Dissecting the Strong-lensing Galaxy Cluster MS 0440.5+0204. I. The Mass Density Profile

Tomás Verdugo1, Eleazar R. Carrasco2, Gael Foëx3, Verónica Motta3, Percy L. Gomez4, Marceau Limousin5, Juan Magaña6, and José A. de Diego7,8

1 Observatorio Astronómico Nacional, Instituto de Astronomía, Universidad Nacional Autónoma de México, Ensenada, B.C., México; tomasv@astro.unam.mx
2 Gemini Observatory/AURA, Southern Operations Center, Casilla 603, La Serena, Chile
3 Instituto de Física y Astronomía, Facultad de Ciencias, Universidad de Valparaíso, Avda. Gran Bretaña 1111, Valparaíso, Chile
4 W. M. Keck Observatory, 65-1120 Mamalahoa Highway, Kamuela, HI, USA 96743
5 CNRS, CNES, LAM, Marseille, France
6 Instituto de Astrofísica, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna, 4860, Santiago, Chile
7 Instituto de Astronomía, Universidad Nacional Autónoma de México, Avenida Universidad 3000, Ciudad Universitaria, C.P. 04510, Ciudad de México, México
8 Instituto de Astrofísica de Canarias (IAC), E-38200 La Laguna Tenerife, Spain

Received 2019 July 3; revised 2020 May 15; accepted 2020 May 23; published 2020 June 26

Abstract

We present a parametric strong-lensing modeling of the galaxy cluster MS 0440.5+0204 (located at \(z = 0.19\)). We have performed a strong-lensing mass reconstruction of the cluster using three different models. The first model uses the image positions of four multiply imaged systems (providing 26 constraints). The second one combines strong-lensing constraints with dynamical information (velocity dispersion) of the cluster. The third one uses the mass calculated from weak lensing as an additional constraint. Our three models reproduce equally well the image positions of the arcs, with an rms image equal to \(\approx 0.05\). However, in the third model, the inclusion of the velocity dispersion and the weak-lensing mass allows us to obtain better constraints in the scale radius and the line-of-sight velocity dispersion of the mass profile. For this model, we obtain \(r_s = 132^{+8}_{-13}\) kpc, \(\sigma_z = 1203^{+78}_{-272}\) km s\(^{-1}\), \(M_{200} = 3.1^{+0.6}_{-0.2} \times 10^{14} M_\odot\), and a high concentration \(c_{200} = 9.9^{+2.2}_{-1.4}\). Finally, we used our derived mass profile to calculate the mass up to 1.5 Mpc. We compare it with X-ray estimates previously reported, finding a good agreement.

Unified Astronomy Thesaurus concepts: Gravitational lensing (670); Galaxy clusters (584); Dark matter (353)

1. Introduction

Studying mass density profiles in galaxy clusters offers the opportunity to probe the formation and evolution of structures in the universe and to probe different cosmological models (see Kravtsov & Borgani 2012). In particular, three methods commonly used to study the mass in galaxy clusters are (1) strong and weak gravitational lensing (e.g., Kneib & Natarajan 2011; Hoekstra et al. 2013 and references therein), (2) X-ray measurements (see Ettori et al. 2013 and references therein), and (3) dynamical analysis from the velocity dispersion of the cluster members (e.g., Girardi et al. 1998; Biviano et al. 2006; Old et al. 2015; Wojtak et al. 2018, and references therein). In fact, several authors have combined two or more of these probes to recover a more robust mass distribution in galaxy clusters (e.g., Kneib et al. 2003; Bradac et al. 2005; Newman et al. 2009; Verdugo et al. 2011, 2016; Morandi et al. 2012; Limousin et al. 2013; Siegel et al. 2018). Indeed, even in the Hubble Frontier Field era, the study of the mass distribution in clusters requires data to probe the gravitational field beyond the nuclei responsible for strong lensing (Limousin et al. 2016).

This is the first paper (Paper I) in a series of two which aim to present a comprehensive and combined analysis of MS 0440.5+0204 by performing lensing modeling (strong and weak), dynamical analysis (through spectroscopy of the galaxy members), and X-ray data. In E. R. Carrasco et al. (2020, in preparation, hereinafter Paper II), we will show the detailed dynamical analysis of the cluster based on the redshift of 93 confirmed member galaxies inside 0.4 \(\times 0.2\) deg\(^2\). In this first paper, we present new lensing models for MS 0440.5+0204, and compare the results with the X-ray data available and the mass previously reported by other authors (Gioia et al. 1998; Hicks et al. 2006; Shan et al. 2010).

This galaxy cluster was selected for its regular morphology with not obvious substructures and therefore easily modeled by a single dark matter halo. The first model MS 0440.5+0204 model was presented in Gioia et al. (1998), hereafter G98, when the only arc with measured redshift was A1 (see Section 2). Nevertheless, these authors were able to put a limit in the mass of the cluster, and they calculated a range of redshifts for the arcs. Using strong lensing, Wu (2000) estimated the mass within the arcs (\(9.0 \times 10^{13} M_\odot\)) and compared it with the mass calculated from the cluster X-ray luminosity, finding that the former was smaller by a factor of 2. This discrepancy was also reported by Shan et al. (2010), who analyzed a sample of 27 clusters including MS 0440.5+0204. This cluster was also explored by different authors using weak-lensing data (e.g., Hoekstra et al. 2012; Mahdavi et al. 2013). The study reported here is the first strong-lensing analysis of MS 0440.5+0204 since G98, using new spectroscopic redshifts for some of the multiple images as model constraints.

In Verdugo et al. (2011), we combined strong-lensing and dynamical information at large scale to estimate the scale radius of a Navarro–Frenk–White (NFW) mass profile (Navarro et al. 1996, 1997), using the constraints as a prior in the strong-lensing analysis. Here, we aim to extend the method of Verdugo et al. (2011) by fitting strong-lensing constraints with dynamical information, combining likelihoods, and adding the mass as an additional constraint. The paper is organized as follows: in the next section we describe the data used. We explain the methodology in Section 3. In Section 4, we present the different models studied in the paper, and we discuss the
Figure 1. HST WFPC2 F702W image, with the local median average of the central region ($60''\times60'')$ of the MS 0440 cluster subtracted. The center of the image coincides with the galaxy 896(B), located at $\alpha = 4:43:09:8, \delta = +2:10:18.2$. The 14 galaxies considered in our models are identified along with the arc systems; each color corresponds to a different arc system. We highlight with ellipses the radial arcs reported and discussed in G98.

results in Section 5. Finally, we present our conclusions in Section 6. Throughout this paper, we adopt a spatially flat cosmological model dominated by cold dark matter and a cosmological constant. We use $\Omega_m = 0.3$, $\Omega_{\Lambda} = 0.7$, and $H_0 = 70$ km s$^{-1}$ Mpc$^{-1}$. With these cosmological parameters, $1''$ corresponds to 3.24 $h^{-1}$ kpc at the redshift of MS 0440.5+0204.

2. Observations

2.1. HST Data

The Hubble Space Telescope (HST) data have been obtained from the Multimission Archive at Space Telescope. They consist of 10 WFPC2 images obtained in the F702W filter with a total integration time of 22,200 s. The image reduction was performed using the IRAF/STSDAS package. First, a warm-pixel rejection was applied to the images using the IRAF task warmpix. The cleaned images were then combined with the task crrej to remove cosmic-ray hits, and finally, the image was aligned with WCS coordinates, i.e., north up and east left. Following G98, we subtracted an image composed of the local median average of the multiple-nucleus cD light distribution. In Figure 1, we show the $60''\times60''$ field obtained after the subtraction, which is used in our strong-lensing analysis. It is important to note that the HST data used in the present paper are the same as those used by G98. Since the publication of that work, there have been no new optical HST observation of MS 0440.5+0204.

