

The distribution of dark matter and gas spanning 6 Mpc around the post-merger galaxy cluster MS 0451–03

Sut-Ieng Tam ¹★, Mathilde Jauzac ^{1,2,3}, Richard Massey ^{1,2}, David Harvey ^{1,4},
Dominique Eckert ^{1,5,6}, Harald Ebeling ⁷, Richard S. Ellis ⁸, Vittorio Ghirardini ⁵,
Baptiste Klein ^{9,10}, Jean-Paul Kneib ¹¹, David Lagattuta ^{1,2}, Priyamvada Natarajan ¹²,
Andrew Robertson ¹ and Graham P. Smith ¹³

¹*Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK*

²*Centre for Extragalactic Astronomy, Durham University, South Road, Durham DH1 3LE, UK*

³*Astrophysics and Cosmology Research Unit, School of Mathematical Sciences, University of KwaZulu-Natal, Durban 4041, South Africa*

⁴*Instituut-Lorentz for Theoretical Physics, Universiteit Leiden, Niels Bohrweg 2, Leiden, NL-2333 CA, the Netherlands*

⁵*Max-Planck Institut für Extraterrestrische Physik, Giessenbachstrasse 1, D-85748 Garching, Germany*

⁶*Astronomy Department, University of Geneva, 16 ch. d'Ecogia, CH-1290 Versoix, Switzerland*

⁷*Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA*

⁸*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK*

⁹*Université de Toulouse, UPS-OMP, IRAP, 14 Avenue E. Belin, F-31400 Toulouse, France*

¹⁰*CNRS, IRAP/UMR 5277, Toulouse, 14 Avenue E. Belin, F-31400 Toulouse, France*

¹¹*Laboratoire d'Astrophysique, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, CH-1290 Versoix, Switzerland*

¹²*Department of Astronomy, Yale University, 260 Whitney Avenue, New Haven, CT 06511, USA*

¹³*School of Physics and Astronomy, University of Birmingham, Birmingham B15 2TT, UK*

Accepted 2020 June 15. Received 2020 June 14; in original form 2020 February 19

ABSTRACT

Using the largest mosaic of *Hubble Space Telescope* images around a galaxy cluster, we map the distribution of dark matter throughout an $\sim 6 \times 6$ Mpc² area centred on the cluster MS 0451–03 ($z = 0.54$, $M_{200} = 1.65 \times 10^{15} M_{\odot}$). Our joint strong- and weak-lensing analysis shows three possible filaments extending from the cluster, encompassing six group-scale substructures. The dark matter distribution in the cluster core is elongated, consists of two distinct components, and is characterized by a concentration parameter of $c_{200} = 3.79 \pm 0.36$. By contrast, *XMM-Newton* observations show the gas distribution to be more spherical, with excess entropy near the core, and a lower concentration of $c_{200} = 2.35^{+0.89}_{-0.70}$ (assuming hydrostatic equilibrium). Such a configuration is predicted in simulations of major mergers 2–7 Gyr after the first core passage, when the two dark matter haloes approach second turnaround, and before their gas has relaxed. This post-merger scenario finds further support in optical spectroscopy of the cluster's member galaxies, which shows that star formation was abruptly quenched 5 Gyr ago. MS 0451–03 will be an ideal target for future studies of the growth of structure along filaments, star formation processes after a major merger, and the late-stage evolution of cluster collisions.

Key words: galaxies: clusters: general – gravitational lensing: weak – cosmology: observations – large-scale structure of Universe.

1 INTRODUCTION

The standard Λ CDM (Lambda Cold Dark Matter) model of cosmology suggests that the large-scale structure (LSS) formed hierarchically via a series of mergers with smaller haloes and accretion of surrounding matter (White & Rees 1978; Springel et al. 2005; Schaye et al. 2015). Reaching total masses of several $10^{15} M_{\odot}$,

galaxy clusters are the largest and rarest structures resulting from this hierarchical formation process. Since their properties depend on the growth of structure (from the seeds provided by primordial density fluctuations, through gravitational collapses, to accretion of matter funnelled on to them along filaments), clusters are ideally suited to test cosmological models (e.g. Bahcall & Cen 1993; Meneghetti et al. 2005; Rozo et al. 2010; de Haan et al. 2016; Jauzac et al. 2016; Schwinn et al. 2017).

Approximately 80 percent of a cluster's mass consists of dark matter. Although this component is invisible, the total mass along a

* E-mail: sut-ieng.tam@durham.ac.uk

line of sight can be mapped through measurements of the deflection of light from background objects by gravitational lensing, a process that is independent of the physical or dynamical state of the lensing matter (see reviews by e.g. Massey, Kitching & Richard 2010; Kneib & Natarajan 2011; Hoekstra 2013; Kilbinger 2015; Treu & Ellis 2015; Bartelmann & Maturi 2017). The distinctive signatures of strong gravitational lensing (multiple images or giant arcs) probe the mass distribution in the inner region of clusters, while weak gravitational lensing provides constraints on the cluster environment on larger scales (Gavazzi et al. 2007; Massey et al. 2007; Shan et al. 2012; Zitrin et al. 2013). Combining strong- and weak-lensing analyses can thus constrain the mass distribution across the entire cluster (Broadhurst et al. 2005; Bradač et al. 2008; Jauzac et al. 2018).

