MS 2053.7–0449: CONFIRMATION OF A BIMODAL MASS DISTRIBUTION FROM STRONG GRAVITATIONAL LENSING

T. Verdugo and J. A. de Diego
Instituto de Astronomía, Universidad Nacional Autónoma de México, AP 70-264, 04510 Mexico DF, Mexico; tomasv@astroscu.unam.mx, jdo@astroscu.unam.mx

AND

Marceau Limousin
Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Marie Vej 30, 2100 Copenhagen, Denmark; marceau@dark-cosmology.dk

Received 2006 June 29; accepted 2007 April 22

ABSTRACT

We present the first strong-lensing study of the mass distribution in the cluster MS 2053–04 based on HST archive data. This massive, X-ray luminous cluster has a redshift $z = 0.583$ and is composed of two structures that are gravitationally bound to each other. The cluster has one multiply imaged system composed of a double gravitational arc. We have performed a parametric strong-lensing mass reconstruction using NFW density profiles to model the cluster potential. We also included perturbations from 23 galaxies, modeled like elliptical singular isothermal spheres, that are approximately within $1' \times 1'$ around the cluster center. These galaxies were constrained in both the geometric and dynamical parameters with observational data. Our analysis predicts a third image that is slightly demagnified. We found a candidate for this counterimage near the expected position and with the same $F702W - F814W$ colors as the gravitational arcs in the cluster. The results from the strong-lensing model show the complex structure in this cluster, the asymmetry and the elongation in the mass distribution, and are consistent with previous spectrophotometric results that indicate that the cluster has a bimodal mass distribution. Finally, the derived mass profile was used to estimate the mass within the arcs and for comparison with X-ray estimates.

Subject headings: galaxies: clusters: general — galaxies: clusters: individual (MS 2053.7–0449) — gravitational lensing

Online material: color figures

1. INTRODUCTION

The possibility that clusters of galaxies might produce gravitational lensing had been considered by several authors (Karoji & Nottale 1976; Dyer & Roeder 1976; Narayan et al. 1984). Twenty years ago Lynds & Petrosian (1986) and Soucail et al. (1987) independently announced the detection of an arc-like structure around the center of Abell 370, but it was Paczynski (1987) who explained the phenomenon in terms of gravitational lensing. Currently the strong and weak regimes of the gravitational lensing effect are often used for constraining the density profiles of galaxy clusters.

The massive clusters of galaxies have a surface density larger than the critical value for strong lensing, approximately $1.0 \text{ g cm}^{-2}$ (Turner et al. 1984). They can produce multiple, strongly elongated images (giant arcs) of background galaxies lying on top of the caustic lines generated by the cluster potential. The first dedicated observational search for gravitational arcs in X-ray clusters was performed by Le Fèvre et al. (1994) using the Einstein Observatory Extended Medium Sensitivity Survey (EMSS). A subsequent search was performed by Luppino et al. (1999) that confirmed the relationship between X-ray luminosities and cluster masses; these authors found strong lensing in eight of 38 clusters. A subsequent survey based on optically selected clusters in the Las Campanas Distant Cluster Survey confirmed these giant-arc fractions (Zaritsky & Gonzalez 2003), and Gladders et al. (2003) report even larger fractions in the Red-Sequence Cluster Survey, showing that giant arcs are quite common in clusters of galaxies.

Since the shapes of these arcs are related to the gravitational potential, strong lensing offers a method to investigate how the baryonic and dark matter are distributed in the inner regions of the cluster. It is also possible to put significant constraints on the cosmological parameters using a set of strong-lensing clusters showing multiple images with spectroscopic redshifts (Soucail et al. 2004). Mass distributions have been modeled for many clusters with giant arcs. In a few cases, accurate models for the mass distribution have led to the identification of additional counterimage candidates (Kneib et al. 1996). In other cases, the detection of the predicted images is uncertain, since they are lost in the cD light distribution (see Gavazzi et al. 2003; Sand et al. 2004). Recently, the gravitational lens cluster A1689 has been modeled in unprecedented detail by Broadhurst et al. (2005) and Limousin et al. (2006). These authors combined HST Deep Advanced Camera images and extensive ground-based spectroscopy, identifying 32 multiply lensed systems.

The aim of this work is to describe the density profile and the mass distribution of MS 2053, using strong lensing. This cluster has a redshift of $z = 0.583$, is not very optically rich in comparison with other $z > 0.5$ EMSS clusters (Luppino et al. 1999), and among these high-redshift clusters has the lowest X-ray luminosity. BeppoSAX observations yield an X-ray luminosity $L_X(2–10 \text{ keV}) = (3.9 \pm 1.0) \times 10^{44} \text{ erg s}^{-1}$ (Della Ceca et al. 2000), and Chandra observations yield $L_X(0.5–2 \text{ keV}) = 3.5 \times 10^{44} \text{ erg s}^{-1}$ (Vikhlinin et al. 2002). Luppino & Gioia (1992) discovered a gravitationally lensed blue arc in deep images of MS 2053, located $\sim 16''$ in the north direction from the brightest cluster galaxy (BCG). The arc is $\sim 11''$ long and breaks into two
clumps, labeled A and B. Recently, Tran et al. (2005) were able to measure the redshift of image A using Keck spectroscopy data and determined the source redshift to be $z = 3.146$.

Hoekstra et al. (2002) detected the weak-lensing (WL) signal induced by this cluster of galaxies from a two-color mosaic of six Hubble Space Telescope (HST) Wide Field Planetary Camera 2 (WFPC2) images. They fitted a singular isothermal sphere model to the observed azimuthally averaged tangential distortion and estimated the cluster velocity dispersion $\sigma = 886$ km s$^{-1}$. Hoekstra et al. (2002) also fitted the predicted profile from the Navarro-Frenk-White (NFW) halo to the observed tangential distortion of MS 2053 and found a concentration parameter (see § 3.1) in good agreement with the predicted value for a cold dark matter (CDM) model without a cosmological constant. Tran et al. (2005) combined images from the wide-field HST WFPC2 and extensive spectral data from the Keck Low Resolution Imager Spectrograph (LRIS) to make a detailed study of the galaxy populations in MS 2053. They found that the cluster is composed of a main cluster (MS 2053-A) and an infalling structure (MS 2053-B) that are gravitationally bound to each other; their study also shows that the extended spatial distribution and lack of spectral segregation in MS 2053-B is consistent with a scenario in which it is in the initial stages of being accreted by the main cluster. Of the 149 spectroscopically confirmed cluster members, 113 belong to MS 2053-A and 36 to MS 2053-B. From this previous research we present in Table 1 a compilation of physical quantities relevant to our work.

