



HAL
open science

Detection of a dark substructure through gravitational imaging

S. Vegetti, L. V. E. Koopmans, A. Bolton, T. Treu, R. Gavazzi

► **To cite this version:**

S. Vegetti, L. V. E. Koopmans, A. Bolton, T. Treu, R. Gavazzi. Detection of a dark substructure through gravitational imaging. *Monthly Notices of the Royal Astronomical Society*, 2010, 408, pp.1969-1981. 10.1111/j.1365-2966.2010.16865.x . hal-03646011

HAL Id: hal-03646011

<https://hal.science/hal-03646011>

Submitted on 4 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Detection of a dark substructure through gravitational imaging

S. Vegetti,^{1*} L. V. E. Koopmans,¹ A. Bolton,² T. Treu³ and R. Gavazzi⁴

¹*Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, the Netherlands*

²*Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822-1897, USA*

³*Department of Physics, University of California, Santa Barbara, CA 93101, USA*

⁴*Institut d'Astrophysique de Paris, CNRS, UMR 7095, Universite Pierre et Marie Curie, 98 bis Bd Arago, 75014 Paris, France*

Accepted 2010 April 16. Received 2010 March 24; in original form 2009 October 13

ABSTRACT

We report the detection of a dark substructure – undetected in the *Hubble Space Telescope HST* ACS F814W image – in the gravitational lens galaxy SDSSJ0946+1006 (the ‘double Einstein ring’), through direct gravitational imaging. The detection of a small mass concentration in the surface density maps, at 4.3 kpc from the galaxy centre, has a strong statistical significance. We confirm this detection by modelling the substructure with a tidally truncated pseudo-Jaffe density profile; in that case the substructure mass is $M_{\text{sub}} = (3.51 \pm 0.15) \times 10^9 M_{\odot}$, precisely where also the surface density map shows a strong convergence peak (Bayes factor $\Delta \log \mathcal{E} = -128.0$; equivalent to a $\sim 16\sigma$ detection). The result is robust under substantial changes in the model. We set a lower limit of $(M/L)_{V,\odot} \gtrsim 120 M_{\odot}/L_{V,\odot}$ (3σ) inside a sphere of 0.3 kpc centred on the substructure ($r_{\text{tidal}} = 1.1$ kpc). The mass and luminosity limit of this substructure are consistent with Local Group results if the substructure had a virial mass of $\sim 10^{10} M_{\odot}$ before accretion and formed at $z \gtrsim 10$. Our detection implies a projected dark matter mass fraction in substructure at the radius of the inner Einstein ring of $f = 2.15^{+2.05}_{-1.25}$ per cent [68 per cent confidence level (CL)] in the mass range $4 \times 10^6 - 4 \times 10^9 M_{\odot}$, assuming $\alpha = 1.9 \pm 0.1$ (with $dN/dm \propto m^{-\alpha}$). Assuming a flat prior on α , between 1.0 and 3.0, increases this to $f = 2.56^{+3.26}_{-1.50}$ per cent (68 per cent CL). The likelihood ratio is ~ 0.5 between these fractions and that from simulations ($f_{N\text{-body}} \approx 0.003$). Hence the inferred dark matter mass fraction in substructure, admittedly based on a single-lens system, is large but still consistent with predictions.

Key words: gravitational lensing; strong – galaxies; structure.

1 INTRODUCTION

In the process of building a coherent picture of galaxy formation and evolution, early-type galaxies play a crucial role. Often unfairly referred to as *red and dead* objects, many aspects about their structure and formation are still unknown. What is the origin of the tight empirical relations between their global properties (Djorgovski & Davis 1987; Dressler et al. 1987; Bower, Lucey & Ellis 1992; Guzman et al. 1992; Bender, Burstein & Faber 1993; Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000)? How do massive early-type galaxies assemble? What is the fraction of mass substructure populating the haloes of early-type galaxies and is this in agreement with the cold dark matter (CDM) paradigm (Kauffmann, White & Guiderdoni 1993; Klypin et al. 1999; Moore et al. 1999, 2001; Macciò & Miranda 2006; Diemand et al. 2008; Springel et al. 2008; Xu et al. 2009)?

Gravitational lensing, especially in combination with other techniques, provides an invaluable and sometimes unique insight in answering these questions (e.g. Treu & Koopmans 2004; Rusin & Kochanek 2005; Koopmans et al. 2009, and references therein).

At the level of small mass structure lensing stands out as a unique investigative method; different aspects of the lensed images can be analysed to extract information about the clumpy component of galactic haloes. Flux ratio anomalies, astrometric perturbations and time delays, in multiple images of lensed quasars, can all be related to substructure at scales smaller than the images separation (Mao & Schneider 1998; Bradač et al. 2002; Chiba 2002; Dalal & Kochanek 2002; Metcalf & Zhao 2002; Keeton, Gaudi & Petters 2003; Bradač et al. 2004; Kochanek & Dalal 2004; Keeton, Gaudi & Petters 2005; Chen et al. 2007; McKean et al. 2007; More et al. 2009).

As described by Koopmans (2005) and Vegetti & Koopmans (2009a), the information contained in multiple images and Einstein rings of extended sources, can also be used. While the former three approaches only provide a statistical measure of the lens clumpiness, the latter allows one to identify and quantify of each single

*E-mail: vegetti@astro.rug.nl

substructure, measuring for each of them the mass and the position on the lens plane. Both approaches are however complementary in that the former is more sensitive to low-mass perturbations, which are potentially present in large numbers, whereas the latter is sensitive to the rarer larger scale perturbations.

The method of direct gravitational imaging of the lens potential – shortly described in the following section – represents an objective approach to detect dark and luminous substructures in individual lens systems and allows on to statistically constrain the fraction of galactic satellites in early-type galaxies. Extensively described and tested in Vegetti & Koopmans (2009a), the procedure is here applied to the study of the double ring SDSSJ0946+1006 (Gavazzi et al. 2008) from the sample of the Sloan Lens ACS Survey (SLACS), yielding the first detection of a dwarf satellite through its gravitational effect only, beyond the local Universe.

The layout of the paper is as follows. In Section 2 we provide a general description of the method. In Section 3 we introduce the analysed data and in Section 4 we discuss the results of the modelling under the assumption of a smooth lens potential. In Section 5 we describe the detection of the substructure. In Section 6 we present the error analysis and the model ranking. In Section 7 we discuss implication for the CDM paradigm and in Section 8 we conclude. Throughout the paper we assume the following cosmological parameters: $H_0 = 73 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_m = 0.25$ and $\Omega_\Lambda = 0.75$.

2 THE METHOD

In this section we provide a short introduction to the lens modelling method. The main idea behind the method of ‘gravitational imaging’ is that effects related to the presence of dwarf satellites and/or CDM substructures in a lens galaxy can be modelled as local perturbations of the lens potential and that the total potential can be described as the sum of a smooth parametric component with linear corrections defined on a grid. We refer to Vegetti & Koopmans (2009a) for a more complete discussion.

2.1 Source and potential reconstruction

As shown in Blandford, Surpi & Kundić (2001), Koopmans (2005), Suyu et al. (2006) and Vegetti & Koopmans (2009a), it is possible to express the relation between perturbations in the lensed data (δd ; i.e. perturbations of the surface brightness distribution of the lensed images), the unknown source surface brightness distribution (s) and perturbations in the lens potential ($\delta\psi$) as a set of linear equations $\delta d = -\nabla s \cdot \nabla \delta\psi$. Through the Poisson equation $\delta\psi$ can be turned into a relation with the convergence correction $\delta\kappa = \nabla^2 \delta\psi / 2$.

For a fixed form of the lensing potential and regularization, the inversion of these equations leads to the simultaneous reconstruction of the source and a potential correction. The source grid is defined by a Delaunay tessellation which automatically concentrates the computational effort in high-magnification regions while keeping the number of degrees of freedom constant, which is critical in assessing the Bayesian posterior probability and evidence for the model (see Vegetti & Koopmans 2009a). The procedure is embedded in the framework of Bayesian statistics which allows us to determine the best set of non-linear parameters for a given potential and the linear parameters of the source to objectively set the level of regularization and to compare different model families (MacKay 1992, 2003; Brewer & Lewis 2006; Suyu et al. 2006). Specifically, for a particular lens system we wish to objectively assess whether it can

be reproduced with a smooth potential or whether mass structure on smaller scales has to be included in the model.

The modelling is performed via a four-step procedure. (i) We start by choosing a form for the parametric smooth lens density profile, generally an elliptical power law, and we determine the non-linear parameters and level of source regularization that maximize the Bayesian evidence, through a non-linear optimization scheme. (ii) In the case that this model is too simple and significant image residuals are left, we allow for grid-based potential corrections. This leads to the initial detection and localization of possible substructures. (iii) The substructure masses and positions are then more precisely quantified by assuming a tidally truncated pseudo-Jaffe (PJ) profile (Dalal & Kochanek 2002) and by simultaneously optimizing for the main lens galaxy and substructure parameters, i.e. its mass M_{sub} and position on the lens plane ($x_{\text{sub}}, y_{\text{sub}}$). (iv) Finally the two models, i.e. the single power law (PL) and the power law plus PJ substructures (PL + PJ), are compared through their total marginalized Bayesian evidences (\mathcal{E}), which represent the (conditional) probabilities of the data marginalized over all variable model parameters.

2.2 Detection threshold of mass substructure

The method has a mass-detection threshold to substructure that depends on the signal-to-noise ratio and spatial resolution of the lensed images; for typical *Hubble Space Telescope* (*HST*) (e.g. SLACS) data quality the mass-detection threshold for a substructure located on the Einstein ring and with a PJ density profile is of the order of a few $\times 10^8 M_\odot$ and quickly increases with the distance from the lensed images (see Vegetti & Koopmans 2009a) because of the decrease in the image surface brightness and local magnification.