Multiple images. In Figure 1 (see also Figure A1 in Appendix A), we identify the six arc systems used in our models. These systems were first reported and used by G98. Systems consisting of a single arc are denoted by S, while those showing multiple images are named by M. We have measured the redshifts of the arcs S1, M1.1, M1.2, M1.3, M1.4, M2.1, M2.2, M3.1, and M3.2 (see Section 2.2 and Table 1). These measured redshifts confirm the proposed arc-system associations used by G98 (even though we used these same systems, we have written our own nomenclature to point out how the multiple images are related to each other). No spectroscopic features were found in the spectra of M3.3, M4.1, M4.2, M4.3, M5.1, and M5.2.

Arc S1 appears to be a highly distorted image of a galaxy (G98), with two substructures that are easily distinguished. However, in spite of its appearance, it is likely to be a single-image system because no counterimages were found. Furthermore, our models are consistent with S1 being a single image (see Section 4.3). The images M2.1 and M2.2, at redshift $z = 0.9543$, have multiple knots. The difference in the numbers of knots in each image is because M2.2 is slightly elongated in the radial direction, splitting the inner part of the object into two different images, which overlap in M2.1.

Observing the image-pattern distribution of system M1, we note that it is similar (with a flip in parity) to that displayed in system M3. Additionally, M1.1 and M1.2 show a mirror symmetry, similar to the one in arcs M3.1 and M3.2, which are resolved into bright knots, with a substructure that divided them in at least two parts. Thus, following G98, who assume that some arc systems belong to the same source, we associate our measured redshifts of arcs M3.1 and M3.2 to their respective counterimages, namely, M3.3, and M3.4. Note that system M3.4 it is not labeled in G98. This association is predicted by our models.

Near the cluster center, there are two radial structures: M4.2 and M5.2. These radial arcs were analyzed and discussed in G98, where the authors studied their geometry and calculated their magnitudes. The importance of these arcs lies in their effect of constraining the mass profile. For an axially symmetric lens, the Jacobian matrix of the lens mapping has two eigenvalues: one relates the tangential critical curve to the total enclosed mass, and the other links the position of the radial arcs to the derivative of the mass (Schneider et al. 1992). These radial systems are key ingredients for characterizing the profile in the inner part. Miralda-Escude (1995) showed that the combination of radial arcs and their counterimages provides powerful constraints on the dark matter density profile as well as on the enclosed mass. Because of our interest in the radial systems, we have included them in our model even though we do not have measured redshifts. Thus, we need to deduce whether M4 and M5 are images of one or two sources. Based on their structure and shape, we infer that these images come from two different sources (see Table 1). The same reasoning was used to assume that M4.3 belongs to system M4. Although we do not have measured redshifts for these systems, given their importance explained in the above lines, we keep them and set their redshifts as free parameters to be estimated during the modeling process.

2.2. GEMINI Data

The spectroscopic data used in this work will be described in Paper II. We refer the interested reader to our forthcoming
publication, Paper II, for a detailed analysis. Here we offer a brief summary. The images were obtained in 2011 (program ID: GS−2011B−Q−59) with the Gemini Multi-Object Spectrograph (GMOS; Hook et al. 2004) at the Gemini South telescope in Chile. MS 0440.5−0204 was imaged in the g′ (3 × 300 s exposures), r′ (3 × 200 s exposures), and i′ filters (3 × 200 s exposures). The images were observed during dark time and under photometric conditions, with median values of seeing of 0.51, 0.50, and 0.57 in the g′, r′, and i′ filters, respectively. Galaxies for spectroscopic follow-up were selected using their magnitudes and colors. A total of 98 galaxies with r′ ≤ 22.5 mag were distributed in four masks, two of them also included the gravitational systems M1, M2, M3, M4, M5, and S1.

Multijob observations (MOS) were carried out in 2013 (program ID: GS−2012B−Q−53) under photometric conditions. The spectra were acquired using the R400+ grating, 1″ slit width, and 2 × 2 binning. Two masks were observed using a central wavelength of 6000 Å. The other two masks were the faintest galaxies and the strongly lensed arcs were included; the nod-and-shuffling technique in band-shuffling mode was used with the blocking filter OG515 and a central wavelength of 7500 Å. The total exposure time in each mask was 4 × 1800 s. All spectra were reduced with the Gemini GMOS package pipeline, following the standard procedures for MOS and nod-and-shuffle observations.

Arc redshifts. While no spectroscopic features (at the 1σ level over the continuum) were found in the spectra of systems M4 and M5, we were able to determine the redshift of the lensed sources in the other four systems (M1, M2, M3, and S1; see Table 1) The spectra of M1.1, M1.2, M1.3, and M1.4 show emission lines associated with [O II]λ3727, at z = 1.10148 (average). In system M2, the spectra of arcs M2.1 and M2.2 show some emission lines ([O II]λ3727, [O III]λ5007, and [O III]λλ4959,5007) placing the object at z = 0.95430 (average). In the spectra of M3.1 and M3.2, a weak emission line is present, consistent with a redshift of z = 2.0834 for [O III]λ5007.

Galaxy redshifts. The redshifts of the galaxies were determined using the programs implemented in the IRAF RV package. The galaxy spectra were separated into early- and late-type populations. For galaxies with clear emission lines (late type), the redshifts were measured with the program RVIDLINES. For early-type galaxies, the redshifts were measured by cross-correlating the spectra with high signal-to-noise templates using the FXCOR program. We were able to determine the redshift of 99 galaxies. We also include in our analysis 113 galaxy redshifts obtained by Yee et al. (1996) and 57 galaxy redshifts derived by G98, with some of the galaxies appearing in both catalogs. The final catalog contains redshift determination for 195 galaxies inside an area of ~0.4 × 0.2 around the MS 0440.5+0204 galaxy cluster. From those 195 galaxies, 10 (from G98) are located at the redshift of the cluster and were not covered by our observations. The average redshift and velocity dispersion of the cluster were calculated using the robust bisection routine of the central location and scale (Beers et al. 1990) with the program ROSTAT and an iterative procedure that applies a 3σ clipping algorithm to remove outliers. We have obtained an average redshift of (Z) = 0.19593±0.00033 (58738±95 km s−1) and a line-of-sight (hereafter LOS) velocity dispersion of σLOS = 771±63 km s−1, with 93 member galaxies.