Despite all the research cited above, only Hoekstra et al. (2002) have made an attempt to describe the strong-lensing effect for this cluster, and it was done when the arc redshift had not yet been measured. Assuming a redshift of $z = 2$ for the arc and adopting a singular isothermal sphere, these authors found a velocity dispersion of about 1030 km s$^{-1}$, consistent with their WL estimate. They argue that the mass distribution is elongated in the direction of the arc, which can be very well seen in the light distribution, and that this elongation is the cause of the high value of their strong-lensing mass estimate. Such elongation in the distribution of the cluster members was also reported by Tran et al. (2005). To assess this mass distribution with precision, we present a detailed strong-lensing model of the bimodal cluster MS 2053.

In § 2, we describe the observational data and present our sample of field galaxies. In § 3, we define the model of the lensing cluster and depict the profiles for both the cluster and the galaxy-scale mass components. In § 4, we present the main results. In § 5, we summarize and discuss our results. Finally, in § 6, we present the conclusions. 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 and at the redshift of MS 2053, $1''$ corresponds to 4.62 $h^{-1}$ kpc.

### Table 1: Cluster Parameters

| Group        | $z$         | $M_{2053}$ ($M_\odot$) | $\sigma$  | $\sigma'_{\text{WL}}$ | $c$ | Reference |
|--------------|-------------|------------------------|-----------|------------------------|-----|-----------|
| MS 2053      | 0.5866 ± 0.0011 | $3.7 \times 10^{15}$  | 1523 ± 95 | 886$^{+121}_{-69}$     | 2.9 | 1, 2      |
| MS 2053-A    | 0.5840 ± 0.0005 | $1.1 \times 10^{15}$  | 865 ± 71  | ...                    | 3.8 | 1         |
| MS 2053-B    | 0.5982 ± 0.0003 | $1.1 \times 10^{14}$  | 282 ± 51  | ...                    | 5.4 | 1         |

1. Proposal ID 5991.

2. THE OBSERVATIONAL DATA

The HST data were obtained from the Multimission Archive at the Space Telescope Science Institute (MAST). They consist of two WFPC2 images obtained in both the F814W and the F702W filters; the total integration time for each filter was 2600 and 2400 s, respectively. The image reduction was performed using the IRAF/STSDASS 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 crrz to remove cosmic-rays hits. Finally, the background was subtracted and the WFPC2 chips were combined using the task mosaic. In Figure 1 we show a $60'' \times 60''$ field obtained from the final four-chip mosaic frame for the F814W filter; this is the region used in our analysis of strong lensing.

The magnitudes and geometric parameters of the galaxies were measured using SExtractor (Bertin & Arnouts 1996). For the arc photometry we used a different procedure because SExtractor often overestimates the sky background (see Smith et al. 2005), and the arcs are highly distorted. These two circumstances reduce the precision of automatic photometric measurements. Therefore, we employed polygonal apertures to obtain more accurate measurements.

Figure 1. HST WFPC2 F814W image of the central region ($60'' \times 60''$) of the MS 2053 cluster. North is up, and east is left. Arches A and B are identified, as well as the 23 galaxies considered in our models (lowercase letters). The square shows the region enlarged in Fig. 3.
measurements. The vertices of the polygons for each arc were
determined using the IRAF task measure. The vertices of the polygons for each arc were
system. The colors for arcs A (0

tem, 0

tures to measure each arc independently.
We also calculated the F702W − F814W color for the AB arc system. The colors for arcs A (0.41 ± 0.03) and B (0.37 ± 0.06) are

TABLE 2

| ID Name | α (J2000.0) | δ (J2000.0) | F814W (mag) | e | θ₀ | z | σ₀ (km s⁻¹) | Reference |
|---------|-------------|-------------|-------------|---|----|---|-------------|-----------|
| a        | 20 56 21.4  | -04 37 50   | 19.13 ± 0.01| 0.259 | 63.9 | 0.583 | 292.0 ± 10.0 | 1         |
| b        | 20 56 21.5  | -04 37 54   | 20.24 ± 0.01| 0.122 | 81.8 | 0.5894| 234.0 ± 23.0 | 1         |
| c        | 20 56 21.5  | -04 38 00   | 20.93 ± 0.02| 0.577 | -41.1 | 0.5724| 151.0 ± 17.0 | 1         |
| d        | 20 56 20.1  | -04 37 35   | 20.29 ± 0.01| 0.368 | -7.5  | 0.5846| 169.0 ± 14.0 | 1         |
| e        | 20 56 20.5  | -04 38 02   | 20.84 ± 0.01| 0.473 | 32.9  | 0.5737| 129.0 ± 15.0 | 1         |
| f        | 20 56 22.0  | -04 38 13   | 21.47 ± 0.02| 0.278 | 80.2  | 0.5800| 138.0 ± 13.0 | 1         |
| g        | 20 56 23.0  | -04 38 08   | 20.19 ± 0.01| 0.136 | -74.2 | 0.5854| 124.0 ± 13.0 | 1         |
| h        | 20 56 22.1  | -04 38 18   | 21.12 ± 0.02| 0.374 | -9.8  | 0.5835| 181.0 ± 13.0 | 1         |
| i        | 20 56 21.3  | -04 37 57   | 21.09 ± 0.02| 0.299 | 7.1   | 0.5854| <200.0 ± 2.0  | 2         |
| k        | 20 56 19.6  | -04 38 05   | 19.98 ± 0.01| 0.867 | -66.5 | 0.5912| <200.0 ± 2.0  | 2         |
| l        | 20 56 19.6  | -04 37 54   | 21.38 ± 0.02| 0.270 | -68.3 | 0.5824| <200.0 ± 2.0  | 2         |
| m        | 20 56 20.0  | -04 37 49   | 20.88 ± 0.01| 0.361 | -10.1 | 0.5813| <100.0 ± 2.0  | 2         |
| n        | 20 56 21.1  | -04 37 22   | 21.34 ± 0.02| 0.240 | 76.1  | 0.5764| 86.0 ± 3.0    | 3         |
| o        | 20 56 21.9  | -04 37 25   | 21.67 ± 0.03| 0.063 | 26.3  | 0.5799| <200.0 ± 2.0  | 2         |
| p        | 20 56 23.2  | -04 37 50   | 20.83 ± 0.01| 0.203 | 11.7  | 0.5990| <100.0 ± 2.0  | 2         |
| q        | 20 56 20.9  | -04 37 47   | 21.74 ± 0.03| 0.335 | -9.0  | 0.5978| <100.0 ± 2.0  | 2         |
| r        | 20 56 20.3  | -04 37 50   | 20.04 ± 0.01| 0.116 | 65.5  | 0.5990| <100.0 ± 2.0  | 2         |
| s        | 20 56 20.7  | -04 37 32   | 21.41 ± 0.02| 0.855 | 84.8  | 0.5993| <100.0 ± 2.0  | 2         |
| t        | 20 56 20.9  | -04 38 06   | 21.90 ± 0.03| 0.093 | -4.4  | 0.5827| 11.7 ± 0.04   | 3         |
| u        | 20 56 21.7  | -04 37 45   | 21.04 ± 0.02| 0.546 | -47.6 | 0.5876| 138.0 ± 13.0  | 1         |
| v        | 20 56 22.2  | -04 37 31   | 21.99 ± 0.04| 0.217 | 22.3  | 0.5892| 138.0 ± 13.0  | 1         |
| w        | 20 56 20.8  | -04 37 28   | 22.57 ± 0.06| 0.910 | 68.7  | 0.5939| 138.0 ± 13.0  | 1         |
| x        | 20 56 20.5  | -04 37 26   | 21.90 ± 0.04| 0.798 | -48.9 | 0.5782| ...          | ...       |