Despite having been developed with the specific task of identifying and constraining the fraction of substructure in lens galaxies, this technique can also be used to model complex lens potentials, which are relatively smooth, but do not have the simple symmetries that are often assumed in mass models (e.g. elliptical power-law density profiles). As shown in Barnabè et al. (2009), we can also reconstruct the lensed images and the relative sources down to the noise level, even for systems that are highly asymmetric and strongly depart from a power-law density profile. The grid-based potential correction is able to correct the inexact initial choice of the lens potential model and recover existing asymmetries in the mass distribution.

In the rest of this paper, we use this method to analyse the double Einstein ring system SLACS SDSSJ0946+1006 and search for deviations from a smooth power-law elliptical mass model.

3 THE DATA

In this section we present a brief overview of the double Einstein ring lens system SLACS SDSSJ0946+1006 (see Fig. 1). We refer to Gavazzi et al. (2008) for a more detailed description.

SLACS selects gravitational lens candidates from the Sloan Digital Sky Survey (SDSS) spectroscopic data base on the basis of multiple emission lines in the spectrum at redshifts larger than that of the lower redshift target galaxies (Bolton et al. 2006). The system was selected by the presence of multiple emission lines at $z_{s1} = 0.609$ in the spectrum of a lensing galaxy at $z_l = 0.222$. Subsequently confirmed as a strong lens with ACS on board the *HST*, the system shows a very peculiar structure in the lensed images which are composed of two concentric partial rings, at radii of 1.43 ± 0.01 and 2.07 ± 0.02 arcsec, respectively, from the centre of the lens galaxy. This particular configuration is related to the presence

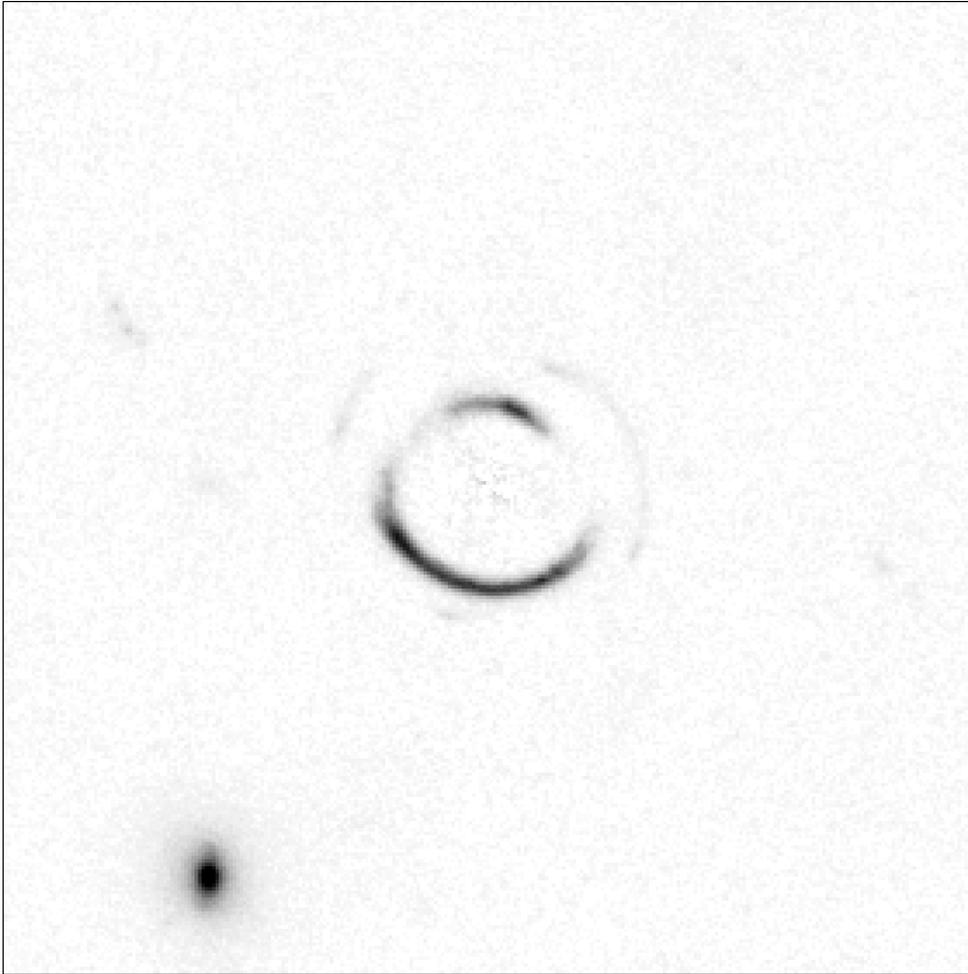


Figure 1. The image of the lens system SDSSJ0946+1006, obtained with *HST* ACS through the filter *F814W*, after subtraction of the lens surface brightness distribution.

of two sources at different redshifts which are being lensed by the same foreground galaxy (see Gavazzi et al. 2008 for the a priori probability of this event in a survey such as SLACS); the nearest source is lensed into the inner ring (Ring 1), while the second one, further away along the optical axis, is lensed into the outermost ring (Ring 2). Ring 1 with a *F814W* magnitude $m_1 = 19.784 \pm 0.006$ is one of the brightest lensed sources in the SLACS sample, while Ring 2 with $m_2 = 23.68 \pm 0.09$ is 36 times fainter. Ring 2 is not observed in the SDSS spectrum, and an upper limit to its redshift $z_{s2} < 6.9$ was set on the basis of ACS imaging. As inferred by Gavazzi et al. (2008) the lens galaxy has a projected dark matter mass fraction inside the effective radius that is about twice the average value of the SLACS lenses (Koopmans et al. 2006; Gavazzi et al. 2007), i.e. $f_{\text{DM}} \approx 73 \pm 9$ per cent, corresponding to a project dark matter mass approximately equal to $M_{\text{DM}} (< R_{\text{eff}}) \approx 3.58 \times 10^{11} h_{70}^{-1} M_{\odot}$.

The high dark matter mass fraction makes this system particularly interesting for CDM substructure studies. If the framework of galaxy formation given by N -body simulations is correct (substructure mass function slope $\alpha_{N\text{-body}} = 1.90$ and projected dark matter fraction in substructure $f_{N\text{-body}} = 0.3$ per cent), we would expect, within an annulus of 0.6 arcsec centred on the Einstein radius, on average $\mu = 6.46 \pm 0.95$ substructures (Vegetti & Koopmans 2009b) with masses between 4×10^6 and $4 \times 10^9 M_{\odot}$ (Diemand, Kuhlen & Madau 2007a,b; Diemand et al. 2008). Whereas these have typical masses a few times that of the lower limit on this range, the probability

of finding a mass substructure above $\gtrsim 10^8 M_{\odot}$ is certainly non-negligible.

4 SMOOTH MASS MODELS

In this section we describe the details and the results of our analysis. Because of the very low surface brightness, and of the low signal-to-noise ratio of the images associated with Ring 2, we limit our study to a tight annulus around Ring 1, in which Ring 2 has been fully excised. This does not affect the lens potential reconstruction which is almost solely constrained by the detailed information given by the high surface brightness distribution of Ring 1. Our potential reconstruction therefore *only* probes the region around the inner ring. The choice of reconstructing the potential (ψ) inside a limited field of view (e.g. mask), rather than the surface mass density (κ), ensures that the potential and the resulting convergence reconstructions are both unbiased (see Koopmans 2005, for a detailed discussion about this subtle point).

4.1 Image reconstruction

At the first level of reconstruction all potential corrections are kept at zero. We start by assuming that to first order the lens is well approximated by a simple smooth elliptical power-law density profile (Barkana 1998) with a convergence (surface density in terms of

critical density Σ_c)

$$k(r) = \frac{b}{2\sqrt{q}r^{\gamma-1}}, \quad (1)$$

where $r = \sqrt{x^2 + y^2/q^2}$. The non-linear parameters describing the lens are the lens strength b , the position angle θ , the flattening q , the centre coordinates x_0 , the density slope γ , the shear strength Γ_{sh} and the shear angle θ_{sh} . We do not optimize for the mass centroid, but centre on the peak of the surface brightness distribution, as precisely determined from the *HST* image. We show in Section 6.4 that this assumption does not alter the main results of the paper, but reduces our substantial computational load.

As described in more detail in Vegetti & Koopmans (2009a), the source grid is constructed from a (sub)sample of pixels in the image plane which are cast back to the source plane using the lens equation. The number of grid points can be objectively set by comparing their Bayesian evidence. In this particular case, we find that using all the image points (e.g. 81×81 pixels) is the most appropriate choice. On the image plane the pixel scale is constant and equal to 0.05 arcsec pixel⁻¹, while on the source plane the Delaunay triangle scale is adaptive and depends on the local lensing magnification. We adopt

an adaptive curvature regularization, weighting the regularization penalty by the inverse of the image signal-to-noise ratio. We find that this significantly improves the modelling of sharp high dynamic range features in the lensed images, where in general all other forms of regularization (e.g. gradient or unweighted curvature) falter and give much lower evidence values.