Table 1
Properties of the Arc Systems

| System |弧| Old Name | α (J2000) | δ (J2000) | zspec | zGiota |
|--------|---|---------|---------|---------|-------|--------|
| 0      | S1 | A1      | 4:43:11.10 | +02:10:10.19 | 0.53223 | 0.53230a |
| I      | M1.1 | A8   | 4: 43: 10.36 | +02: 10: 33.48 | 1.10125 | [0.53, 1.1] |
|   | M1.2 | A9   | 4: 43: 10.65 | +02: 10: 29.04 | 1.10139 |        |
|   | M1.3 | A12  | 4: 43: 10.55 | +02: 10: 05.50 | 1.10176 |        |
|   | M1.4 | A24  | 4: 43: 09.23 | +02: 10: 18.44 | 1.10147 |        |
| II     | M2.1 | A6   | 4: 43: 10.40 | +02: 09: 57.60 | 0.95434 | [0.60, 1.6] |
|       | M2.2 | A5   | 4: 43: 09.13 | +02: 10: 26.49 | 0.95426 |        |
| III    | M3.1 | A3   | 4: 43: 08.45 | +02: 10: 27.29 | ‡        | [0.59, ∞] |
|       | M3.2 | A2   | 4: 43: 08.35 | +02: 10: 19.15 | ‡        |        |
|       | M3.3 | A20  | 4: 43: 09.38 | +02: 09: 55.92 | ...      |        |
|       | M3.4 | ...  | 4: 43: 10.29 | +2: 10: 23.82 | ...      |        |
| IV     | M4.1 | A18  | 4: 43: 10.09 | +02: 09: 53.41 | ...      | [0.59, 1.5] |
|       | M4.2 | A17  | 4: 43: 09.88 | +02: 10: 24.43 | ...      |        |
|       | M4.3 | A19  | 4: 43: 09.91 | +02: 10: 28.95 | ...      |        |
| V      | M5.1 | A7   | 4: 43: 09.43 | +02: 09: 56.70 | ...      | [0.59, 1.5] |
|       | M5.2 | A16  | 4: 43: 10.20 | +02: 10: 23.98 | ...      |        |

Notes. Columns 2 and 3 list the names used in this work and the original names given by G98, respectively. Column 6 lists the measured spectroscopic redshift (from Paper II). The last column shows in parentheses the range of values predicted by G98. [‡] There is a weak emission line present in the spectrum of both M3.1 and M3.2 at an observed wavelength of 7156 Å that may be [O III]λ2321 at z = 2.0834.

a Spectroscopic value reported by G98. There is a weak emission line present in the spectrum of both M3.1 and M3.2 at an observed wavelength of 7156 Å that may be [O III]λ2321 at z = 2.0834.
2.3. CFHT Data

MS 0440.5+0204 was observed using MegaCam in the Canada–France–Hawaii Telescope (CFHT). This cluster was part of the Canadian Cluster Comparison Project, a comprehensive multiwavelength study targeting 50 massive X-ray-selected clusters of galaxies (e.g., Hoekstra et al. 2012). The images were preprocessed using the Elixir pipeline at CHFT, and the resulting images were then calibrated and combined into a single image using the Megapipe stacking pipeline (Gwyn 2008) at the Canadian Astronomy Data Center. The final g’ and r’ images have a total exposure times of 2440 s and 6060 s, with an average seeing of 0.89 and 0.85, respectively. With these data, we constructed the catalogs used in Section 4.2 to perform the weak-lensing analysis.

2.4. X-Ray Data

The X-ray data of MS 0440.5+0204 were retrieved from the Chandra Data Archive. We processed the level 2 event files with the CIAO 4.11 software package. To map out the extended X-ray emission, we first removed point sources using the “wavedetect” CIAO task. Next, we filtered the image using the 0.2–6 keV energy band. Finally, we adaptively smoothed the data with the “Csmooth” task. The final image has been smoothed with kernels with a minimum scale of 0”5 and a maximum scale of 5”, with a minimum signal to noise of 3” per kernel. The smoothed Chandra image of the cluster in the 0.2–6 keV energy band is shown in Figure 2.

The peak of the X-ray image coincides with the bright central galaxies, and the X-ray isophotes are elongated in the same direction as the mass contours obtained from the lensing analysis (see Section 5). The final X-ray image has a higher spatial resolution than the one reported by Gioia et al. (1998). Note that this is a new reduction of the same data used by Hicks et al. (2006) and Shan et al. (2010). Notice that these X-ray data are used to provide morphological information; however, the X-ray mass estimation supplied by the literature is compared to our lensing mass results in Section 5.

3. The Method

In this section, we present a general review of the mass density profile used in our models and derive some useful relations. We also define our different figure-of-merit functions and explain how they are constructed and computed in our models.

3.1. Mass Density Profile

We model the mass density profile of MS 0440.5+0204 as an NFW profile (Navarro et al. 1996, 1997):

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

This profile, predicted in cosmological N-body simulations, is characterized by the scale radius $r_s$ that corresponds to the region where the logarithmic slope of the density equals the isothermal value and the density $\rho_s$. The scale radius is related to the radius $r_{200}$ through the expression $c_{200} = r_{200}/r_s$, which is commonly called the concentration. The $r_{200}$ parameter is defined as the radius of a spherical volume, inside of which the mean density is 200 times the critical density, $\rho_{\text{crit}}$, at redshift $z$, i.e., $M_{200} = 200 \times (4\pi/3) \rho_{\text{crit}} r_{200}^3$. By integrating Equation (1), it is straightforward to show that the mass contained within a radius $r$ of the NFW halo is given by

$$\bar{M}(r) = 4\pi r^3 \rho_s \left[ \ln (1 + r/r_s) - \frac{r/r_s}{1 + r/r_s} \right].$$  (2)

Similarly, but integrating along one axis, we obtain the surface mass density of the NFW profile:

$$\Sigma(\xi) = 2\rho_s\xi F(\xi),$$  (3)

where $F(\xi)$ is a function (e.g., Golse & Kneib 2002) of the dimensionless coordinate $\xi$, the radius in the YX plane in units of the scale radius, $\xi = (x/r_s, y/r_s)$. Integrating the surface mass density, we obtain the 2D mass inside the radius $\xi$:

$$m(\xi) = 4\pi r^2 \rho_s g(\xi),$$  (4)

where $g(\xi)$ is an expression that can be calculated analytically (Bartelmann & Menechetti 2004).

The mass and the density profile are related to the velocity dispersion (at radius $r_s$) in the LOS through the expression (Mamon & Łokas 2005; Mamon et al. 2013)

$$\sigma_v^2 = \frac{2G}{\Sigma(r_s)} \int_{r_s}^{\infty} K(r/r_s) \bar{M}(r) \rho(r) dr,$$  (5)

where $K(r/r_s) = \sqrt{1 - (r/r_s)^2}$ for an isotropic velocity dispersion. Note that in this case, Equation (5) is equivalent to Equation (7) in Verdugo et al. (2011).

3.2. The Figure-of-merit Functions

The $\chi^2$ has been previously used to quantify the goodness of fit of the lens model (e.g., Jullo et al. 2007; Verdugo et al. 2007, 2011); therefore, we summarize the method here. Let us assume a model whose parameters are $\theta$, with $N$ sources, and $n_i$ the number of multiple images for source $i$. For every system $i$, we compute the position in the image plane $x^i(\theta)$ of image $j$, i.e.,
using the lens equation. Therefore, the total $\chi^2$ from multiple-image system $i$ is

$$
\chi_i^2 = \sum_{j=1}^{n_i} \frac{\left( x_{\text{obs}}^i - x_j(\Theta) \right)^2}{\Delta_y^2},
$$

(6)

where $\Delta_y$ is the measured error in the position of image $j$ and $x_{\text{obs}}^i$ is the observed position. The $\chi_i^2$ from all image systems are added to form $\chi^2_{\text{lens}}$.

We also define a $\chi^2_{\sigma}$ associated with the velocity dispersion as

$$
\chi^2_{\sigma} = \frac{\left[ \sigma_{\text{obs}} - \sigma(\rho_s, r_s) \right]^2}{\Delta_{\sigma}^2},
$$

(7)

where $\sigma_{\text{obs}}$ is the velocity dispersion obtained from the spectroscopic analysis (Paper II), $\Delta_{\sigma}$ is the error, and $\sigma(\rho_s, r_s)$ is calculated through Equation (5). Note that we are assuming that the velocity dispersion at $r_s$ is similar to the velocity dispersion measured at a greater radius.