Notes.—Units of right ascension are hours, minutes, and seconds, and units of declination are degrees, arcminutes, and arcseconds. Col. (1): Identification for each MS 2053 galaxy. Cols. (2) and (3): Right ascension and declination. Col. (4): F814W magnitude. Cols. (5) and (6): Geometric parameters derived from SExtractor: ellipticity e and position angle θ₀, which gives the orientation of the semimajor axis from the horizontal line in the image, measured counterclockwise. Col. (7): Redshift. Cols. (8) and (9): Value and reference for the central velocity dispersion.

a The ellipticities are defined as ε = (a² − b²)/(a² + b²), where a and b are the semimajor and seminor axis, respectively, of the elliptical shape.

b Galaxies that belong to the second cluster component, according to Tran et al. (2005).

References.—(1) Wuyts et al. 2004; (2) Tran et al. 2005; (3) Tran et al. 2003.

3. THE MODEL OF THE LENS

3.1. Dark Matter Component

To study the mass distribution in MS 2053 we use standard lens modeling techniques implemented in the LENSTOOL2 ray-tracing code (Kneib 1993).

We model the cluster with a NFW profile that has been predicted in cosmological N-body simulations (Navarro et al. 1997):

$$\rho(r) = \frac{\rho_{\text{cen}}(z) \delta_{c}(z)}{(r/r_s)^{2}} (1 + r/r_s)^{-3};$$

where \(\delta_{c}(z)\) is a characteristic density contrast, \(\rho_{\text{cen}}(z)\) is the critical density of the universe as a function of the redshift \(z\), and \(r_s\) is a scale radius that corresponds to the region where the logarithmic slope of the density equals the isothermal value. However, higher resolution simulations have suggested that the inner cusp of the NFW profile is too shallow (e.g., Moore et al. 1998; Ghigna et al. 1998).
2000), giving origin to the generalized NFW-type profile. It is worth noting that Gavazzi (2005) studied MS 2137−2353 and obtained similar results by using both profiles for the strong-lensing regime. Therefore, for simplicity we adopt a profile similar to the original NFW. However, we consider an elliptical profile since in modeling the gravitational lens it is important to take into account the asphericity in the cluster potential.

To generalize the spherical model to an ellipsoidal one, we used the “pseudoelliptical” NFW proposed by Golse & Kneib (2002). This potential is characterized by six parameters: the position \(X, Y\); the ellipticity \(c\); the position angle \(\theta\); and \(r_s\) and the velocity dispersion \(\sigma_v\). We define \(\sigma_v\) as

\[
\sigma_v^2 = 4\left(1 + \ln \frac{1}{2}\right)G\rho_{\text{crit}}\delta_c. \tag{2}
\]

This is approximately half of the value proposed by Golse & Kneib (2002), and it is defined in such a way that it represents a realistic velocity dispersion at radius \(r_s\) and not only a scaling parameter (see the Appendix).

We can relate this \(\sigma_v\) to the virial mass \(M_\Delta\) and the virial radius \(r_\Delta\) of the cluster. If we define \(r_\Delta\) as the radius of a spherical volume within which the mean density is \(\Delta\) times the critical density at the given redshift \(z\) (\(M_\Delta = \frac{3}{2}\pi r_\Delta^3\rho_{\text{crit}}\Delta\)), then the virial radius can be expressed by

\[
r_\Delta = \left[\frac{2M_\Delta G}{\Delta H(z)^2}\right]^{1/3} . \tag{3}
\]

The value of the characteristic overdensity \(\Delta\) is taken from the solution to the collapse of a spherical top-hat model. Its value depends on \(\Omega\) and can be approximated by \(\Delta = 178\Omega_m^{0.45}\) when \(\Omega_m + \Omega_\Lambda = 1\) (Lacey & Cole 1993; Eke et al. 1996, 1998). By integrating equation (1) and using the above definition, it is straightforward to show that the ratio between the virial and scale radii, which is commonly called the concentration, \(c = r_\Delta/r_s\), is related to \(\delta_c\) by

\[
\delta_c = \frac{\Delta}{3\ln(1 + c) - c/(1 + c)}. \tag{4}
\]

By solving equation (4), we can determine the concentration value given \(z\), \(\sigma_v\), and \(r_s\).