We use the results obtained by Gavazzi et al. (2008), for a single-lens plane, as starting point η_0 and then optimize for the potential parameters and the level of the source regularization. The resulting source and image reconstruction are presented in Fig. 2. In Table 1 the recovered lens parameters and level of source regularization $\{\eta_b, \lambda_{s,b}\}$ are listed. The recovered parameters for the smooth mass component of the lens potential are somewhat different from the results in Gavazzi et al. (2008), which we attribute to the fact that Gavazzi et al. (2008) make use of both Einstein rings and match conjugate points instead of the full surface brightness distribution. Some notable results for the smooth mass model are a density slope $\gamma \approx 2.20$ and a mass axial ratio of $q = 0.96$, indicating that the galaxy is very close to an isothermal sphere mass model, although it has a slightly steeper density profile. The quantity cannot be compared directly with the slope measured by Gavazzi et al. (2008) because

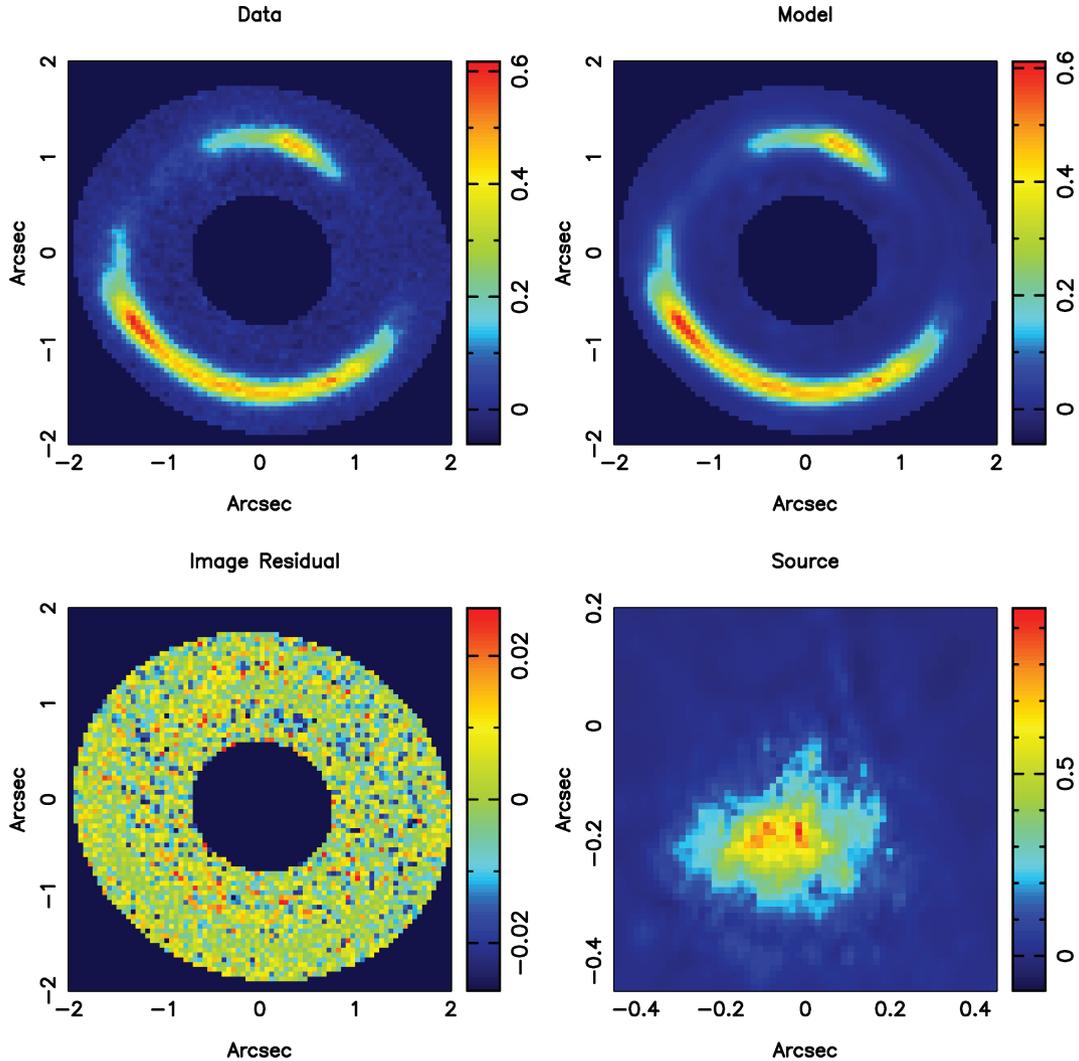


Figure 2. Results of the lens and source reconstruction under the hypothesis of a smooth potential. The top right-hand panel shows the original lens data, while the top left-hand one shows the final reconstruction. On the second row the image residuals (left-hand panel) and the source reconstruction (right-hand panel) are shown.

Table 1. Parameters of the mass model distribution for the lens SDSSJ0946+1006. For each parameter we report the best recovered value and the relative likelihood for a smooth model (PL) in column (2); for a smooth over-regularized smooth model in column (3); for a perturbed model (PL + PJ) in column (4); for a smooth and perturbed model (PL + PJ) with rotated PSF, respectively, in columns (5) and (6); a smooth and perturbed model (PL + PJ) for different galaxy subtraction, respectively, in columns (7) and (8) and a smooth and perturbed model (PL + PJ) with varying coordinates for the lens centre, respectively, in columns (9) and (10). We note that the models in columns (7) and (8) use a different (also rotated) data set, and the evidence values, position angles and positions can therefore not be directly compared.

	(PL) ₀	PL _{0,over}	(PL + PJ) ₀	PL _{psf90}	(PL + PJ) _{psf90}	PL _{subt}	(PL + PJ) _{subt}	PL _{cntr}	(PL + PJ) _{cntr}
b	1.329	1.329	1.328	1.329	1.328	1.280	1.272	1.329	1.328
θ (°)	65.95	65.80	69.26	64.97	71.04	−60.99	−60.96	63.45	70.15
q	0.961	0.961	0.962	0.962	0.963	0.982	0.982	0.962	0.962
x_c (arcsec)								−0.006	−0.005
y_c (arcsec)								0.014	0.012
γ	2.196	2.199	2.198	2.194	2.200	2.282	2.292	2.209	2.210
Γ_{sh}	0.081	0.081	0.086	0.080	0.087	−0.092	−0.097	0.083	0.082
θ_{sh} (°)	−20.83	−20.65	−22.32	−20.63	−22.12	−39.83	−40.58	−21.13	−21.00
$\log(\lambda_s)$	1.152	2.028	2.028	1.059	1.988	0.036	0.052	1.414	1.151
$m_{sub}(10^{10} M_\odot)$			0.323		0.333		0.342		0.325
x_{sub} (arcsec)			−0.686		−0.682		−1.286		−0.683
y_{sub} (arcsec)			0.989		0.996		−0.391		0.993
$\log \mathcal{L}$	20 350.97	20 328.11	20 511.14	20 358.49	20 525.32	61 520.63	61 674.63	20 418.74	20 578.90

we are measuring the slope at the location of the inner ring, while Gavazzi et al. (2008) is measuring the average slope in-between the rings.

4.2 Image residuals after reconstruction

In Fig. 2, we clearly see remaining image residuals above the noise level, in particular near the uppermost arc feature. The source appears to be a normal, although not completely symmetric, galaxy. Structure in the source (e.g. brightness peaks and the faint tail-like feature to the upper right-hand side of the source) can also be one-to-one related to structure in the arcs. This provides strong confidence to the overall reconstruction of the system, as being remarkably accurate despite its complexity. The source still shows significant structure on small scales, which is due to a preferred low level of regularization, when optimizing for the Bayesian evidence. (Note that at this level the evidence is simply the posterior probability of the free parameters, including the source regularization.)

The image residuals can be related either to different aspects of the reconstruction procedure, for example the modelling of the point spread function (PSF), the choice of the simply parametrized model for the lens potential, the number and scale of the image pixels, the lens galaxy subtraction, or to features in the galaxy brightness profile. To test whether these residuals are related to the presence of substructures, however, we now first proceed by considering a more general model in which we allow for very general potential corrections (see above). We discuss the effects of systematic errors in a later section, but stress that none of the above systematic errors is expected to mimic localized lensing features.

5 THE DETECTION OF MASS SUBSTRUCTURE

From the ‘Occam’s razor’ point of view, it is more probable that uncorrelated structures in the lensed images are related to local small-scale perturbations in the lens potential, rather than features in the source distribution itself (Koopmans 2005; Vegetti & Koopmans 2009a). It is, therefore, possible to describe galaxy substructure or satellites as linear local perturbations to an overall smooth parametric potential and separate them from changes in the

surface brightness distribution due to the source model (Koopmans 2005). Given that the remaining image residuals are small, we can assume that the values for the lens parameters recovered in the previous section are sufficiently close to the real smooth component of the lens potential such that our linearized reconstruction of the source and the grid-based lens potential corrections are fully justified, as discussed in Section 2.1.

5.1 Grid-based substructure modelling

The potential corrections are defined on a regular Cartesian grid with 21×21 pixels. Both the source and the potential have a curvature regularization (in the case of the source inversely weighted by the local image signal-to-noise ratio) and are initially over-regularized in order to keep the potential corrections in the linear regime, where the formalism of the method is valid. The potential corrections are repeated (adding the previous correction to the current total potential) until convergence is reached in the evidence value. Results for this linear reconstruction are presented in Fig. 3.