Finally, the weak-lensing analysis provides us with a 2D mass for the system measured at radius $\xi$; thus, we define

$$
\chi^2_{\text{mass}} = \frac{[m_{\text{WL}}(\xi) - \bar{m}(\rho_s, r_s, \xi)]^2}{\Delta_m^2},
$$

(8)

where $\Delta_m$ is the error in the mass, and $\bar{m}(\rho_s, r_s, \xi)$ is calculated through expression (4). Hence, with a likelihood $L \propto \exp(-\chi^2/2)$, we define

$$
\chi^2_{\text{lens-}r_s} = \chi^2_{\text{lens}} + \chi^2_{\sigma},
$$

(9)

with $\chi^2_{\text{lens}}$ and $\chi^2_{\sigma}$ being the statistical parameters obtained from the fit. We also define

$$
\chi^2_{\text{lens-}r_s, \text{mass}} = \chi^2_{\text{lens}} + \chi^2_{\sigma} + \chi^2_{\text{mass}},
$$

(10)

where we have added the $\chi^2_{\text{mass}}$ related to the mass.

Note that we are not performing a simultaneous fit between data sets, for example, such as those conducted by Verdugo et al. (2016) with dynamics and strong lensing, or the ones by Newman et al. (2013) using weak lensing and strong lensing. Our approach is simpler: the idea is to construct a model that is consistent with other data sets that are calculated independently. In other words, the mass and velocity are added only as constraints to the modeling. This is an alternative when the modeller does not have access to the original data to perform simultaneous fitting.

Finally, we assume the same weight for strong lensing, dynamics, and weak lensing in the total likelihood, but when combining strong lensing and weak lensing, this could be different. As discussed in Umetsu et al. (2015), the contribution of strong lensing to the total fit could bias the result if image systems having similar configurations are not taken into account. However, this is not the case in the model of MS 0440.5+0204, where the four systems are different.

## 4. Mass Modeling

### 4.1. Strong Lensing

We operate our models with the latest version of the LENSTOOL\textsuperscript{11} ray-tracing code, which uses a Bayesian Markov Chain Monte Carlo (MCMC) method (Jullo et al. 2007). To model the dark matter component of this cluster, we consider first a single large-scale clump, and then we add small-scale clumps as perturbations associated with individual cluster galaxies. A single clump was chosen given the elliptical pattern of the lensed arcs and the lack of a secondary luminous subclump. In addition, the single-cluster model has sufficient accuracy (quantified by a small $\chi^2$) to reliably reproduce the image positions of the arcs.

**Large-scale clump.** This component was modeled as an NFW mass density profile (see Equation (1)). To take into account the ellipticity in the lens modeling, we consider the pseudo-elliptical NFW potential proposed by Golve & Kneib (2002). This potential is characterized by six parameters: the position $X, Y$; the ellipticity $e$; the position angle $\theta$; the scale radius $r_s$, and the velocity dispersion $\sigma_s$. In practice, all of the parameters describing the large-scale dark matter clump are allowed to vary in the optimization procedure. In particular, the priors for the velocity and the scale radius are broad and set as follows: $780 \text{ km s}^{-1} \leq \sigma_s \leq 1600 \text{ km s}^{-1}$ and $50 \text{ kpc} \leq r_s \leq 400 \text{ kpc}$, respectively.

**Small-scale clumps.** The cluster galaxy population is incorporated into the lens model as pseudo-elliptical isothermal ellipsoid mass distribution potentials (Limousin et al. 2005; Elaïsdottir et al. 2007). This mass distribution is parameterized by a velocity dispersion, $\sigma_0$, related to the central density of the profile. It has two characteristic radii, $r_{\text{core}}$ and $r_{\text{cut}}$ that define changes in its slope. In the inner region, the profile has a core with a central density $\rho_0$, then in a transition region ($r_{\text{core}} < r < r_{\text{cut}}$), it becomes isothermal, with $\rho \sim r^{-2}$. These parameters are scaled as a function of luminosity:

$$
\begin{align*}
&r_{\text{core}} = r_{\text{core}}^* \left( \frac{L}{L^*} \right)^{1/2}, \\
&r_{\text{cut}} = r_{\text{cut}}^* \left( \frac{L}{L^*} \right)^{1/2}, \\
&\sigma_0 = \sigma_0^* \left( \frac{L}{L^*} \right)^{1/4}.
\end{align*}
$$

(11)

The scaling relation for $\sigma_0$ assumes that mass traces light, and its origin resides in the Faber–Jackson relation (Faber & Jackson 1976), which has been reliable for describing early-type cluster galaxies (e.g., Wuyts et al. 2004; Fritz et al. 2005). These scaling relations are common in lens-modeling techniques (e.g., Limousin et al. 2007b).

The 13 member galaxies in the central part of the cluster (see Figure 1) are early type and, therefore, they satisfy the above scaling relations (Equation (11)). Then, for a given $L^*$ luminosity, we will search for those values of $\sigma_0^*$ and $r_{\text{cut}}^*$ that yield the best fit, while $r_{\text{core}}^*$ is fixed at 0.15 kpc (the value is arbitrary; see the discussion in Limousin et al. 2009). These two parameters, $\sigma_0^*$ and $r_{\text{cut}}^*$, describing the cluster galaxy population add two more free parameters to the optimization procedure. For a galaxy with $r'$ magnitude equal to 17.70

\textsuperscript{11}This software is publicly available at http://projets.lam.fr/projects/lenstool/wiki.
(galaxy 100606(A); see Paper II), we set the following limits: 150 km s^{-1} < \sigma_0 < 300 km s^{-1} and 5 kpc < r_{cut} < 50 kpc. These limits in the parameters are motivated by galaxy–galaxy lensing studies in clusters (Natarajan et al. 1998, 2002a, 2002b; Limousin et al. 2007a) and numerical simulations (Limousin et al. 2009). The other parameters describing the galaxy-scale clumps are set as follows: the center of the dark matter halo is assumed to be the same as for the luminous component, and the ellipticity and position angle of the mass are assumed to be the same as those of the light. The luminosity distribution of a given galaxy may not trace the dark matter distribution in its halo; however, there is evidence that the projected mass and light distributions tend to be aligned (Keeton et al. 1998).

Considering the perturbation produced by galaxies 896(B) and 599(D) on different arclets (see Figure 1), some of their dark matter halo parameters were allowed to vary in the optimization procedure. Therefore, every model is computed and optimized in the image plane with a total of 17 free parameters that we detail below: the 6 parameters, \{X, Y, e, \theta, r_{c}, \sigma_{i}\}, for the main halo; 2 parameters, \{r_{cut}, \sigma_{0}\}, for the smaller-scale clumps; the redshift, \{z_{MS}\}, for the arc system M5; and finally, 4 parameters, \{e, \theta, r_{cut}, \sigma_{0}\}, for each of the galaxies 896(B), and 599(D). All of the parameters are allowed to vary with uniform priors. As an additional test of our best models (the three models presented in Section 4.3), we used the fact that no counterimage was identified for the spectroscopically confirmed background source, S1. Namely, with the parameters obtained with our best models, we use LENSTOOL to search for possible counterimages of S1, but in all cases, S1 is returned as a single image. We try to include system M4 in our calculations with the redshift as a free parameter, but we were unsuccessful in reproducing the configuration. When we include this system, the \chi^2/DOF is more than seven times the one obtained in our best models, i.e., \chi^2/DOF \approx 15. Therefore, we exclude this system from our final models. It is important to stress that we are not claiming that M4 is not an arc system, just that the information is not enough to reproduce either its configuration or its redshift. Including this system would have increased the number of free parameters of the central galaxies, because their perturbation may affect the shape and position of the arcs.