### 3.3. Galaxy-Scale Mass Components

For simplicity, we modeled the 23 galaxies using elliptical singular isothermal spheres (ESIs) for the individual galaxies (Blandford & Kochanek 1987). These galaxy-scale mass components were included as perturbations into the cluster potential. The potential for ESIs is given by

\[
\phi(r, \theta) = 4\pi \frac{\sigma_e^2}{c^2} D_{\text{ES}} \frac{D_{\text{OS}}}{r} \sqrt{1 + c \cos[2(\theta - \theta_0)]}, \tag{5}
\]

where \(\sigma_e^2\) corresponds exactly to the true three-dimensional velocity dispersion and \(\theta_0\) is the position angle. Note that equation (5) is defined completely if we can obtain the values for \(\sigma_e^2\), the ellipticity \(\epsilon\), and \(\theta_0\) from observational data. The luminosity distribution of a given galaxy does not trace the dark mass distribution in its halo, but there is evidence that the projected mass and light distributions tend to be aligned (Keeton et al. 1998). Koopmans et al. (2006) report this alignment between light and mass for a sample of early-type lens galaxies, for which the isophotal and isodensity contours trace each other. These authors also report an alignment between the stellar component and the singular isothermal ellipsoids used in their lens models. Following these authors, we infer the parameters of the galaxy-scale mass components from the position, the ellipticity, and the position angle of the luminous component.

Wynter et al. (2004) have measured the velocity dispersions of several elliptical and lenticular galaxies in the cluster, and Tran et al. (2005) have estimated the internal velocity dispersion for the whole cluster sample with \(M_{\text{Be}} \leq -18 + 5 \log h\). Nevertheless, the individual data for each galaxy have not been published for all the objects. To minimize the number of free parameters in the model, the galaxies for which the velocity dispersion is not available (see Table 2) are scaled as a function of luminosity following the Faber-Jackson relation (Faber & Jackson 1976),

\[
\sigma_0 = \sigma_0 \left(\frac{L}{L_\odot}\right)^{1/4}, \tag{6}
\]

which transformed to magnitudes is expressed by \(\log \sigma_0 = -0.1m + K\), where \(K\) is a scaling factor. In this way the cluster members are incorporated into the model as a galaxy-scale perturbation with all their individual parameters fixed.

### 3.3. The Model Optimization

In principle, the optimizations in the source plane and in the image plane are mathematically equivalent. However, we chose to use the source plane because the solutions are numerically simpler and faster to compute. The fit of the model is optimized by mapping the positions of the multiple images back to the source plane and requiring them to have a minimal scatter. To achieve this goal we use a figure-of-merit function, \(\chi^2\), to quantify the goodness of the fit for each trial of the lens model, i.e., a function that measures the agreement between our data and the fitting model for a particular choice of the parameters:

\[
\chi^2 = \chi^2_{\text{pos}} + \chi^2_{\text{shape}}. \tag{7}
\]

The first term is constructed as follows: given a model, we compute the position in the source plane \(x^S_i, y^S_i\) for each observed image \(1 \leq i \leq N\), using the lens equation. The position in the source plane \(x^S_i, y^S_i\) is constructed from the N sources, and the \(\chi^2\) for the position is defined by

\[
\chi^2_{\text{pos}} = \frac{1}{N} \sum_{i=1}^{N-1} \frac{(x^B_i - x^S_i)^2 + (y^B_i - y^S_i)^2}{\sigma_{\text{pos}}} , \tag{8}
\]

where \(\sigma_{\text{pos}}\) is the error in determining the position of the source.

The second term, \(\chi^2_{\text{shape}}\) measures how well the shape of the arcs is reproduced by the model. Given a model, we compute the complex deformation, \(\vec{\tau}\), for each observed image \(1 \leq i \leq N\), using the lens equation for complex deformation (see the discussion in Kneib et al. [1996] and their detailed derivation). The complex deformation of the image represents how much the intrinsic source shape changes due to the induced deformation caused by the potential. This complex deformation is a function of the ellipticity, the position angle, and the lens potential. The complex deformation is an equivalent form in the source plane, \(\vec{\tau}^S = \vec{\tau}^f(\epsilon^S, \theta^S, \varphi)\).

We define \(\tau_{x,i}^S\) and \(\tau_{y,i}^S\) as the components of the
complex deformation in the source plane, and $\sigma_{\text{shape}}$ as the error in determining the shape of the source, we get

$$
\chi^2_{\text{shape}} = \frac{1}{N} \sum_{i=1}^{N-1} \left( \frac{(\tau_{x,i} - \tau_{x,i-1})^2 + (\tau_{y,i} - \tau_{y,i-1})^2}{\sigma_{\text{shape}}} \right). \tag{9}
$$

The errors, $\sigma_{\text{pos}}$ and $\sigma_{\text{shape}}$, have to be measured in the source plane, but these are not directly observable quantities. To avoid this problem, we assumed that the errors in the source plane arise from the measurements performed in the image plane. Thus, we arbitrarily set $\sigma_{\text{pos}}$ and $\sigma_{\text{shape}}$ to 0.2" and 0.8" respectively, because these are the accuracies achieved in the position and shape measurements from the HST images. This is qualitatively correct in the sense that for poor models, the merit function (eq. [7]) gives large values and thus indicates a poor fit.

In version 6.2 of LENSTOOL the Bayesian Markov chain Monte Carlo (MCMC) package BAYESIS$^3$ has been implemented. The Bayesian MCMC method will be described in a dedicated forthcoming publication (Jullo et al. 2007) This package allows us to obtain the errors and to avoid local $\chi^2$ minima. The MCMC sampler draws random models according to a user-defined model parameter space and computes their $\chi^2$. Progressively, it converges to the most likely parameter space and outputs the best solution. We used these random models to perform the statistics to compute error bars on the estimation of the parameters and to constrain the positions of the bimodal cluster (see the next section).

4. RESULTS

4.1. A Simple Model for Arc A

The only arc for which the redshift has been measured is arc A (Tran et al. 2005). The first step was to demonstrate that arcs A and B belong to the same source. But a model that includes only arc A and not arc B is unconstrained, and the data available are not adequate to break the degeneracy. Given the general characteristics of the cluster and a gravitational lensed source with the properties inferred from arc A (i.e., redshift and arc position), we investigate whether or not this source would produce multiple images. To achieve this task, we constructed a simple spherical model with a NFW profile and its central 23 galaxies modeled like elliptical isothermal spheres.