The potential correction and convergence show a clear signature of a concentrated mass overdensity (i.e. a substructure) observed around the position (−0.5, 1.0 arcsec). We have tested the effect of the potential-correction regularization on the stability of the reconstruction by using three different levels of regularization, in particular $\lambda_{\delta\psi} = 10^7, 10^8$ and 10^9 . As expected, the convergence correction becomes smoother as the regularization increases; however, the feature near (−0.5, 1.0 arcsec) in the convergence-correction map remains clearly visible in each reconstruction, together with a minor mass gradient from the lower right-hand side of the ring to the upper left-hand side. This gradient is associated with the presence of the substructure itself (curvature regularization of the potential implies $\langle \kappa \rangle = 0$ in the annulus and thus neither the total mass nor the average convergence gradient changes in the annulus; this is an advantage of the method). For the nearly under-regularized case with $\lambda_{\delta\psi} = 10^7$, the source is slightly twisted and the reconstruction becomes more noisy. This suggests that the potential allows for a minor amount of shear (we note that shear has $\kappa = 0$ everywhere and that is not penalized by a curvature form of regularization), but that the substructure, although noisier, is still present near the same

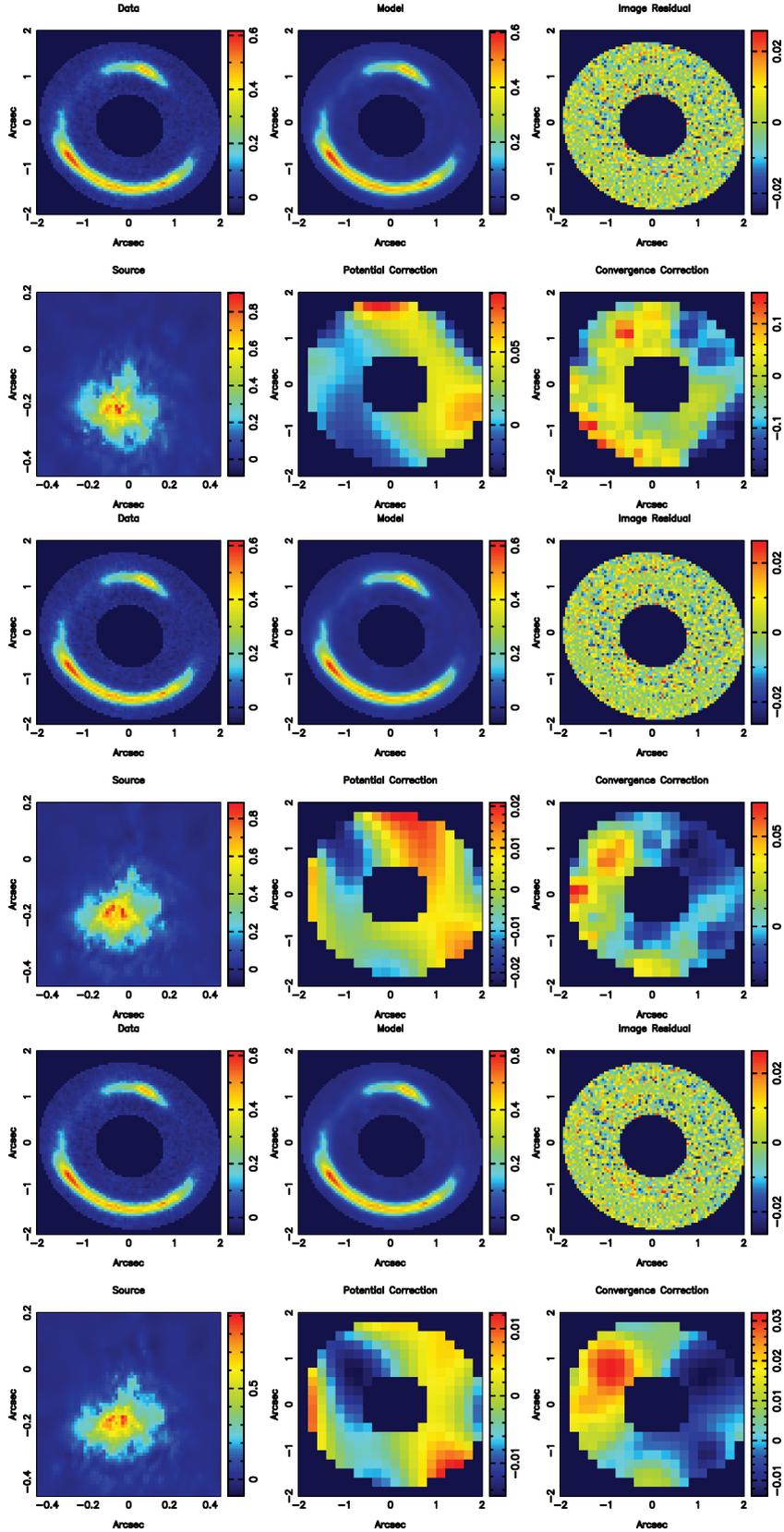


Figure 3. Results for the pixelized reconstruction of the source and lens potential corrections for three different value of the potential corrections regularization $\lambda_{\delta\psi} = 10^7$ (top panels), $\lambda_{\delta\psi} = 10^8$ (middle panels) and $\lambda_{\delta\psi} = 10^9$ (low panels). For each panel on the first row from left to right the original lens data, the final reconstruction, and the image residuals are shown. On the second row from left to right the source reconstruction, the potential correction and the potential correction convergence are shown.

position as in the other reconstruction. We therefore believe, given the data, that this feature is genuine. This statement, however, requires quantification. This is difficult at this moment, based on the grid-based method, but can be done if the substructure is modelled through a simply parametrized mass component.

5.2 Parameterized substructure modelling

We quantify the mass of this substructure by assuming an analytic PL + substructure model. We assume the structure to have tidally truncated PJ profile (Dalal & Kochanek 2002) with a convergence

$$k(r) = \frac{b_{\text{sub}}}{2} \left[r^{-1} - (r^2 + r_t^2)^{-1/2} \right], \quad (2)$$

where r_t is the substructure tidal radius and b_{sub} its lens strength; both are related to the main galaxy lens strength b and to M_{sub} by $r_t = \sqrt{b_{\text{sub}} b}$ and $M_{\text{sub}} = \pi r_t b_{\text{sub}} \Sigma_c$. Combining the last two relations leaves its total mass and position on the lens plane as free parameters for the substructure model. Fig. 4(b) shows the final result of the Bayesian evidence maximization for both the main lens and substructure parameters.

Remarkably, this procedure requires a substructure *right at* the position of the convergence overdensity found in the grid-based reconstruction. In terms of likelihood, the PL + PJ model is favoured with a $|\Delta \log \mathcal{L}| = +161.0$ over the PL model (i.e. roughly comparable to a $\Delta \chi^2 \sim 2 \Delta \log \mathcal{L}$ improvement). One might note that the two models still seem to have similar levels of image residuals. This can be attributed to a significant difference in the source regularization. The smooth model, in order to fit the data, has to allow for more freedom to the source and has a lower level of source regularization. Hence, part of the potential structure is ‘absorbed’ in the source brightness distribution. The two models have in fact an absolute difference between their χ^2 values of $|\Delta \chi^2| = 120$, while the difference in the regularization term of the evidence is

1120. Because the change in the rms of the image residuals is related to the difference in the χ^2 per number of degrees of freedom, the two models do not seem to differ significantly in their residual maps. However, this does not take into account source structure, which is smoother when a substructure is included in the model. The latter model thus yields a substantially higher evidence as will be shown in the next section. To better assess the level at which the source regularization contributes to the image residual level, we run a non-linear optimization for the smooth model while keeping fixed the regularization constant at the level of the best PL + PJ model; we call this over-regularized model PL_{0,over} (see Fig. 4a). The likelihood difference between a perturbed model and a smooth one is now further increased to $|\Delta \log \mathcal{L}| = +183.0$. Hence, indeed there is some covariance between the potential and source models. However, no smooth potential model does as well as models that include the PJ substructure, near the position found in the grid-based reconstruction. We are therefore convinced, based on the likelihood ratio, that a PL + PJ model provides a much more probable explanation of the data than a PL model with a more structured source model.

Despite the difference in $\mathcal{L}(\eta)$, at this stage of the modelling, it is not possible to state whether the detection is statistically significant because the effective number of degrees of freedom have not yet been accounted for. As shown in the next section, a nested-sampling exploration and marginalization of the posterior probability density can be used to clarify this point and provide the Bayesian evidence values for the PL and PL + PJ models that can objectively be compared.

We note that the exclusion of the outer Einstein ring (Ring 2) from our analysis does not have influence on the substructure detection nor its quantification. First, our potential reconstruction method is based only on *local* quantities, e.g. the gradient of the local potential and the local intensity difference between the lensed images and the source model. Exclusion of data does therefore not bias

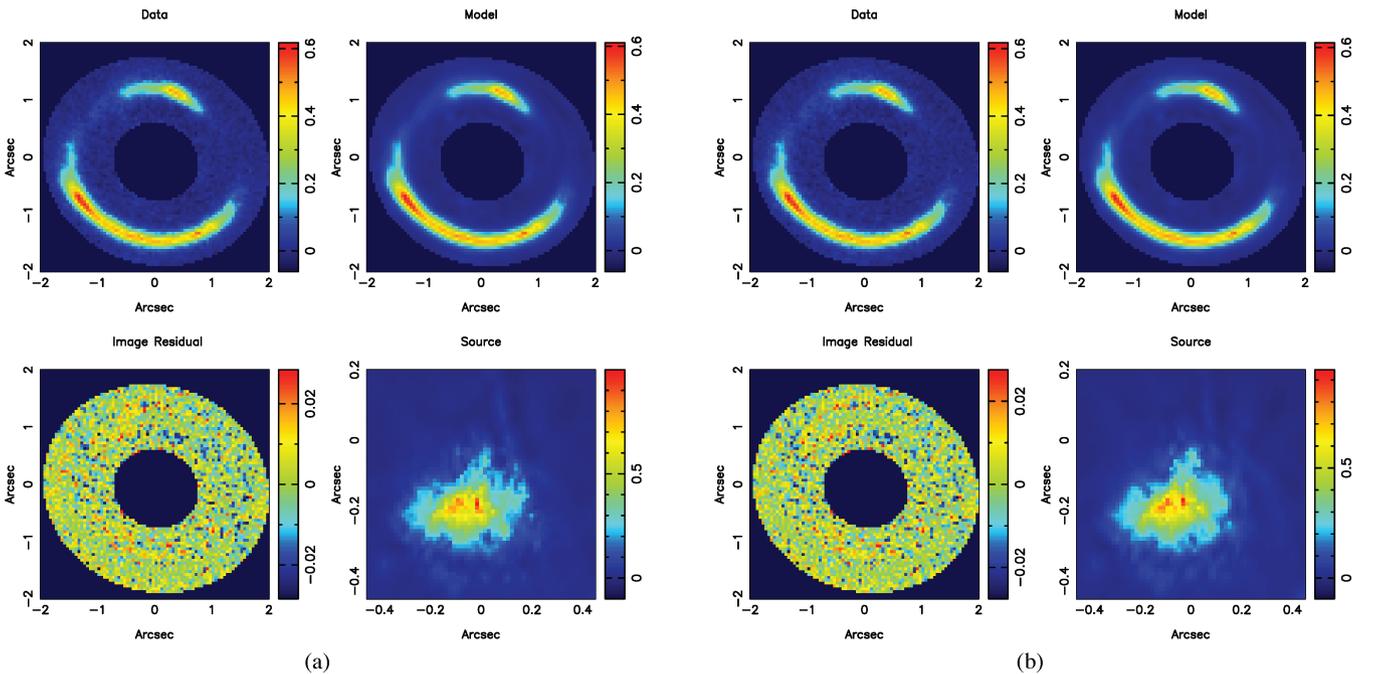


Figure 4. Shown are the results for the over-regularised PL model (left panel) and for the PL+PJ model (right panel). The residuals for the PL model are still subtle but have become more pronounced and that the source model also still has more structure, despite over-regularization. Some of these residuals are reduced by lowering the source regularisation (see text), but the evidence difference between the two models remains large.