4.2. Weak Lensing

To estimate the weak-lensing mass of MS 0440.5+0204, we followed the methodology described in detail in Foëx et al. (2012), which can be summarized as follows. The detection and selection of objects were performed in the r' band with SEXTTRACTOR (Bertin & Arnouts 1996) in dual mode. To sort stars, galaxies, and false detections, we combined different criteria: the size of the objects compared with the point-spread function (PSF hereafter), the position in the magnitude/central flux diagram with respect to the star branch, and their stellarity index as given by the CLASS_STAR parameter of SEXTTRACTOR. After this first step, we obtained a number density of \sim 23 arcmin^{-2} for galaxies and \sim 2 arcmin^{-2} for stars. To select the lensed sources, we removed galaxies located in the red sequence, estimated in the r'-(g' - r') diagram, down to a magnitude of \m_r = 23 mag. All galaxies brighter than \m_r = 21 mag and fainter than \m_r = 25 mag were also removed. These cuts led to a number density of \sim 12 arcmin^{-2}.

Next, we estimated the shape of the background galaxies with the software IM2SHAPE (Bridle et al. 2002), as presented in Foëx et al. (2012, 2013). Briefly, the catalog of stars was used to estimate the local PSF, which was then interpolated at the position of each galaxy. The code convolves the PSF field to an elliptical model for the galaxy shape and runs an MCMC sampler to find the parameters that minimize the residuals. The method was calibrated in the STEP1 simulations (Heymans et al. 2006), and the measured shear was corrected accordingly.

To estimate the lensing strength (i.e., the factor needed to convert shear into mass), we relied on the approach that consists of assuming that similar selection criteria applied to similar data sets must generate the same redshift distribution. We used the photometric redshifts from the T0004 release of the CFHTLS-DEEP survey that were computed with HYPERZ (Bolzonella et al. 2000). After applying the same photometric selection criteria to the CFHTLS sample, we calculated for each galaxy the angular-diameter distance–source distance \Dls and observer–source distance \Dos. By setting \Dls = 0 whenever the photometric redshift is smaller than the cluster redshift, we automatically correct for the contamination of our sources catalog by foreground galaxies. We estimate the average lensing strength to be \beta = \langle \Dls/\Dos \rangle = 0.66, a similar value to that of Hoekstra et al. (2015) who used \beta = 0.656 to derive their lensing mass estimate.

The shear profile was computed in logarithmically spaced annuli, and fitted from 100 kpc to 2.5 Mpc. We used two parametric mass models: a singular isothermal sphere (SIS) and the classical NFW model. We propagated the uncertainties in the galaxy’s shape parameters to the mass estimate by generating 5000 Monte Carlo realizations of the shear profile. Each of these profiles was obtained after randomly sampling the ellipticity parameters of each galaxy, assuming a normal distribution for the mean and the standard deviation given by Im2shape. The SIS and NFW best-fit parameters and their corresponding 1\sigma uncertainties are given in Figure 3, where we plot the average density contrast \Delta \Sigma (see Foëx et al. 2012, 2013). This quantity is related to the shear and is
The density contrast of a circular-symmetric lens is estimated from the tangential shear $\gamma_t$, which produces a galaxy source $i$ located at the concentric radius $r_i$. For the SIS and NFW profiles, respectively, the critical density reads $\Sigma_{\text{crit}} = \frac{c^2}{4\pi G D_{\text{ls}}}$, and $D_{\text{ls}}$ is the angular-diameter observer–lens distance.

12 The density contrast of a circular-symmetric lens is estimated from the tangential shear $\gamma_t$, which produces a galaxy source $i$ located at the concentric radius $r_i$. For the SIS and NFW profiles, respectively, the critical density reads $\Sigma_{\text{crit}} = \frac{c^2}{4\pi G D_{\text{ls}}}$, and $D_{\text{ls}}$ is the angular-diameter observer–lens distance.

13 The parameter files (from modeling) containing the information presented in this section are available on request from the corresponding author.

The Astrophysical Journal, 897:4 (17pp), 2020 July 1

Verdugo et al.

4.3. The Three Models

Model $M_\text{NFW}$. Our first strong-lensing model uses only the positions of the arcs as constraints. Some of the arc systems show multiple subcomponents (bright knots) that can be conjugated as different multiple-image systems (e.g., Broadhurst et al. 2005; Verdugo et al. 2011, 2016), increasing the number of constraints. Considering systems M1 (four images), M2 (decomposed in three systems with two images each), M3 (arranged in two systems with four images each), and M5 (two images), we get a total of 20 images (see Figure A1 in Appendix A). For these systems, we have 26 constraints (if there is a total of $\sum_{i=1}^{N} n_i = N$ images, then there are $\sum_{i=1}^{N} (n_i - 1) = N_e$ constraints in the models assuming that the position $(x, y)$ of the images is fitted). This means 9 degrees of freedom (DOFs) for this model (17 free parameters). The results of our fits are summarized in Table 2, where we present all the parameters of each model in such a way that they can be used by the community. The first column identifies the model parameters. The first two rows are the position $X$, $Y$ in arcseconds relative to the BCG. Rows 3–8 list the parameters associated with the NFW profile: ellipticity, position angle, scale radius, velocity, concentration, and mass. Rows 9–11 provide the parameters related to the galaxy-scale clumps. The results presented in this table allow us to compare parameters among the three different models. Figure 5 shows the critical lines (for a source at $z = 0.95$, $z = 1.10$, and $z = 2.08$) and the predicted positions of the images for systems M1, M2, M3, M5, and S1. We obtain $X^2/DOF = 2.2$ for this model (see Table 3), with the most precise value for the redshift of system M5 equal to $z = 1.5^{+0.2}_{-0.1}$.

In Appendix C (see Figure C1), we present the 2D contours and the PDFs for the parameters of the main halo as well as the $r_{\text{vir}}$ and $\sigma_0$ values for the galaxy-scale clumps. Although the degeneracy between parameters is common in lensing modeling, and has been studied previously (see discussion in Jullo et al. 2007), we want to show how these contours change when adding the additional constraints to the lensing models. Note that the position of the large-scale dark matter clump is found to be offset from the BCG.

Model $M_\text{NFW-CL}$. Our second strong-lensing model adds an additional constraint, the velocity dispersion of the cluster (see Equation (7)). This implies 10 DOFs for this model. The value of the LOS velocity dispersion was calculated in Paper II, $\sigma_{\text{obs}} = 771^{+61}_{-41}$ km s$^{-1}$, using 93 confirmed members. This velocity is slightly lower, but within the errors, than the value calculated from weak-lensing analysis using an NFW profile, $\sigma_{\text{WL}} = 894^{+77}_{-44}$ km s$^{-1}$ (see Section 4.2). In Table 2, we present the best-fit values, and in the left panel of Figure B1, the critical lines and the predicted positions of the images. Comparing this figure with Figure 5, we conclude that both models recover the image positions equally well, which is consistent with their similar