Using equation (3) we calculated the virial radius of the cluster. For a virial mass of $3.7 \times 10^{13} M_\odot$ and a cluster redshift of $z = 0.583$ (see Table 1), $r_s \approx 2.9$ Mpc. Therefore, $r_s \approx 1$ Mpc with a concentration parameter $c = 2.9$ (according to the prescription described in § 3.3 of Eke et al. 2001). The characteristic velocity, $\sigma_c = 1600$ km s$^{-1}$, is obtained from equation (2). With these values ($r_s$ and $\sigma_c$), the NFW profile is fully characterized. At this stage, the ellipticity and the position angle are set to zero.

The velocity dispersion in the galaxy-scale mass components was fixed using the scaling relation given by equation (6). We found that the early-type cluster galaxies in the sample of Wuyts et al. (2004) satisfy the Faber-Jackson relation with $\log \sigma_e = -(0.1 \pm 0.05) m + (4.4 \pm 1.1)$. For the remaining subset of galaxies not included in the Wuyts et al. (2004) sample, only galaxy $n$ has a measured $\sigma_0$ value. We supposed that these galaxies also satisfy the relation but with a different scaling factor $K$. In order to construct the new correlation, we considered the luminosity $L^*$ corresponding to the faintest galaxy in Table 2 (galaxy $w$,

$$
M_{\text{FB14W}} = 22.57). \text{ We set } \sigma_n = 64.7 \text{ km s}^{-1} \text{ such that galaxy } n \text{ satisfies this correlation. With these values, the velocity dispersion in the subset of galaxies not included in the Wuyts et al. (2004) sample is expressed by } \log \sigma_0 = -0.1 m + 4.19. \text{ Figure 2 shows both the previous log } \sigma_0 \text{ and the Faber-Jackson relations for the early-type galaxies of Wuyts et al. (2004).}

With all these parameters we calculate the counterimages of arc A. The model obtained with these fixed parameters predicts two more images, one very demagnified and the other only 1 mag below arc A; the position does not agree with arc B, but this shows that for these cluster characteristics and with this redshift of the source, arc A would have multiple counterimages. Therefore, it is entirely plausible that images A and B arise from the same source.

4.2. Bimodality Evidence

Two kinds of models are possible depending on whether we assume that MS 2053 is composed of one or two clusters, as reported by Tran et al. (2005). We performed the fits by setting up the position (two parameters), the velocity, the scale radius, the ellipticity, and the position angle of the cluster as free parameters. These six parameters define the lens potential and comprise all the allowed free parameters for this single cluster case. With just two images and four constraints [if there is a total of $\sum_{i=1}^n n_i = N$ images, then there are $\sum_{i=1}^n 4(n_i - 1) = N_c$ constraints on the models assuming that the position, ellipticity, and position angle of the images are fitted], the model is underconstrained. And it gets worse in the second model, where we have 12 parameters defining the lens potential. Therefore, we expect that there may be a family of models that can fit the data, even with unrealistic parameters. To try to avoid this problem, we have limited one parameter in the NFW, the scale radius $r_s$.

The $r_s$ cutoff was made in the following way: Given the cluster redshift and using equation (3), we can write the scale radius as $r_s \propto M^{1/3}$, where $M$ is the mass found from lensing. Since lensing is sensitive to the integrated mass along the line of sight, we expect discrepancies between the true three-dimensional mass, $M_{\text{true}}$, and the lensing mass estimates of our models, $M_{\text{lens}}$. Following Gavazzi (2005) we consider an axisymmetric NFW density profile with either the major or the minor axis aligned toward the line of sight. Therefore, the net effect of the projection is

$\chi^2_{\text{shape}}$ is written by J. Skilling and is available at http://www.inference.phy.cam.ac.uk/bayesys/.

\footnotesize

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig2}
\caption{Cluster galaxies and Faber-Jackson relation. The solid line shows the correlation for the early-type galaxies of Wuyts et al. (2004) for which the velocity dispersion is available. The dashed line is a similar correlation but for galaxies not included in the Wuyts et al. (2004) sample. The triangle represents galaxy $n$, whose velocity dispersion has been measured by Tran et al. (2003). [See the electronic version of the Journal for a color version of this figure.]}\end{figure}
given by $M_{\text{Lens}} = qM_{\text{true}}$, where $q$ is the ratio between the semi-axis aligned with and the semi-axis perpendicular to the line of sight. Thus, we can write

$$r_s / q^{1/3} = 3^{1/3} M_{\text{true}}.$$  \(\text{(10)}\)

Using the concentrations shown in Table 1 and the axis values from Gavazzi (2005) we can constrain the scale radius for each model. For the one-component model, $0.60 \text{ Mpc} < r_s < 1.74 \text{ Mpc}$, and for the two-component model, $0.30 \text{ Mpc} < r_s < 0.87 \text{ Mpc}$ and $0.10 \text{ Mpc} < r_s < 0.28 \text{ Mpc}$ for components one and two, respectively.

We constructed three models for the gravitational lens system. The first model, which is unconstrained, consists of a single cluster and data from arcs A and B to make the fit. Our best fit for the single-component model predicts a third image which is slightly demagnified ($M_{F814W} \approx 23$) in a position where no object is detected in the HST exposures. We explored the field near its predicted position and identified an object around 6" to the north (R.A. = 20h56m22.5s, decl. = -04°37'37") which seems to be a deformed image of a galaxy. Figure 3 shows the isocontour map superimposed onto the optical image of this blurry object. We measured the magnitude of this counterimage candidate ($m_{F814W} = 23.37 \pm 0.05$), and we found that it is similar to the magnitude predicted by our model. Moreover, its F702W – F814W color turns out to be analogous to that of arc A, $0.46 \pm 0.07$ (see discussion in \S~5 below).

In the second model, we repeated the previous fit but including the third image candidate. With this third image the model is well constrained because we have eight constraints and only six parameters. The fit was very accurate for the position, but the total $\chi^2$ is affected by the shape. It is interesting to note that in this well-constrained model, there are two different solutions in the X-Y space. In Figure 4 we show the X-Y parameter space for this lens model constructed from all the trials coming out from the MCMC optimization procedure. We think that this result is due to the fact that the lens is a double cluster; we used the regions defined by these clumps to construct a third model, the double-component model, which accounts for a bimodal distribution of mass.