the resulting local potential correction (see Koopmans 2005 for a discussion on this important point), although of course it could reduce its significance if those data are near the substructure position in the image plane and substantially contributes to the information we can gain on it. Secondly, low-mass substructures have a rather localized effect on the surface brightness of the lensed images. Ring 2 is localized about 1 arcsec away from the substructure in the radial direction, whereas the effect of the substructure is mostly along the tangentially stretched arc. Hence, whereas the inclusion of the second ring could lead to more stringent constraints on the global model, it has negligible influence on our ability to detect and quantify the substructure.

As a final remark, we note that there is the possibility that the substructure is not associated with the lens galaxy, but lies along the line of sight. Despite the fact that lensing effects by substructure in the lens galaxy are most efficient and decrease in efficiency further from the lens, the cumulative number density of substructures along the line of sight could be significant. Chen, Kravtsov & Keeton (2003), however, showed that substructures along the line of sight are probably responsible for a contamination of only up to 20–30 per cent. This result was based on comparatively low-resolution numerical simulations where only a few substructures were located in the inner 10 per cent of the lens virial radius. In current high-resolution simulations this fraction is shown to be substantially larger, whereas the overall substructure mass fraction that dominates the line-of-sight contamination has not changed. We therefore expect their contamination rate to be a conservative upper limit and currently well within the Poisson errors derived on the substructure mass fraction (see below).

6 ERROR ANALYSIS AND MODEL RANKING

In this section we present the statistical analysis on the model parameters and the total marginalized evidence computation for model comparison. We are interested to test whether the lensed images are compatible with a single smooth potential or whether the data indeed objectively require the presence of a mass substructure. We consider therefore two models, one defined by a smooth lens with a power-law density profile and the other containing an additional mass substructure. In general, two models can only be objectively and quantitatively compared in terms of the total marginalized Bayesian evidence and the Bayes factor, $\Delta \log \mathcal{E} \equiv \log \mathcal{E}_0 - \log \mathcal{E}_1$, which expresses their relative probability given a specific data set.

Heuristically the Bayesian evidence (\mathcal{E}) can be compared to the classic reduced χ^2 (i.e. per degree of freedom), but without assumptions about Gaussianity of the posterior probability distribution function about lack of covariance between parameters (which could reduce the effective number of degrees of freedom).

6.1 Prior probabilities

Prior to the data taking, little is known about the non-linear parameters describing the lens potential model. A natural choice is therefore a uniform prior probability. We centre this prior on the best smooth values $\eta_{b,i}$ as recovered in Section 4.1, although the choice of prior range is not very relevant as long as the likelihood is sharply peaked inside the prior volume:

$$P(\eta_i) = \begin{cases} \text{constant} & \text{for } |\eta_{b,i} - \eta_i| \leq \delta\eta_i, \\ 0 & \text{for } |\eta_{b,i} - \eta_i| > \delta\eta_i. \end{cases} \quad (3)$$

Hence, the sizes of the intervals are taken in such a way that they enclose the bulk of the evidence (i.e. the likelihood multiplied by prior

volume). Exactly identical priors for η are used for both the smooth and perturbed model. Also in the latter case, the prior is centred on the mass model parameters of the smooth model. This ensures that we are comparing their evidences in a proper manner. The regularization constant has a prior probability which is logarithmically flat in a symmetric interval around $\lambda_{s,b}$. The mass substructure is assumed to have a PJ density profile and a mass with a flat prior between $M_{\min} = 4.0 \times 10^6 M_{\odot}$ and $M_{\max} = 4.0 \times 10^9 M_{\odot}$ (Diemand et al. 2007a,b) and a position with a flat prior over the complete data grid. We note that our recovered mass, although close to the upper limit, is well inside this range (see below). We choose this range to make a comparison with simulations easier, but could have chosen a smaller or larger range. The results, however, are similar (only the evidence is offset by a constant value for both the PJ and PL + PJ models).

6.2 The evidence and posterior probability exploration

One of the most efficient methods for exploring the posterior probability within the framework of Bayesian statistics is the nested-sampling technique developed by Skilling (2004). Although being faster than thermodynamic integration, the nested sampling can still be computationally expensive as the overall computational cost rapidly grows with the dimension D of the problem as $O(D^3/e^2)$, where e is the desired level of accuracy (Chopin & Robert 2008). Most of the nested-sampling computational effort is required by the simulations of points from a prior probability distribution $\pi(\eta)$ with the constraints that the relative likelihood $\mathcal{L}(\eta)$ has to be larger than an increasing threshold \mathcal{L}^* . Different approaches have been suggested in order to increase the performance of this simulation. Chopin & Robert (2008), for example, proposed an extension of the nested sampling, based on the principle of importance sampling, while Mukherjee, Parkinson & Liddle (2006) developed an ellipsoidal nested sampling by approximating the isolikelihood contours by D -dimensional ellipsoids. Shaw, Bridges & Hobson (2007) subsequently improved the ellipsoidal nested sampling with a cluster nested sampling which allows efficient sampling also of multimodal posterior distributions.

In our analysis, we replace the standard nested sampling used in Vegetti & Koopmans (2009a) with MULTINEST, a multimodal nested sampling algorithm developed by Feroz & Hobson (2008). As further improved by Feroz, Hobson & Bridges (2009), MULTINEST allows us to efficiently and robustly sample posterior probabilities even when the distributions are multimodal or affected by pronounced degeneracies. The possibility of running the algorithm in parallel mode further reduces the computational load.

The most appealing property of nested-sampling-based techniques is that they also efficiently explore the model parameter posterior probabilities and simultaneously compute the marginalized Bayesian evidence of the model. The former provide error determinations for the parameters of a given model, while the latter allows for a quantitative and objective comparison between different and not necessarily nested (i.e. one model is not necessarily a special case of the other) models. The Bayesian evidence automatically includes the Occam's razor and penalizes models which are unnecessarily complicated. This means that a PL + substructure model is preferred over a single PL only if the data require the presence of extra free parameters and the likelihood of the model increases sufficiently to offset the decrease in prior probability (i.e. extra model parameters lead to a larger prior volume and hence

a smaller prior probability density near the peak of the likelihood function).

6.3 The substructure evidence and model parameters

The main result of the nested-sampling analysis is that the PL + PJ model has a substructure with mean mass

$$M_{\text{sub}} = (3.51 \pm 0.15) \times 10^9 M_{\odot},$$

located at a position $(-0.651 \pm 0.038, 1.040 \pm 0.034 \text{ arcsec})$ (see Table 1); the quoted statistical errors do not take into account the systematic uncertainties, but fully account for all covariance in the mass model. In our case, systematic errors are mostly related to the PSF and to the procedure for the subtraction of the lens galaxy surface brightness (see Marshall et al. 2007, for a discussion). Effects related to systematic uncertainties are explored in Section 6.4. Note that the results of this section are in agreement with those in the previous section and in particular that the substructure is exactly located where the positive convergence correction is found by the pixelized potential reconstruction (see Fig. 3).

Finally, we find that the perturbed PL + PJ model is strongly favoured by the data with $\Delta \log \mathcal{E} = \log \mathcal{E}_{\text{PJ}} - \log \mathcal{E}_{\text{PL+PJ}} = 20\,353.90 - 20\,482.1 = -128.2$. Heuristically, and ignoring the difference in degrees of freedom between the PL and PL + PJ models, this would correspond in classical terms to more or less a dramatic $\Delta \chi^2 \sim 256$ improvement in the model. Given that we have thousands of data pixels and no major residual features this shows that adding only a few extra parameters to the lens model improves the agreement between the model and the data over a wide range of data pixels. Heuristically one might further estimate the substructure mass error to be $\delta M_{\text{sub}} \sim M_{\text{sub}} / \sqrt{2|\Delta \mathcal{E}|} \sim 0.2 \times 10^9 M_{\odot}$, which is indeed close to the proper determination of this error. We are therefore confident about this detection and its strong statistical significance. This represents the first gravitational imaging detection of a dark substructure in a galaxy.

However, to test the robustness of this detection (i.e. systematics) we will now subject our reconstructions against several substantial changes in the model and the data, some of these going far beyond what could be regarded as reasonable changes.

6.4 Robustness and systematic errors

A number of major sources of systematic error might still affect the lens modelling: the PSF modelling, the pixel scale and lens galaxy subtraction from the lens plane. To determine at which level systematic errors influenced the substructure detection we tested the PL + PJ modelling (see Section 5) by rotating the PSF model through 90° from the original one; we call this model $(\text{PL} + \text{PJ})_{\text{psf90}}$. We also used a different data set with smaller drizzled pixels (0.03 arcsec) and a different lens galaxy subtraction procedure (using a Sérsic profile rather than a b-spline surface brightness profile); we call this model $(\text{PL} + \text{PJ})_{\text{subt}}$. We refer to the corresponding smooth models as PL_{psf90} and PL_{subt} , respectively. The results are shown in Figs 5(a)–(d) and listed in Table 1.