![Figure 4. PDFs and contours of the $c_{200}$ and $M_{200}$ parameters using weak-lensing analysis. The three contours stand for the 68%, 95%, and 99% confidence levels. The values obtained for our best-fit model are marked by a green square and vertical lines in the 2D contour and in the 1D histograms, respectively (the asymmetric errors are presented in Table 2).]
Table 2

| Parameter | $M_{\text{lens}}$ | $M_{\text{lens}-\sigma_l}$ | $M_{\text{lens}-\sigma_l-\text{mass}}$ |
|-----------|------------------|-----------------------------|----------------------------------|
| $X^a (\text{arcsec})$ | 4.1$^{+0.9}_{-0.7}$ | 3.9$^{+0.9}_{-0.8}$ | 3.6$^{+0.7}_{-0.4}$ |
| $Y^a (\text{arcsec})$ | -2.0$^{+0.8}_{-0.5}$ | -1.4$^{+0.3}_{-0.4}$ | -1.9$^{+0.3}_{-0.4}$ |
| $\epsilon^b$ | 0.25$^{+0.04}_{-0.03}$ | 0.26$^{+0.03}_{-0.02}$ | 0.25$^{+0.04}_{-0.02}$ |
| $\theta (\text{arcsec})$ | 165.3$^{+1.2}_{-1.7}$ | 165.8$^{+0.9}_{-0.9}$ | 166.8$^{+0.6}_{-0.6}$ |
| $r_s$ (kpc) | 172$^{+32}_{-28}$ | 177$^{+32}_{-30}$ | 132$^{+19}_{-22}$ |
| $\sigma_l$ (km s$^{-1}$) | 1248$^{+64}_{-76}$ | 1226$^{+62}_{-76}$ | 1203$^{+47}_{-51}$ |
| $c_{200}$ | 8.3$^{+3.2}_{-1.0}$ | 9.8$^{+3.4}_{-1.4}$ | 9.9$^{+2.2}_{-1.4}$ |
| $M_{200} (10^{14}M_\odot)$ | 3.9$^{+0.9}_{-1.2}$ | 3.3$^{+0.6}_{-0.8}$ | 3.1$^{+0.6}_{-0.6}$ |
| $\epsilon_{\text{cut}}$ (kpc) | ... | ... | ... |
| $\epsilon_{\text{cut}}$ (kpc) | 47$^{+12}_{-14}$ | 41$^{+8}_{-9}$ | 35$^{+9}_{-10}$ |
| $\sigma_0$ (km s$^{-1}$) | 325$^{+51}_{-53}$ | 314$^{+59}_{-53}$ | 279$^{+89}_{-85}$ |

Notes.

$^a$ The position in arcseconds relative to galaxy 896(B) located at $\alpha = 4:43:09.8, \delta = +2:10:18.2$.

$^b$ The ellipticity is defined as $\epsilon = (a^2 - b^2)/(a^2 + b^2)$, where $a$ and $b$ are the semimajor and semiminor axes, respectively, of the elliptical shape. Some of the dark matter halo parameters of galaxies 896(B) and 599(D) were allowed to vary in the optimization procedure. Values in square brackets are not optimized.

Figure 5. Critical and caustic lines for model $M_{\text{lens}}$ (for a source at $z = 0.95, z = 1.10$, and $z = 2.08$, from inner to outer radii, respectively). The circles show the positions of the images (input data for the model) and the crosses the predicted positions of the lensed images. We follow Figure 1 code colors for each family of lensed images.

$\chi^2/\text{DOF}$ (see also Section 5). The 2D contours and the PDFs for the free parameters are shown in Figure C2. We note that the model $M_{\text{lens}-\sigma_l}$ produces similar critical lines and predicted positions to those for the model $M_{\text{lens}}$. We obtain $\chi^2/\text{DOF} = 2.5$ for this second model, with the best value for the redshift of system M5 equal to $z = 1.5^{+0.2}_{-0.1}$.

Model $M_{\text{lens}-\sigma_l-\text{mass}}$. Finally, we include the weak-lensing mass in our last model. From our weak-lensing data, we calculate the mass at radius $r_s$ and found $\tilde{m}_{\text{WL}} = 8.4 \pm 2.5 \times 10^{13} M_\odot$. This value is used in Equation (8) to set the final constraint to our model, resulting in 11 DOFs. In Table 2, we present the best-fit values, and the right panel of Figure B1 shows the critical lines and the predicted positions of the images. The 2D contours and the PDFs are shown in Figure C3. For this model, we obtain $\chi^2/\text{DOF} = 2.2$, with the best value for the redshift of system M5 equal to $z = 1.5^{+0.2}_{-0.1}$.

5. Discussion

5.1. Including Additional Constraints

The degeneracies depicted in Figures C1–C3 are expected, and similar results have been presented previously by different authors (e.g., Elíasdóttir et al. 2007; Julio et al. 2007). We can appreciate that the 2D contours and the best values for $\sigma_l$ and $r_s$ change significantly in each model, reflecting a degeneracy in mass between the smooth cluster component and the galaxies. Another interesting degeneracy is the one between the lens mass and $z$. The right panel of Figure 6 shows the PDFs and the contours of these parameters for the three models discussed in this work. Note that the inclusion of dynamical constraints reduces the 2D contours (compare the top and middle-right panels of Figure 6), which reflects the fact that the parameters are related through Equation (5). Because the parameter $\sigma_l$ is constrained by the measured velocity dispersion of the cluster (see Equation (7)), the constraint is spread to the scale radius parameter $r_s$. The tightening of the contours also decreases the errors associated with the parameters. For example, for $r_s$, the relative error between $M_{\text{lens}}$ and $M_{\text{lens}-\sigma_l}$ changes from 0.30% to 0.25%.
Figure 6. Left column: PDFs and contours of the parameters $c_{200}$ and $M_{200}$. Right column: PDFs and contours of the parameters $r_s$ and $\sigma_s$. From top to bottom, $M_{\text{lens}}$, $M_{\text{lens, min}}$, and $M_{\text{lens, min, max}}$, respectively. The three contours stand for the 68%, 95%, and 99% confidence levels. The values obtained for our best-fit model are marked by a green square and by vertical lines in the 1D histograms (the asymmetric errors are presented in Table 2).
The left panel of Figure 6 also presents the PDFs and the contours for \(c_{200}\) and \(M_{200}\). We observe an analogous behavior: a reduction of the contours when adding additional constraints. As discussed in Verdugo et al. (2016), this result is expected because dynamics do constrain the scale radius of the NFW mass profile, a parameter that is not accessible to strong lensing alone. Note also (Figure 6) that the inclusion of the mass as an additional constraint in model \(M_{\text{lens}−\sigma,\text{mass}}\) improves, although slightly, the result and tightens the 1D histograms (compare the middle and bottom panels). In general, the contours in \(M_{\text{lens}−\sigma,\text{mass}}\) are better defined, with the best solution nearly in the center of the 1σ region. This suggests that the inclusion of the weak-lensing mass is complementary to the use of velocity dispersion. This is interesting because weak lensing (as well as strong lensing) provides the total amount of mass and its distribution with no assumptions about the dynamical state of the matter producing the lensing effect.

Interestingly, the three models are equally good to reproduce the image positions of the arcs, which is quantified not only through the \(\chi^2_{2\text{DOF}}\) (very similar in the models) but also through their rms image (rmsi).14 This value provides another way to quantify the goodness of the model (see Elíasdóttir et al. 2007; Limousin et al. 2007b). We obtain \(\text{rmsi} = 0.50, 0.56, \) and \(0.54\) for \(M_{\text{lens}−\sigma}, M_{\text{lens}−\sigma,\text{mass}}, \) and \(M_{\text{lens}−\sigma,\text{mass}}\), respectively. Their similar values are consistent with the predicted positions of the lensed images and the critical lines depicted in Figures 5 and B1; they look extremely similar. We can also appreciate that our models do not predict any counterimages for the single lens, in accordance with our assumption that this is a distorted image of a single source.