The results of these fits are summarized in Table 3. Column (1) identifies the model: "single" for models with one component and "double" for the model with two components; "AB" and "ABC" indicate that the model consists of arcs A and B, or these two arcs plus the third image C, respectively. Columns (2) and (3) show the position in arcseconds relative to the BCG. Columns (4)–(8) list the parameters associated with the NFW profile: ellipticity, position angle, concentration, scale radius, and $\sigma_r$, respectively. Columns (9) and (10) show the $\chi^2$ for the image positions only and for the whole fit, respectively.

4.3. Mass Profile

For the purpose of quantifying the mass distribution, we traced a projected mass map from the best-fit model, double ABC (Fig. 5). Figure 6 shows the corresponding isocontour map superimposed onto the F814W image of MS 2053. By integrating this two-dimensional map, we can get the total mass of the cluster. In
other words, we can determine the mass, making an azimuthal average of its mass in circles around the BCG.

In Figure 7 we compare the projected mass as a function of radius for the three components of the cluster: the two clusters (component 1 and component 2) and the individual galaxies. To calculate the contribution of the individual galaxies, we extrapolate the masses of the individual cluster components to their virial radii and add them. We can see that within the central 0.08 Mpc, the dominant component corresponds to the mass of the galaxies. However, at the arc radius \((r \approx 0.1 \text{ Mpc})\) component 1 has the main contribution to the lens mass. Figure 7 also shows the total mass of the cluster as the sum of these three components.

Figure 8 shows the sum of the projected mass of component 1 and component 2 (without including the galaxies) as a function of radius. We have extrapolated the NFW profiles up to 1 Mpc in order to compare with the virial mass estimate by Tran et al. (2005) and with a \(\beta\)-model derived by Vikhlinin et al. (2002). The vertical line shows the distance of arc A with respect to the BCG, and it indicates the limit below which the mass estimate from strong lensing is reliable. The mass determined using X-ray data depends on assumptions involving spherical symmetry and hydrostatic equilibrium. However, MS 2053 shows evidence of being two clusters enduring a merger process (Tran et al. 2005), and therefore we do not expect to find a good agreement between the \(\beta\)-model and the strong-lensing mass profile (see § 5).

5. DISCUSSION

We found that the colors for the ABC arc system turned out to be virtually the same, with arc B slightly bluer than but still coinciding with the color of arc A and image C within the errors. This result supports the assumption that images A, B, and C belong to the same source. Despite this fact, a serious concern is the lack of specularity; arcs A and B should be mirror images of each other, for which there is no evidence in the \(HST\) image. However, this would not be the first case of lack of specularity. For example, the cluster A383 shows a giant tangential arc with substructures and with no evidence of being formed by two mirror images; these substructures have different colors (Smith et al. 2001) and lie at the same redshift, \(z = 1.01\) (Sand et al. 2004). Smith et al. (2005) have shown that these arc features can be understood in terms of just one galaxy positioned over the tangential and radial caustic lines simultaneously, with three different portions of the same galaxy producing the arc substructure. We think that something similar may be occurring in MS 2053, i.e., the proximity of the

![Fig. 5.—Projected mass distribution derived from the double-component lens model. Note the elongation in the cluster mass, which traces the distribution of the galaxies.](image)

![Fig. 6.—\(HST\) WFPC2 F814W image of MS 2053. The contours correspond to projected surface densities of \(2.3 \times 10^{10}\), \(4.8 \times 10^{10}\), and \(7.3 \times 10^{10} \, M_\odot \text{arcsec}^{-2}\). The shape of the main component is elongated, as in Fig. 5. [See the electronic edition of the Journal for a color version of this figure.]](image)
source to the caustic line (see Fig. 9). The color distribution in galaxies often changes at different optical radii (e.g., bulge, disk, and halo populations). In the case of a spiral galaxy for which the disk (but not the bulge) is located over a caustic line, part of the disk is mapped onto a stretched arc such as arc B. This effect may explain the slight difference, if any, in arc B’s color, the difference in shape between arcs A and B, and the lack of specularity between the arcs.

For the rest of this section we assume that the three images A, B, and C correspond to the same source. This assumption, along with the evidence presented along the following lines, in the sense that the quality of the fits is better for the model with two cluster components, justifies our decision to focus our discussion on our model 3. The $\chi^2$ for the whole fit is <2.0 for the double

ABC model and >3.0 for the other two, favoring the bimodal model (see $\chi^2$ values in Table 3). In addition, the $\sigma_v$ values, 746 ± 114 and 392 ± 133 for component 1 and component 2, respectively, agree within the errors with the values reported by Tran et al. (2005); but the other models yield larger velocities (cf. Tables 1 and 3).

Hoekstra et al. (2002) suggested that the mass distribution in MS 2053 is elongated in the direction of the giant arc; this projected spatial distribution of the cluster was confirmed by Tran et al. (2005) and was attributed to a real elongation in the spreading of cluster members. The critical lines for our model, shown in Figure 9, reveal that the main component of the double cluster has a position angle of ≈80° and support the existence of a mass elongation. In addition, the mass map as well as the isocontour map of the projected mass distribution (see Figs. 5 and 6) reflect MS 2053’s elongated distribution. The offset in the position of the cluster components with respect to the BCG (see Table 3) also confirms the cluster asymmetry. Similar offsets have been found by other authors for different clusters (e.g., Oguri et al. 2004; Covone et al. 2006), and the effects can reach values of up to 120 kpc in unrelaxed clusters (Smith et al. 2005).

