davis, M. 1987, ApJ, 313, 59

Dressler, A., Lynden-Bell, D., Burstein, D., et al. 1987, ApJ, 313, 42

Faber, S. M. & Jackson, R. E. 1976, ApJ, 204, 668

Ferrarese L., Merritt D., 2000, ApJ, 539, L9

Frittelli, S., Kling, T. P., and Newman, E. T. 2000 Phys. Rev. D 63, 023007

Gebhardt K., et al., 2000, ApJ, 539, L13

Grillo, C., Suyu, S. H., Rosati, P., et al. 2015, ApJ, 800, 38

Grillo, C., Wong, K. C., Suyu, S. H., Chen, G. C.-F., et al. 2020, ApJ, 898, 87

Hubble, E. P. 1926, ApJ, 64

Jullo, E., Kneib, J.-P., Limousin, M., et al. 2007, New J. Phys., 9, 447

Kneib, J.-P., Ellis, R. S., Smail, I., Couch, W. J., & Sharples, R. M. 1996, ApJ, 471, 643

Laporte, C. F. P., & White, S. D. M. 2015, MNRAS, 451, 1177 Le Fèvre, O., Saisse, M., Mancini, D., et al. 2003, in Instrument Design and Performance for Optical/Infrared Ground-based Telescopes, eds. M. Iye, & A. F. M. Moorwood, SPIE Conf. Ser., 4841, 1670

Limousin, M., Kneib, J.-P., & Natarajan, P. 2005, MNRAS, 356, 309

Lotz, J. M., Koekemoer, A., Coe, D., et al. 2017, ApJ, 837, 97

Magorrian J., et al., 1998, AJ, 115, 2285

McConnell, N. J., and Ma, C.-P. 2013, ApJ, 764, 184

Meštrić, U., Vanzella, E., Zanella, A., et al. 2022, Monthly Notices of the Royal Astronomical Society, 516, 3

Monna A., Seitz S., Greisel N. et al. 2014 MNRAS 438 1417

Ohanian H. C., 1983, ApJ, 271, 551

Pacucci, F., et al 2023 ApJL 957 L3

Peterson B. M., Ferrarese L., Gilbert K. M., Kaspi S., Malkan M. A., Maoz D., Merritt D., et al., 2004, ApJ, 613, 682

Postman, M., Coe, D., Bentez, N., et al. 2012, ApJS, 199, 25

Ricarte A., Tremmel M., Natarajan P., Zimmer C., Quinn T., 2021a, MNRAS, 503, 6098

Roberts-Borsani, G., Treu, T., Chen, W. et al., 2023, Nature, 618, 480

Rood, H. J., & Sastry, G. N. 1971, Astronomical Society of the Pacific, 83(493), 313–319

Rosati, P., Balestra, I., Grillo, C., et al. 2014, Msngr, 158, 48

Sargent, W. L. W., Turner, E. L. 2017, ApJ 212

Schneider P., Ehlers J., Falco E. E., 1992, Gravitational Lenses, doi:10.1007/978-3-662-03758-4

Schneider, P. 2006, Extragalactic Astronomy and Cosmology (Springer Nature)

Seyfert, C. K. 1943, ApL, 97, 28

Skilling, J., https://www.inference.org.uk/bayesys/, 2004

Suyu, S.H., Chang, TC., Courbin, F. et al. 2018, Space Sci Rev, 214, 91

Umetsu, K., Medezinski, E., Nonino, M., et al. 2014, ApJ, 795, 163

Van Dokkum, P., Brammer, G., Wang, B. et al. 2024, Nat Astron 8, 119–125

Vanzella, E., Guo, Y., Giavalisco, M., et al. 2012, ApJ, 751, 70

Zwicky, F. 1937, ApJ, 86, 217