Finally, as a supplementary comparison between models, we calculate the inverse of the square root of the the covariance matrix. This FoM (see Albrecht et al. 2006 and also Magaña et al. 2015 for an application to lensing models) quantifies the ability of the observational data set to constrain the parameters; a larger FoM means stronger constraints. For the parameters \(r_s−\sigma_s\), the ratios between FoM’s relative to model \(M_{\text{lens}−\sigma,\text{mass}}\) are \(0.58, 0.93,\) and 1, respectively (see Table 3). This confirms that the inclusion of dynamical restrictions (\(M_{\text{lens}−\sigma}\)) gives slightly more stringent constraints than those provided by the model \(M_{\text{lens}}\). The incorporation of the mass in model \(M_{\text{lens}−\sigma,\text{mass}}\) increases the FoM, although not appreciably. Similarly, we observe the same behavior for the parameters \(c_{200}−M_{200}\). We conclude that the three models are equally good in reproducing the strong-lensing features presented in MS 0440.5+0204. However, they are different at large scale, with a slightly better performance for model \(M_{\text{lens}−\sigma,\text{mass}}\). This is because the strong-lensing model is sensitive to the mass distribution at inner radii, whereas the dynamics (Verdugo et al. 2016) as well as the weak-lensing mass provide constraints at larger radius. Thus, in the next sections, we concentrate our discussion on the \(M_{\text{lens}−\sigma,\text{mass}}\) model as the favored model.

### 5.2. Overconcentration

Galaxy clusters with prominent strong-lensing features commonly have high concentration values when modeled with NFW profiles (e.g., Gavazzi et al. 2003; Kneib et al. 2003; Broadhurst et al. 2005; Comerford & Natarajan 2007; Oguri et al. 2009), larger than those predicted by \(N\)-body simulations of cluster formation and evolution in a \(\Omega_m\)-dominated universe (e.g., Bullock et al. 2001). This overconcentration is interpreted as the cluster dark matter halo departure from spherical symmetry (e.g., Gavazzi 2005), with the major axis of their triaxial geometry oriented toward the line of sight (e.g., Corless & King 2007; Corless et al. 2009). Indeed, strong-lensing clusters could represent a biased population with their major axes aligned with the line of sight (Hennawi et al. 2007; Oguri & Blandford 2009; Meneghetti et al. 2010; Giocoli et al. 2016). Thus, a value of 9.9±4.2 (Table 2) for a galaxy cluster such as MS 0440.5+0204 is not surprising. Besides that, this value is consistent with the value obtained from weak lensing, \(c_{200} = 9.3±3.3\), which supports the idea of a highly concentrated cluster. In the left panel of Figure 6, we show the PDFs and the contours for the parameters \(c_{200}\) and \(M_{200}\), which display similar shape among models. This high concentration is consistent with the analysis presented in Paper II, because the cluster displays evidence of having an elongation along the LOS. Interestingly, the redshift distribution of the galaxies reveals three subclumps (see Paper II), indicating that the cluster core might be experiencing a merging process along the LOS. This evidence of substructures is important because, as mentioned before, they can bias the strong-lensing results (Bayliss et al. 2014), as they are sensitive to the projected mass along the LOS.

#### 5.3. Is MS 0440.5+0204 Part of a Larger Structure?

We discuss here the projected mass map from the best-fit model of \(M_{\text{lens}−\sigma,\text{mass}}\). Given our constraints, the strong-lensing mass is reliable up to the scale radius \(r_s = 132\pm10\), a value beyond which the mass is extrapolated. However, as we will see later, this is consistent with other measurements. Figure 7 shows the 2D mass map from the strong-lensing model (magenta lines) superimposed onto the 2D mass map from weak lensing. The strong-lensing contours show the projected surface mass density at values of 3, 5, 10, 20, and 150 \(\times 10^5 M_\odot\) arcsec\(^{-2}\). Interestingly, their direction (position angle) is consistent with the one from the contours of the 2D weak-lensing signal. Furthermore, we use the task “ellipse” in IRAF to fit ellipses to the smoothed X-ray image of the cluster to obtain the position angle. We find a value of 160° ± 4°, in agreement with the position angle obtained from strong lensing (see Table 2). The elongation in the MS 0440.5+0204 mass is likely produced by the influence of a second structure (a group or galaxy cluster) in the northeast direction. The red cross in the top-left corner of Figure 7 depicts the position of the cluster ZwCl 0441.1+0211; unfortunately, this object does not have a reported redshift (see discussion in Paper II).

Note also that the peak of the weak-lensing mass contours (dark green square on Figure 7) is shifted with respect to the center of the cluster (i.e., the BCG), which is tempting to associate with the presence of another galaxy cluster. However, the uncertainty in the position of the mass peak has a lower limit of 29″ (≈94 kpc), which means that the position of the BCG is within the measurement error. Analyzing 25 galaxy clusters, Oguri et al. (2010) found that the mass centroid obtained from weak lensing can be constrained with an accuracy of 50 kpc in radius, but some clusters showed significant offsets, up to 100 kpc. These kinds of offsets are common and could be caused by shape noise (Dietrich et al. 2012). Martinet et al. (2016) showed that fake lensing peaks are expected in weak-lensing maps due to sampling and shape

---

14 The rmsi is defined through the expression: \(\text{rmsi} = \sqrt{\frac{1}{n−2}\sum_{i=1}^{n}[(\xi_{\text{obs}} − \xi(i))\sigma_i]^2}\), where \(n\) is the number of images for the system.
noise. This effect could be responsible for the weak-lensing peak near the position of ZwCl 0441.1+0211.

5.4. Comparison with Other Works

The projected mass as a function of radius obtained by integrating the 2D dimensional map (for model $M_{\text{lens} - \alpha, \text{mass}}$) is shown in Figure 8. The gray-shaded region shows the area where the arc systems lie, i.e., from the nearest radial arc to the farthest tangential one from the center. Assuming a flattened potential (Blandford & Kochanek 1987), G98 derived a projected mass distribution of $(6.6-9.5) \times 10^{13} M_\odot$ for the central $24^\circ$ region encircled by the arcs. Their model was constructed assuming that the multiple-image systems were the product of five different background sources (see Table 1). Our results are consistent with this upper value (see the blue arrow in Figure 8).