Several researchers have reported large concentration values by fitting NFW profiles to clusters of galaxies (e.g., Gavazzi et al. 2003; Kneib et al. 2003; Broadhurst et al. 2005). However, these results were obtained assuming spherically symmetric mass distributions, although the CDM model predicts triaxial halos as a consequence of the collisionless dark matter (Jing & Suto 2002). Using N-body simulations, Clowe et al. (2004) argue that the halo triaxiality affects the concentration parameter measurements obtained through gravitational lensing. Gavazzi (2005) concludes that a prolate halo aligned toward the line of sight is a natural explanation for the high concentration found in MS 2137−2353, but he also leaves open the possibility that the lensing mass estimates can be underestimated. Oguri et al. (2005) also investigate the importance of the halo triaxiality and demonstrate that this could cause a significant bias in estimating the virial mass and concentration parameter from the lensing information.
In contrast, the concentration parameters obtained for either the single- or the double-component clusters may be slightly lower (cf. Tables 1 and 3), although they agree, within the errors, with the expected values. According to Gavazzi (2005) a halo with axis ratio $\geq 1$ could explain the possible small discrepancies between our model and those inferred by Bullock et al. (2001) and Eke et al. (2001; $c \approx 4$ and $\sim 5$) for a cluster with $1.1 \times 10^{15}$ and $1.1 \times 10^{14} M_{\odot}$ in a $\Omega_m$-dominated universe.

A lower concentration changes the mass estimate. At large radii, the effect tends to vanish and the weak-lensing and true masses are the same (see Fig. 8), but inside the arcs ($r \lesssim 0.1$ Mpc) the effect is not negligible, producing an overestimate with respect to its actual X-ray mass. But the differences in the masses can also be explained as a result of the merging process. Rasia et al. (2006) studied a set of five galaxy clusters resolved at high resolution in a hydrodynamic simulation, examining the systematics affecting the X-ray mass estimates. They showed that for a cluster undergoing a merger, the assumption of hydrodynamic equilibrium led to the underestimation of mass by 20%, and that a $\beta$-model gave even more discrepant results with typical deviations of about 40%. In Figure 8 we compare the strong-lensing mass profile to the X-ray mass profile corrected for the 40% underestimation; we can see that with this correction, the X-ray mass estimate agrees with the mass obtained from the NFW components at the positions of the arcs, but for smaller radii the X-ray mass is significantly lower. It is important to note that we did not include the masses of the individual cluster components, and thus we are not considering in Figure 8 the total mass within the arcs. In addition, the extrapolation of the NFW mass inferred from gravitational lensing up to 1 Mpc from the BCG, well beyond the distance of the arcs, is a factor of $\sim 2$ lower than the mass shown in Table 1 obtained by Tran et al. (2005).

NFW ellipsoids have been successfully used for strong-lensing modeling in galaxy clusters (e.g., Comerford et al. 2006). However, the nonlinearity of the gravitational lensing effect, as well as its high dependence on the core densities, the asymmetries of the mass distribution, and the cluster neighborhoods, put limits to the use of analytical solutions in the case of nonrelaxed systems, such as MS 2053. In this sense, the work of Torri et al. (2004), who have computed the strong-lensing effect in a merging cluster with complex and irregular mass distribution, has shown that the lensing cross sections can grow by 1 order of magnitude and that during the merger, the shape of the critical and caustic lines change substantially. These topics have been discussed in a recent paper by Fedeli et al. (2006), who propose a fast method to calculate cross sections for complex and asymmetric mass distributions.

6. CONCLUSION

We have modeled the mass distribution of the cluster MS 2053 using the LENSTOOL ray-tracing code developed by Kneib (1993). The fits were performed considering both a single and a double NFW profile and 23 galaxy-scale mass components as perturbations to the cluster potential. We used elliptical singular isothermal spheres (ESIs) for the individual galaxies and set constraints to their parameters using observational data. We measured the arc positions and shapes, and the galaxies’ positions using the HST archive image, and calculated the NFW parameters for the single and double galaxy cluster models. Our main results can be summarized as follows:

1. The $X$-$Y$ parameter space of the single ABC model has a bimodal distribution that strongly suggests a double cluster undergoing a merger process. The quality of the fits also favor a model with a bimodal mass distribution.
2. The models consistently predict a third slightly demagnified counterimage for the AB arc system.
3. We found a candidate for this counterimage near the predicted position. This candidate shows a fuzzy object that could be the image of a distorted galaxy due to the effect of gravitational lensing. In addition, it has the same F702W $- F814W$ colors as arcs A and B.
4. Using the strong-lensing effect we confirm the asymmetry and the elongation in the mass distribution of MS 2053 reported by other authors (Hoekstra et al. 2002; Tran et al. 2005), and we estimate the total lensing mass within the arcs to be $4.7 \times 10^{13} M_{\odot}$. We find a significant discrepancy between the mass estimates from lensing and from X-ray measurements, which we attribute to the fact that the cluster is not in hydrodynamic equilibrium.
5. The concentration parameters obtained for either the single- or the double-component clusters may be slightly lower than the expected values. These possible small discrepancies can be explained by asymmetries in the mass distribution and the projection effects.

Much more work is necessary in order to understand the MS 2053 mass distribution. Future spectroscopic follow-up of arcs B and C will provide a test of our model. In addition, an extensive study of the internal velocity dispersion of the cluster members could be used to further improve the accuracy of the mass reconstruction.

This work was partially supported by DGAPA-UNAM grant IN113502. T. Verdugo also acknowledges scholarship support by CONACyT (register number 176553) and CONACyT grant 54799.

Some of the data presented in this paper were obtained from the Multimission Archive at the Space Telescope Science Institute (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-HST data is provided by the NASA Office of Space Science via grant NAG5-7584 and by other grants and contracts. The Dark Cosmology Centre is funded by the Danish National Research Foundation.

The authors thank Vladimir Avila for useful comments and the anonymous referee for invaluable remarks and suggestions. We also thank J. Benda for helping with proofreading.

APPENDIX
THE EINSTEIN ANGLE FOR A NFW PROFILE

Consider a spherical NFW density profile acting like a lens. The analytical solutions for this lens were given by Bartelmann (1996) and have been studied by different authors (Wright & Brainerd 2000; Golse & Kneib 2002; Meneghetti et al. 2003). The positions of the source and the image are related through the equation

$$u^2 = u^l - \nabla \varphi(u^l) = u^l - \alpha(u^l),$$

(A1)
where \( u^l \) and \( u^s \) are the angular position in the image and the source planes, respectively, \( \alpha \) is the deflection angle between the image and the source, and \( \phi \) is the two-dimensional lens potential. We introduce the dimensionless radial coordinate \( x = u^l/u^l_c \), where \( u^l = r_s/D_{OL} \), and \( D_{OL} \) is the angular diameter distance between the observer and the lens. In the case of an axially symmetric lens, the relations become simpler, as the position vector can be replaced by its norm.