More precisely, $(\text{PL} + \text{PJ})_{\text{subt}}$ is not only rotated but also has a different pixel scale (0.03 arcsec pixel $^{-1}$), a different number of pixels, a different noise level and a different PSF, so that we also test against all these changes. We also check whether the form of source regularization has any effect by running a PL and a PL + PJ modelling for a non-adaptive regularization constant and for a gradient regularization. Finally we run an optimization for both the smooth and the perturbed model in which the centre of the lens is

allowed to change [PL_{ctr} and $(\text{PL} + \text{PJ})_{\text{ctr}}$] and an optimization with a larger PSF.

All tests (see Table 1) lead to results that are consistent with each other for both the main lens and the substructure parameters. First, we note that rotating the PSF changes the evidence by a value that could be expected based on the sampling error in the nested sampling. Hence we conclude that PSF effects are negligible. In the case of the ‘subt’ model, we note that we are no longer comparing the same data sets and that the evidence values have dramatically changed. This simply reflects the large increase by a factor of $\sim (0.05/0.03)^2$ in the number of data points. Bayesian evidence cannot be used to compare different data sets, but we can compare the PL_{subt} and $(\text{PL} + \text{PJ})_{\text{subt}}$ models amongst each other. First, we remark that the pixel scale in this data set is considerably smaller than the resolution and pixel scale in the image; hence neither the data pixels nor their errors are fully independent. This leads to a rather odd stripped source reconstruction, not observed for the original data set. Despite this difference, we note that image residuals in the PL + PJ models are reduced, especially near the substructure positions, compared to the PL model. The likelihood difference is $\Delta \log \mathcal{L}_{\text{subt}} = \log \mathcal{L}_{\text{PJ}} - \log \mathcal{L}_{\text{PL+PJ}} = -154$ in favour of the substructure model.

We have considered a wide range of possible sources of systematic errors and we have applied rather extreme changes, beyond what one would expect realistic scenarios. This shows that the substructure detection is stable. Overall, we are therefore confident that the substructure detection is not only a statistically sound detection, but also robust against dramatic changes in the model and the data.

6.5 The substructure mass-to-light ratio

Based on the residual images, we determine an upper limit on the magnitude of the substructure in two different ways: by setting the limit equal to three times the estimated (cumulative) noise level or by aperture-flux fitting, both inside a 5×5 pixel ($0.25 \times 0.25 \text{ arcsec}^2$) window. The aperture is chosen to gather most of the light of the substructure, which is expected to be effectively point-like, given the typical size of the optical counterpart of galaxies of $\approx 10^9 M_{\odot}$. They are in good agreement because the image residuals are very close to the noise level. The 3σ limit is found to be $I_{F814W,3\sigma} > 27.5$ mag. At the redshift of the lens galaxy this corresponds to a 3σ upper limit in luminosity of $5.0 \times 10^6 L_{V,\odot}$. Within the inner 0.3 and 0.6 kpc, we therefore find that the integrated masses are, respectively, $(5.8 \pm 0.3) \times 10^8$ and $(1.1 \pm 0.05) \times 10^9 M_{\odot}$, and hence lower limits are $(M/L)_{V,\odot} \gtrsim 120 M_{\odot}/L_{V,\odot}$ (3σ) and $(M/L)_{V,\odot} \gtrsim 218 M_{\odot}/L_{V,\odot}$ (3σ).

The mass of the substructure is about one order of magnitude larger than the mass of Milky Way (MW) satellites inside 0.3 kpc (see Figs 6a and b; e.g. Strigari et al. 2008). Whereas this might appear high, this is not necessarily in disagreement with MW observations. Diemand et al. (2007b), for example, showed that these massive substructures can have a mass at the present time of only 40 per cent of the virial mass before accretion, due to tidal effects. The total mass within the tidal radius of the substructure is currently $3.5 \times 10^9 M_{\odot}$, which then translates into $M_{\text{vir,acc}} \approx 10^{10} M_{\odot}$ before accretion. As shown by Kravtsov (2010) a subhalo with $M_{\text{vir,acc}} \approx 10^{10} M_{\odot}$, which formed at a redshift $z \gtrsim 10$, corresponds to the mass of a few $\times 10^8 M_{\odot}$ that we find within 0.3 kpc. To be in agreement with the MW observations, the $M_{\text{vir,acc}}-L_V$ relationship proposed by Kravtsov (2010) yields a satellite luminosity of $L_V = 1.2 \times 10^6 L_{\odot}$, which is lower than our detection limit. Hence, if the detected substructure was formed early (i.e. $z \gtrsim 10$), not unlikely

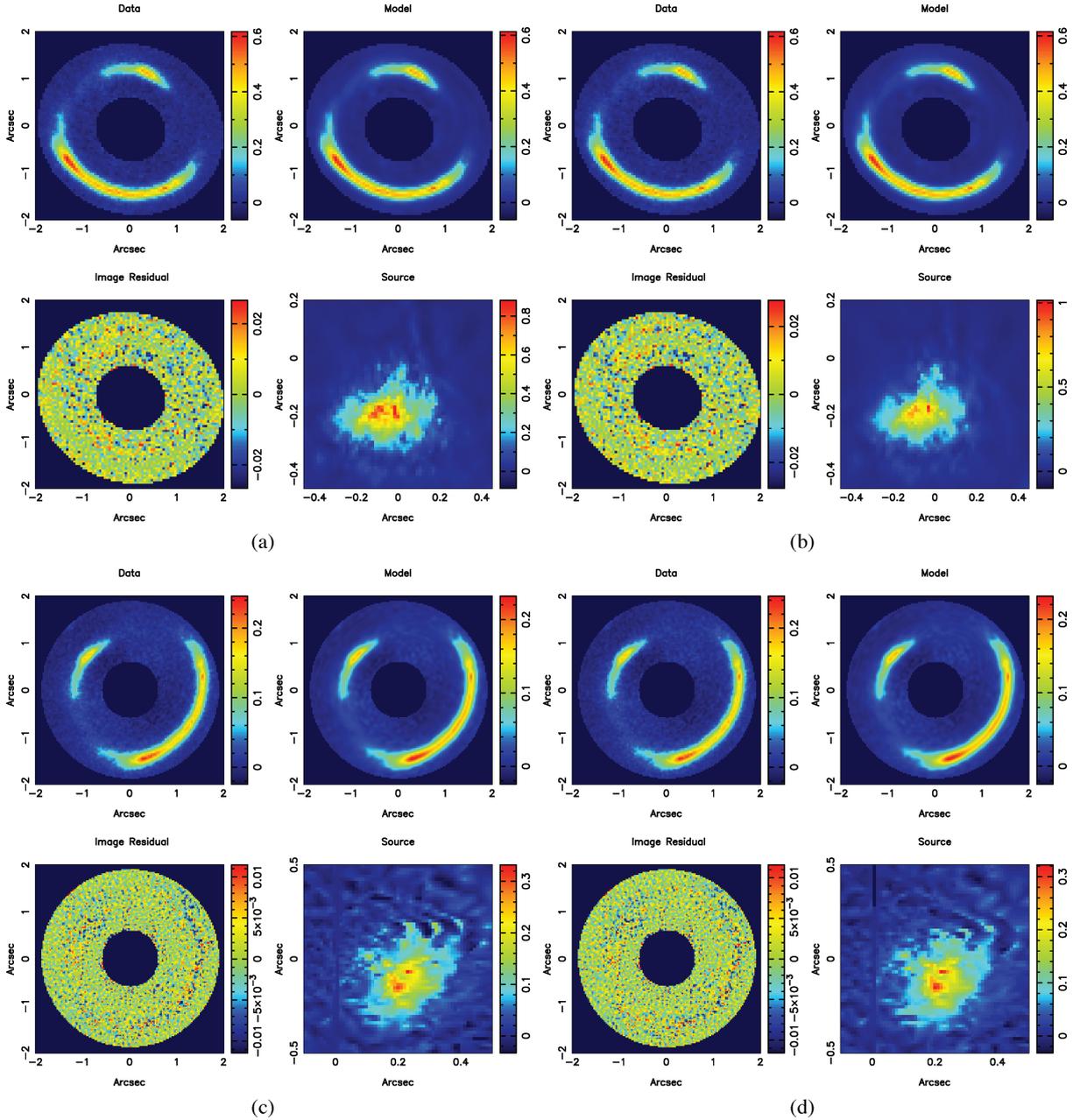


Figure 5. The over-regularised PL model with a rotated PSF ((a) panels). The PL + PJ model with a rotated PSF ((b) panels). The over-regularised PL model with a smaller pixel scale and a different procedure for the lens galaxy subtraction ((c) panels). The PL + PJ model with a smaller pixel scale and a different procedure for the lens galaxy subtraction ((d) panels).

for massive early-type galaxies whose building blocks formed in the highest density peaks in the early universe, our detection is fully consistent with all observations. It does, however, require an earlier typical formation redshift of the substructure in this galaxy compared to those in the Local Group.

Also, the normalization of the mass function number density scales with the total mass of the host galaxy. Because SDSSJ0946+1006 is substantially more massive than the MW at fixed radius, it is therefore not surprising to find a satellite of this mass around such a massive elliptical galaxy. Madau, Diemand & Kuhlen (2008), for example, showed that a dark matter halo eight times more massive than a MW halo contains almost a factor of 2 more substructures with larger circular velocities. Gao et al. (2004) looked at haloes of different scales (i.e. cluster and galaxy) and they con-

cluded that the subhalo abundance in high-mass haloes is higher than that in low-mass haloes. The difference between a rich cluster halo and a galaxy halo is a factor of 2. Similar results were found with the semi-analytical modelling by Zentner & Bullock (2003). This is expected in the hierarchical scenario, where the more massive haloes assemble later and their satellites have not been destroyed yet.