MS 0440.5+0204 was part of the sample of Sand et al. (2005), who studied gravitationally lensed arcs in clusters of galaxies. The authors reported the length-to-width ratio of the arcs and their magnitudes, but not the mass. Shan et al. (2010) calculated the mass from strong lensing, but assuming a spherical matter distribution and using $m_{\text{lens}} = \pi r^2 \alpha_\text{arc} \Sigma_{\text{crit}}$. As the authors point out, their model is not realistic and probably overestimates the mass. They reported a value $m_{\text{lens}} = 1.23 \times 10^{14} M_\odot$ within a radius of $r_{\text{arc}} = 0.1 \text{ Mpc}$. From our best model, we find a projected mass of $5.6 \pm 0.3 \times 10^{13} M_\odot$ inside the same radius, which could explain the discrepancy between the $m_{\text{lens}}$ and $m_{\text{X-ray}}$ reported by those authors ($m_{\text{lens}}/m_{\text{X-ray}} \sim 3$). Indeed, using the values of $kT$, $c$, and $\beta$ reported by Hicks et al. (2006), we calculated the projected mass inside the radius $R_{500}$ following the procedure of Wu (1994, 2000). We found $M_{\text{X-ray}} = 3.5 \pm 0.7 \times 10^{14} M_\odot$, which agrees, within the errors, with our mass estimate for the same radius (see Figure 8). In the same figure, we note also that although the values for the mass at $r_{500}$ reported by Hoekstra et al. (2012, 2015) are slightly smaller than those obtained from our lensing model, they agree within the errors. Hoekstra et al. (2015) performed a weak-lensing analysis of MS 0440.5+0204 using the same data set but with a different approach. In particular, they did not directly fit the concentration parameter of the NFW model but rather relied on the mass–concentration [c(M)] relation from Dutton & Macciò (2014). Our results suggest that the concentration of MS 0440.5+0204 is higher than expected from these $c(M)$ relations; therefore, we can anticipate different results in the inferred weak-lensing mass estimate. Hoekstra et al. (2015) quote a mass $M_{500} = 3.0 \pm 1.5 \times 10^{14} M_\odot$ at a radius $R_{500} = 1.28 \text{ Mpc}$. Within the same aperture, our best-fit model gives a mass $M(1.28 \text{ Mpc}) = 3.5 \pm 0.9 \times 10^{14} M_\odot$, which is $\approx 15\%$ larger than the results reported by Hoekstra et al. (2015) but in agreement within the uncertainties.

6. Conclusions

In this work, we have reconstructed the 2D mass distribution of MS 0440.5+0204. We have presented a detailed strong-lensing analysis of the core of this galaxy cluster, including four multiply imaged systems as constraints, three of them with spectroscopic confirmation. Extending the investigation to large clustercentric distance, we have combined the strong-lensing analysis with other data sets calculated independently, namely the velocity dispersion and the weak-lensing mass. Our analysis is threefold: a strong-lensing-only mass model (in which 13 galaxies have been added as a perturbation to the cluster potential), a model combining strong lensing with
dynamical information (following the method of Verdugo et al. 2011), a model combining strong lensing, dynamics, and the weak-lensing-estimated mass.

Verdugo et al. (2016) have demonstrated that it is possible to determine both the scale radius and the concentration parameter when combining lensing and dynamical constraints. In the present work, we have reached a similar result. Including additional information (from velocity dispersion and mass) allows us to reduce the degeneracy between \( \sigma_s \) and the scale radius \( r_s \), or the mass and the concentration. However, those constraints do not improve the model (regarding the \( \chi^2 \) and the rmsi) in the sense that the new data sets do not considerably affect the predicted positions of the lensing images. Although rmsi values smaller than 1" as those found in this paper (0'50, 0'56, and 0'54) have also been reported in well-modeled clusters (e.g., Newman et al. 2013; Jauzac et al. 2015; Kawamata et al. 2016), they remain unaffected even after imposing tight constraints on the velocity \( \sigma_s \) and the scale radius.

MS 0440.5+0204 seems to be a highly concentrated galaxy cluster (either calculated with weak lensing \( c_{200} = 9.3^{+3.8}_{-3.5} \) or with strong lensing alone \( c_{200} = 8.3^{+3.2}_{-1.0} \)), suggesting that the major axis is probably oriented along the LOS. This result is supported by the analysis presented in Paper II. Although it is tempting to associate MS 0440.5+0204 with ZwCl 0441.1+0211, our weak-lensing analysis is not conclusive. More spectroscopic data are necessary to shed light on this problem.

The authors thank the anonymous referee for a careful reading of the submitted paper and the thoughtful comments that helped to improve the document. This research has been carried out thanks to PROGRAMA UNAM-DGAPA-PAPIIT IA102517. M.L. acknowledges CNRS and CNES for support. J.M. acknowledges the support from CONICYT/FONDECYT project 316064. V.M. acknowledges support from Centro de Astrofísica de Valparaíso. J.A.D. is grateful for the support from the UNAM-DGAPA-PASPA 2019 program and the kind hospitality of the IAC. We thank E. Ellingson for sharing her data.

Facilities: Gemini South: GMOS-S; CFHT: Megaprime/Megacam.

Appendix A
Substructures in Arc Systems

Figure A1 shows the position of the images used to construct the models presented in our work. To emphasize the substructure in some of the arc systems, we zoomed in on arc M2.1, M2.2, M3.1, and M3.2. Note the multiple knots in system M2.

Figure A1. HST WFPC2 F702W image, with the local median average of the central region (60'' x 60'') of the MS 0440 cluster subtracted. The circles show the positions of the images (input data for the model). We follow the color code in Figure 1 for each family of lensed images.
Appendix B

Results for Models $M_{\text{lens-}\sigma}$ and $M_{\text{lens-}\sigma_{\text{mass}}}$

In order to compare the results of model $M_{\text{lens}}$ (see Figure 5) with the two other models, i.e., $M_{\text{lens-}\sigma}$ and $M_{\text{lens-}\sigma_{\text{mass}}}$, Figure B1 shows the predicted positions of the lensed images. Note how the three models recover the image positions equally well, and the critical and caustic lines are very similar.

**Figure B1.** Critical and caustic lines for a source located at $z = 0.95$, $z = 1.10$, and $z = 2.08$, from inner to outer radii. The circles show the positions of the images (input data for the model) and the crosses the predicted positions of the lensed images. Left panel: results from model $M_{\text{lens-}\sigma}$. Right panel: results from model $M_{\text{lens-}\sigma_{\text{mass}}}$. We follow the color code in Figure 1 for each family of lensed images.
Appendix C

PDFs and 2D PDFs of the Model Parameters

Figures C1–C3 show the PDFs for the three models discussed in this work. The parameters exhibit the same degeneracies in all models because those depend on the lensing configuration (see Jullo et al. 2007). However, incorporating additional constraints to the lensing models, namely the velocity and the mass, yields better constrained parameters with reduced contours.

Figure C1. PDFs and contours of the model parameters for $M_{\text{lim}}$. The three contours stand for the 68%, 95%, and 99% confidence levels. The values obtained for our best-fit model are marked by a green square and with vertical lines in the 1D histograms (the asymmetric errors are presented in Table 2).
Figure C2. PDFs and contours of the model parameters for $M_{\text{lim}-\alpha}$. The three contours stand for the 68%, 95%, and 99% confidence levels. The values obtained for our best-fit model are marked by a green square and with vertical lines in the 1D histograms (the asymmetric errors are presented in Table 2).
Figure C3. PDFs and contours of the model parameters for $M_{\text{lens},\text{mass}}$. The three contours stand for the 68%, 95%, and 99% confidence levels. The values obtained for our best-fit model are marked by a green square and with vertical lines in the 1D histograms (the asymmetric errors are presented in Table 2).

ORCID iDs

Tomás Verdugo  https://orcid.org/0000-0003-4062-6123
Eleazar R. Carrasco  https://orcid.org/0000-0002-7272-9234
Gael Foëx  https://orcid.org/0000-0001-9271-4155
Verónica Motta  https://orcid.org/0000-0003-4446-7465
Percy L. Gomez  https://orcid.org/0000-0003-0408-9850
José A. de Diego  https://orcid.org/0000-0001-7040-069X

References

Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, arXiv:astro-ph/0609591
Bartelmann, M., & Meneghetti, M. 2004, A&A, 418, 413
Bayliss, M. B., Johnson, T., Gladders, M. D., Sharon, K., & Oguri, M. 2014, ApJ, 783, 41
Beers, T. C., Flynn, K., & Gebhardt, K. 1990, AJ, 100, 32
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Biviano, A., Murante, G., Borgani, S., et al. 2006, A&A, 456, 23