The deflection angle then becomes (Golse & Kneib 2002)

\[
\alpha(x) = 4k_x u^l x g(x)\hat{e}_x, \quad (A2)
\]

where \( g(x) \) is a function related to the surface density inside the dimensionless radius \( x \) and is given by (Bartelmann 1996)

\[
g(x) = \begin{cases} 
\ln x + \frac{1}{\sqrt{1-x^2}} \arccosh \frac{1}{x} & \text{if } x < 1, \\
1 + \ln \frac{1}{2} & \text{if } x = 1, \\
\ln x + \frac{1}{\sqrt{x^2-1}} \arccos \frac{1}{x} & \text{if } x > 1,
\end{cases} \quad (A3)
\]

where \( k_x = r_s \rho_c \delta_c \Sigma_{\text{crit}}^{-1} \) is the lensing strength, which is an estimate of the convergence parameter (see the discussion in Oguri et al. 2004). The quantity \( \Sigma_{\text{crit}} = (c^2/4\pi G)(D_{OS}/D_{OL} D_{LS}) \) is the critical surface mass density for lensing, and \( D_{LS} \) and \( D_{OS} \) are the angular diameter distances between the lens and the source, and the observer and the source, respectively.

If we calculate the deflection angle at radius \( r_s \), we obtain \( x = 1 \), and equation (A2) can be expressed as

\[
\alpha(x) = \frac{16\pi G \rho_c \delta_c r_s^2 D_{LS}}{c^2 D_{OS}} \left( 1 + \ln \frac{1}{2} \right) \hat{e}_x. \quad (A4)
\]

Hence, the lens equation (A1) can be written as

\[
u^s = u^l \left[ 1 + \frac{u^l_{E,r_s}}{u^l} \right], \quad (A5)
\]

where \( u^l_{E,r_s} \) is the Einstein angle that for the case \( x = 1 \) is expressed by

\[
u^l_{E,r_s} = \frac{16\pi G \rho_c \delta_c r_s^2 D_{LS}}{c^2 D_{OS}} \left( 1 + \ln \frac{1}{2} \right). \quad (A6)
\]

For a circularly symmetric lens the Einstein radius is given by \( u^l_{E,r_s} = 4\pi \sigma^2 D_{LS}/c^2 D_{OS} \). Therefore, equation (A6) takes the form

\[
u^l_{E,r_s} = \frac{4\pi \sigma^2 D_{LS}}{c^2 D_{OS}}, \quad (A7)
\]

where \( \sigma^2 = 4(1 + \ln \frac{1}{2})G r_s^2 \rho_c \delta_c \). Therefore, \( \sigma \) represents the velocity at radius \( r_s \). We adopt this characteristic velocity in order to compare the velocity predicted by our models with the velocity dispersion measured in dynamical studies.

REFERENCES

Bartelmann, M. 1996, A&A, 313, 697
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Blandford, R. D., & Kochanek, C. S. 1987, ApJ, 321, 658
Broadhurst, T., et al. 2005, ApJ, 621, 53
Bullock, J. S., Kolatt, T. S., Sigad, Y., Somerville, R. S., Kravtsov, A. V., Klypin, A. A., Primack, J. R., & Dekel, A. 2001, MNRAS, 321, 599
Clowe, D., De Lucia, G., & King, L. 2004, MNRAS, 350, 1038
Comerford, J. M., Meneghetti, M., Bartelmann, M., & Schirmer, M. 2006, ApJ, 642, 39
Covone, G., Kneib, J.-P., Soucail, G., Richard, J., Jullo, E., & Ebeling, H. 2006, A&A, 456, 409
Della Ceca, R., Scaramella, R., Gioia, I. M., Rosati, P., Fiore, F., & Squires, G. 2000, A&A, 353, 498
Dyer, C. C., & Roeder, R. C. 1976, Nature, 260, 764
Eke, V. R., Cole, S., & Frenk, C. S. 1996, MNRAS, 282, 263
Eke, V. R., Navarro, J. F., & Frenk, C. S. 1998, ApJ, 503, 569
Eke, V. R., Navarro, J. F., & Steinmetz, M. 2001, ApJ, 554, 114
Faber, S. M., & Jackson, R. E. 1976, ApJ, 204, 668
Fedeli, C., Meneghetti, M., Bartelmann, M., Kolag, K., & Moscardini, L. 2006, A&A, 447, 419
Gavazzi, R., 2005, A&A, 443, 793
Gavazzi, R., Fort, B., Mellier, Y., Pelló, R., & Daniel-Fort, M. 2003, A&A, 403, 11
Ghigna, S., Moore, B., Governato, F., Lake, G., Quinn, T., & Stadel, J. 2000, ApJ, 544, 616
Gladders, M. D., Hoekstra, H., Yee, H. K. C., Hall, P. B., & Barrientos, L. F. 2003, ApJ, 593, 48
Golse, G., & Kneib, J.-P. 2002, A&A, 390, 821
Hoekstra, H., Franx, M., Kuijken, K., & van Dokkum, P. G. 2002, MNRAS, 333, 911
Jing, Y. P., & Suto, Y. 2002, ApJ, 574, 538
Jullo, E., Kneib, J.-P., Limousin, M., Eliasdóttir, Á., Marshall, P., & Verdugo, T. 2007, preprint (0706.0048)
Karoji, H., & Nottale, L. 1976, Nature, 259, 31
Keeton, C. R., Kochanek, C. S., &Falco, E. E. 1998, ApJ, 509, 561
Kneib, J.-P. 1993, Ph.D. thesis, Univ. Paul Sabatier
Kneib, J.-P., Ellis, R. S., Smail, I., Couch, W. J., & Sharples, R. M. 1996, ApJ, 471, 643
Kneib, J.-P., et al. 2003, ApJ, 598, 804
Koopmans, L. V. E., Bolton, A. S., Burles, S., & Moustakas, L. A. 2006, ApJ, 649, 599
Lacey, C., & Cole, S. 1993, MNRAS, 262, 627

No. 2, 2007 STRONG GRAVITATIONAL LENSING OF MS 2053.7–0449 711