7 THE SUBSTRUCTURE MASS FUNCTION

What does this imply for the Λ CDM model and the expected fraction of mass in substructure? Given the statistical formalism presented in Vegetti & Koopmans (2009b), we can use this detection to constrain the projected dark matter mass fraction in substructure

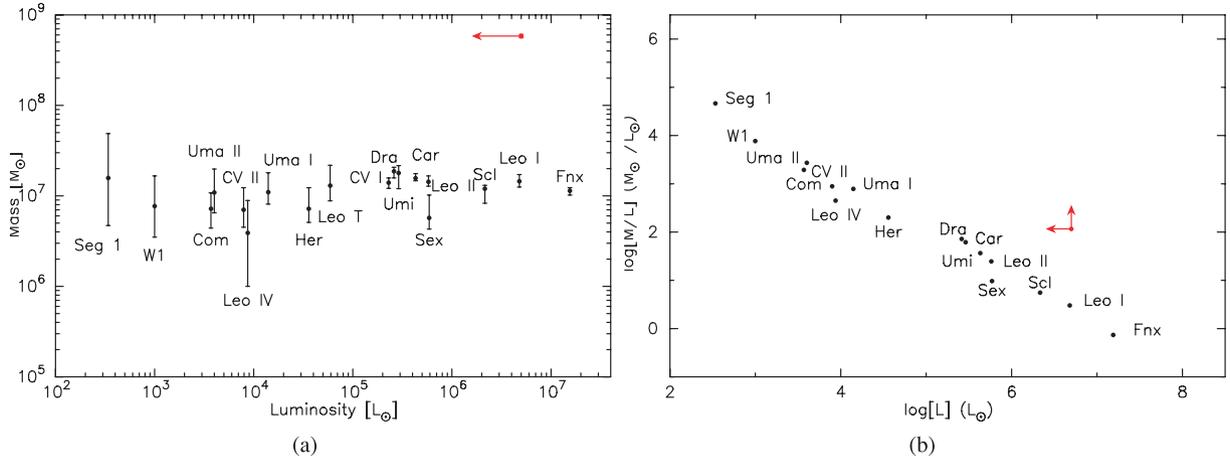


Figure 6. Left-hand panel: the integrated mass in units of solar masses, within the inner 0.3 kpc as a function of the total luminosity, in units of solar luminosity for the MW satellites (black points) and the substructure detected in this paper (red arrow). Note the small error on the substructure mass. Right-hand panel: the integrated M/L , within the inner 0.3 kpc as a function of the total luminosity, in units of solar luminosity for the MW satellites (black points) and the substructure detected in this paper (red arrow). We note that the results for the MW, a spiral galaxy, and the system analysed in this paper, an elliptical galaxy, may not be directly comparable as explained in the text.

f and the substructure mass function slope $dN/dm \propto m^{-\alpha}$. We note that we can ignore the baryonic content in the substructure because of its large total mass-to-light ratio (M/L).

To make a proper comparison with simulations, we assume that the substructure mass can assume any value from $M_{\min} = 4.0 \times 10^6 M_{\odot}$ to $M_{\max} = 4.0 \times 10^9 M_{\odot}$ (Diemand et al. 2007a,b, 2008) and that the mass we can detect varies from $M_{\text{low}} = 0.15 \times 10^9 M_{\odot}$ to $M_{\text{high}} = M_{\max}$. The current lower limit is set by the resolution of the current simulations. If that limit decreases, then fraction of substructure increases in that widened range. However, this does not modify the fraction within the well-defined range where the current simulations have already converged in the substructure mass function. We note that different limits would scale both the simulation and observed mass fraction in the same way. The mass fractions quoted throughout the papers are for this mass range only. We ignore the error on the measured substructure mass; this does not influence the results because our detection is well beyond the error $\sigma_m = 0.15 \times 10^9 M_{\odot}$ level. Given the quality of the fit for a PL + PJ model, we are confident that there are no other substructures with mass above our detection threshold.

Fig. 7 shows the joint posterior probability density function $P(\alpha, f | n_s = 1, m = M_{\text{sub}}, \mathbf{p})$ contours and the marginalized probability densities $P(f | n_s = 1, m = M_{\text{sub}}, \mathbf{p})$ and $P(\alpha | n_s = 1, m = M_{\text{sub}}, \mathbf{p})$, given one detected substructure $n_s = 1$ with mass $m = M_{\text{sub}}$, where \mathbf{p} is a vector containing the model parameters, $M_{\min}, M_{\max}, M_{\text{low}}$ and M_{high} . Specifically, from the marginalized probability density distributions we find $f = 2.56_{-1.50}^{+3.26}$ per cent and $\alpha = 1.36_{-0.28}^{+0.81}$ at a 68 per cent confidence level (CL) for a flat prior on α and $f = 2.15_{-1.25}^{+2.05}$ per cent and $\alpha = 1.88_{-0.10}^{+0.10}$ at a 68 per cent CL for a Gaussian prior centred in 1.90 ± 0.1 . The same results are found if an error on the mass measurement and a detection threshold $M_{\text{low}} = 3 \times \sigma_m$ are assumed.

As already discussed in Vegetti & Koopmans (2009b), while even a single-lens system is enough to set upper and lower limits on the mass fraction, a larger number of lenses is required in order to constrain the mass function slope, unless a stringent prior information on the parameter itself is adopted. By assuming a Gaussian prior of the mass function slope centred at 1.90, we can quantify the probability that the dark matter mass fraction is the one given by

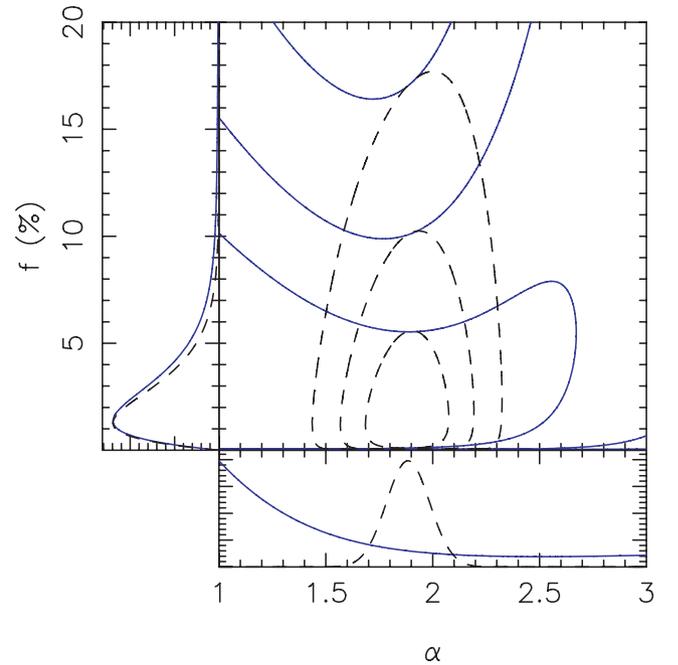


Figure 7. Joint probability $P(\alpha, f | \{n_s, m\}, \mathbf{p})$ contours and marginalized probabilities $P(f | \{n_s, m\}, \mathbf{p})$ and $P(\alpha | \{n_s, m\}, \mathbf{p})$ for a uniform prior (solid lines) and for a Gaussian prior in α (dashed lines). Contours (inside out) are set at levels $\Delta \log(P) = -1, -4, -9$ from the peak of the posterior probability density.

N -body simulations $f_{N\text{-body}} \approx 0.3$ per cent (Diemand et al. 2007a,b, 2008), by considering the ratio between the posterior probability densities $P(f_{N\text{-body}}, \alpha = 1.9 | n_s = 1, m = M_{\text{sub}}, \mathbf{p})$ and $P_{\max}(f, \alpha = 1.9 | n_s = 1, m = M_{\text{sub}}, \mathbf{p})$ and find that this ratio is 0.51. Hence, currently our measurement and that inferred from N -body simulations still agree as a result of the rather larger error bar on the measured value of f . The combination of more lens systems is, of course, required to set more stringent constraints also on α . We plan such an analysis in forthcoming papers.

Given our best value of $f = 0.0215$ for $\alpha = 1.9$, we might expect to detect ~ 1 mass substructures above our 3σ mass threshold of $4.5 \times 10^8 M_\odot$. It is therefore unlikely that we have missed many substructures with a mass slightly below that of our detection. Given this result and image residuals already at the noise level, we believe that adding a second substructure is not warranted and missing lower mass substructure leads only to a minor bias (note that logarithmic bins have nearly equal amounts of mass for $\alpha = 1.9$).

8 SUMMARY

We have applied our new Bayesian and adaptive-grid method for pixelized source *and* lens potential modelling (Vegetti & Koopmans 2009a) to the analysis of *HST* data of the double Einstein ring system SLACS SDSSJ0946+1006 (Gavazzi et al. 2008). This system was chosen based on its large expected dark matter mass fraction near the Einstein radius and the high signal-to-noise ratio of the lensed images. Although these two facts should be uncorrelated to the mass fraction of CDM substructure, both incidences maximize the change of detection.

We find that a smooth elliptical power-law model of the system leaves significant residuals near or above the noise level; these residuals are correlated and spread over a significant part of the lensed images. Through a careful modelling of this data including either lens potential corrections or an additional (low-mass) simply parametrized mass component, we conclude that the massive early-type lens galaxy of SLACS SDSSJ0946+1006 hosts a large- M/L substructure with a mass around $M_{\text{sub}} \sim 3.5 \times 10^9 M_\odot$, situated on one of the lensed images. A careful statistical analysis of the image residuals, as well as a number of more drastic robustness tests (e.g. changing the PSF, pixel number and scale, regularization level and form, galaxy subtraction and image rotation), confirm and support this detection. Based on this detection, the first of its kind, we derive a projected CDM substructure mass fraction of ~ 2.2 per cent for the inner regions of the galaxy, using the Bayesian method of Vegetti & Koopmans (2009b); this fraction is high, but still consistent with expectations from numerical simulations due to the large (Poisson) error based on a single detection.

The numerical details of our results can be further summarized as follows.

(1) Using a Bayesian MULTINEST Markov-chain exploration of the full model parameter space, we show that the identified object has a mass of $M_{\text{sub}} = (3.51 \pm 0.15) \times 10^9 M_\odot$ (68 per cent CL) and is located near the inner Einstein ring at $(-0.651 \pm 0.038, 1.040 \pm 0.034)$ arcsec. The Bayesian evidence is in favour of a model that includes a substructure versus a smooth elliptical power law only, with $\Delta \log(\mathcal{E}) = -128.2$. This is roughly equivalent to a 16σ detection.

(2) At the redshift of the lens galaxy a 3σ upper limit in luminosity is found to be $5.0 \times 10^6 L_{V,\odot}$. Within the inner 0.3 and 0.6 kpc, we find that the integrated masses are, respectively, $(5.8 \pm 0.3) \times 10^8$ and $(1.1 \pm 0.05) \times 10^9 M_\odot$, and hence lower limits are $(M/L)_{V,\odot} \gtrsim 120 M_\odot/L_{V,\odot}$ (3σ) and $(M/L)_{V,\odot} \gtrsim 218 M_\odot/L_{V,\odot}$ (3σ). This is higher than that of MW satellites, but may be not unexpected for satellites near massive elliptical galaxies. The mass and luminosity limit of this substructure are consistent with Local Group results if the substructure had a virial mass of $\sim 10^{10} M_\odot$ before accretion and formed at $z \gtrsim 10$.

(3) The CDM mass fraction and a mass function slope are equal to $f = 2.15_{-1.25}^{+2.05}$ per cent and $\alpha = 1.88_{-0.10}^{+0.10}$, respectively, at a 68 per cent CL for a Gaussian prior on α centred on 1.90 ± 0.1 . For a

flat prior on α between 1.0 and 3.0, we find $f = 2.56_{-1.50}^{+3.26}$ per cent. Asking whether the $f = 2.15$ per cent is consistent with $f_{N\text{-body}} = 0.3$ per cent, we find a likelihood ratio of 0.51; indeed both are consistent. This is the result of the considerable measurement error found for f because it is based on only a single detection.

This is the first application of our adaptive ‘gravitational imaging’ method to real data and clearly shows its promise. In the near future we will apply the method to a larger set of SLACS lenses in order to constrain, via the statistical formalism presented in Vegetti & Koopmans (2009b), the dark matter fraction in substructure and the substructure mass function.

ACKNOWLEDGMENTS

SV is grateful to Farhan Feroz for the help provided in the implementation of MULTINEST. We thank Simon White and Chris Kochanek for useful suggestions, and the anonymous referee for useful suggestions to further improve the discussion in the paper. SV and LVEK are supported through an NWO-VIDI programme subsidy (project number 639.042.505). TT acknowledges support from the NSF through CAREER award NSF-0642621, from the Sloan Foundation through a Sloan Research Fellowship and from the Packard Foundation through a Packard Fellowship. This research is supported by NASA through *HST* programmes SNAP-10174, GO-10494, SNAP-10587, GO-10798 and GO-10886. Based on observations made with the NASA/ESA *HST*, obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. Funding for the creation and distribution of the SDSS Archive has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the US Department of Energy, the Japanese Monbukagakusho and the Max Planck Society. The SDSS web site is <http://www.sdss.org>. The SDSS is managed by the Astrophysical Research Consortium (ARC) for the Participating Institutions. The Participating Institutions are the University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Korean Scientist Group, Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory and the University of Washington.

REFERENCES

- Barkana R., 1998, *ApJ*, 502, 531
- Barnabè M., Nipoti C., Koopmans L. V. E., Vegetti S., Ciotti L., 2009, *MNRAS*, 393, 1114
- Bender R., Burstein D., Faber S. M., 1993, *ApJ*, 411, 153
- Blandford R., Surpi G., Kundić T., 2001, in Brainerd T. G., Kochanek C. S., eds, *ASP Conf. Ser. Vol. 237, Gravitational Lensing: Recent Progress and Future Goals*. Astron. Soc. Pac., San Francisco, p. 65
- Bolton A. S., Burles S., Koopmans L. V. E., Treu T., Moustakas L. A., 2006, *ApJ*, 638, 703
- Bower R. G., Lucey J. R., Ellis R. S., 1992, *MNRAS*, 254, 589
- Bradač M., Schneider P., Steinmetz M., Lombardi M., King L. J., Porcas R., 2002, *A&A*, 388, 373
- Bradač M., Schneider P., Lombardi M., Steinmetz M., Koopmans L. V. E., Navarro J. F., 2004, *A&A*, 423, 797
- Brewer B. J., Lewis G. F., 2006, *ApJ*, 637, 608

- Chen J., Kravtsov A. V., Keeton C. R., 2003, *ApJ*, 592, 24
 Chen J., Rozo E., Dalal N., Taylor J. E., 2007, *ApJ*, 659, 52
 Chiba M., 2002, *ApJ*, 565, 17
 Chopin N., Robert C., 2008, preprint (arXiv:0801.3887)
 Dalal N., Kochanek C. S., 2002, *ApJ*, 572, 25
 Diemand J., Kuhlen M., Madau P., 2007a, *ApJ*, 657, 262
 Diemand J., Kuhlen M., Madau P., 2007b, *ApJ*, 667, 859
 Diemand J., Kuhlen M., Madau P., Zemp M., Moore B., Potter D., Stadel J., 2008, *Nat*, 454, 735
 Djorgovski S., Davis M., 1987, *ApJ*, 313, 59
 Dressler A., Lynden-Bell D., Burstein D., Davies R. L., Faber S. M., Terlevich R., Wegner G., 1987, *ApJ*, 313, 42
 Feroz F., Hobson M. P., 2008, *MNRAS*, 384, 449
 Feroz F., Hobson M. P., Bridges M., 2009, *MNRAS*, 398, 1601
 Ferrarese L., Merritt D., 2000, *ApJ*, 539, L9
 Gao L., White S. D. M., Jenkins A., Stoehr F., Springel V., 2004, *MNRAS*, 355, 819
 Gavazzi R., Treu T., Rhodes J. D., Koopmans L. V. E., Bolton A. S., Burles S., Massey R. J., Moustakas L. A., 2007, *ApJ*, 667, 176
 Gavazzi R., Treu T., Koopmans L. V. E., Bolton A. S., Moustakas L. A., Burles S., Marshall P. J., 2008, *ApJ*, 677, 1046
 Gebhardt K. et al., 2000, *ApJ*, 539, L13
 Guzman R., Lucey J. R., Carter D., Terlevich R. J., 1992, *MNRAS*, 257, 187
 Kauffmann G., White S. D. M., Guiderdoni B., 1993, *MNRAS*, 264, 201
 Keeton C. R., Gaudi B. S., Petters A. O., 2003, *ApJ*, 598, 138
 Keeton C. R., Gaudi B. S., Petters A. O., 2005, *ApJ*, 635, 35
 Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, *ApJ*, 522, 82
 Kochanek C. S., Dalal N., 2004, *ApJ*, 610, 69
 Koopmans L. V. E., 2005, *MNRAS*, 363, 1136
 Koopmans L. V. E., Treu T., Bolton A. S., Burles S., Moustakas L. A., 2006, *ApJ*, 649, 599
 Koopmans L. V. E. et al., 2009, *ApJ*, 703, L51
 Kravtsov A., 2010, *Advances Astron.*, 2010, 8
 Macciò A. V., Miranda M., 2006, *MNRAS*, 368, 599
 MacKay D. J. C., 1992, PhD thesis, Caltech
 MacKay D. J. C., 2003, *Information Theory, Inference, and Learning Algorithms*. Cambridge Univ. Press, Cambridge
 McKean J. P. et al., 2007, *MNRAS*, 378, 109
 Madau P., Diemand J., Kuhlen M., 2008, *ApJ*, 679, 1260
 Magorrian J. et al., 1998, *AJ*, 115, 2285
 Mao S., Schneider P., 1998, *MNRAS*, 295, 587
 Marshall P. J. et al., 2007, *ApJ*, 671, 1196
 Metcalf R. B., Zhao H., 2002, *ApJ*, 567, L5
 Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, *ApJ*, 524, L19
 Moore B., Calcáneo-Roldán C., Stadel J., Quinn T., Lake G., Ghigna S., Governato F., 2001, *Phys. Rev. D*, 64, 063508
 More A., McKean J. P., More S., Porcas R. W., Koopmans L. V. E., Garrett M. A., 2009, *MNRAS*, 394, 174
 Mukherjee P., Parkinson D., Liddle A. R., 2006, *ApJ*, 638, L51
 Rusin D., Kochanek C. S., 2005, *ApJ*, 623, 666
 Shaw J. R., Bridges M., Hobson M. P., 2007, *MNRAS*, 378, 1365
 Skilling J., 2004, in Fischer R., Preuss R., Toussaint U. V., eds, *AIP Conf. Ser. Vol. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering*. Am. Inst. Phys., New York, p. 395
 Springel V. et al., 2008, *MNRAS*, 391, 1685
 Strigari L. E., Bullock J. S., Kaplinghat M., Simon J. D., Geha M., Willman B., Walker M. G., 2008, *Nat*, 454, 1096
 Suyu S. H., Marshall P. J., Hobson M. P., Blandford R. D., 2006, *MNRAS*, 371, 983
 Treu T., Koopmans L. V. E., 2004, *ApJ*, 611, 739
 Vegetti S., Koopmans L. V. E., 2009a, *MNRAS*, 392, 945
 Vegetti S., Koopmans L. V. E., 2009b, *MNRAS*, 400, 1583
 Xu D. D. et al., 2009, *MNRAS*, 398, 1235
 Zentner A. R., Bullock J. S., 2003, *ApJ*, 598, 49

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.