Wide-field observations of weak lensing with ground-based telescopes have successfully measured clusters' properties, including their mass (e.g. Umetsu et al. 2014, 2019; Okabe & Smith 2016; Medezinski et al. 2018a; Herbonnet et al. 2019; McClintock et al. 2019; Miyatake et al. 2019; Rehmman et al. 2019) and halo shape (e.g. Evans & Bridle 2009; Oguri et al. 2010; Clampitt & Jain 2016; van Uitert et al. 2017; Chiu et al. 2018; Shin et al. 2018; Umetsu et al. 2018). The latter is of particular interest since it reflects the nature of dark matter (specifically whether dark matter is collisionless; Robertson et al. 2019). On larger scales, the mass distribution's shape is governed by accretion of matter from the surroundings. As substructures are accreted on to clusters along filaments (Bond, Kofman & Pogosyan 1996; Yess & Shandarin 1996; Aragón-Calvo et al. 2007; Angulo et al. 2012), cluster mass haloes tend to align with the directions of infall (e.g. Warren et al. 1992; Jing & Suto 2002). Direct detection of the mass in filamentary LSS through gravitational lensing is, however, extremely challenging with ground-based observations because of the filaments' low mass and the low density of resolved galaxies behind them (e.g. Kaiser et al. 1998; Gray et al. 2002; Gavazzi et al. 2004; Clowe et al. 2006; Dietrich et al. 2012; Martinet et al. 2016).

The higher angular resolution afforded by space-based imaging increases the signal-to-noise ratio (S/N) of lensing measurements. However, efforts to exploit this advantage are currently limited by the small field of view of the *Hubble Space Telescope* (*HST*). In the next decade, high-resolution observations from space over a much wider field of view will become possible with *Euclid* (Laureijs et al. 2011), *Nancy Grace Roman Space Telescope* (Spergel et al. 2013), and the balloon-borne telescope *SuperBIT* (Romualdez et al. 2016; Romualdez et al. 2018). It is thus timely to hone analysis methods that will exploit this new era of wide-field, high-resolution lensing data.

Multiwavelength data, including X-ray spectral imaging of the intracluster medium (ICM), are crucial to our understanding of the dynamics in clusters. Since dark matter and baryons interact differently during a merger, a combined study of the distributions of dark matter and ICM provides insights into clusters' evolutionary history (e.g. Bradač et al. 2006; Merten et al. 2011; Jauzac et al. 2015; Ogorean et al. 2015; Molnar & Broadhurst 2018). Furthermore, X-ray analyses usually assume that the ICM is in hydrostatic equilibrium (HSE) and spherically symmetric. Therefore, a comparison between X-ray and lensing mass measurements can be used to test the validity of the HSE assumption.

In this paper, we use a wide-field *HST/ACS* imaging mosaic and *XMM-Newton* observations to conduct a combined strong- and weak-lensing and X-ray analysis of the galaxy cluster MS 0451–03 ($z = 0.54$; hereafter MS 0451), also known as MACS J0454.1–0300 (Ebeling, Edge & Henry 2001; Ebeling et al. 2007) and the most

X-ray luminous cluster in the Extended Medium Sensitivity Survey (EMSS; Gioia et al. 1990). MS 0451 was extensively studied previously at optical wavelengths (Luppino et al. 1999; Moran et al. 2007a, b; Jørgensen & Chiboucas 2013; Soucail et al. 2015; Martinet et al. 2016), in X-rays (Molnar et al. 2002; Donahue et al. 2003; Jeltema et al. 2005; Jørgensen et al. 2018), and via the Sunyaev–Zel'dovich (SZ) effect (De Filippis et al. 2005; Sayers et al. 2019). Strong gravitational lensing analyses have built a model of the cluster core (Borys et al. 2004; Berciano Alba et al. 2010; Zitrin et al. 2011; MacKenzie et al. 2014; Jauzac et al. 2020), and a ground-based weak-lensing analysis detected a possible filamentary structure (Martinet et al. 2016). In 2004, MS 0451 was extensively observed with *HST* over a large area, providing the community with the largest *HST* mosaic centred on a galaxy cluster to date. In this paper, we exploit these wide *HST* observations, combining strong and weak gravitational lensing to map the mass distribution out to a projected radius of ~ 3 Mpc.

This paper is organized as follows. The multiwavelength observations of MS 0451 upon which we base our analysis are summarized in Section 2. Our methods for measuring gravitational lensing and reconstructing the distribution of mass are described in Section 3, while our X-ray data analysis is presented in Section 4. Our measurements of the main cluster halo and surrounding LSS are the subject of Section 5. We discuss the cluster's dynamical state in Section 6, before presenting our conclusions in Section 7.

Throughout this paper, we adopt a Λ CDM cosmology with $\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$, and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$; 1 arcsec corresponds to 6.49 kpc at the redshift of the cluster. All magnitudes are quoted in the AB system.

2 OBSERVATIONS

2.1 Hubble Space Telescope observations

2.1.1 HST/ACS

A mosaic of 41 high-resolution images spanning ~ 20 arcmin \times 20 arcmin around MS 0451 was obtained with the *Advanced Camera for Surveys* onboard *HST* (ACS; Ford et al. 1996) between 2004 January 19 and February 3 (GO-9836, PI: R. Ellis), in the F814W passband, with an exposure time of 2 ks per pointing (single-orbit depth). We reduced the data with the PYHST software package,¹ which corrects charge transfer inefficiency using ARCTIC (Massey et al. 2014), removes detector bias and applies flat-field corrections using CALACS (Miles et al. 2018), and finally stacks the dithered images using ASTRODRIZZLE (Hoffman & Avila 2018). We use these high-resolution images to measure the effect of weak gravitational lensing on the shapes of background galaxies.

2.1.2 HST/WFC3

A 2 arcmin \times 2 arcmin region in the cluster core was imaged with the *Wide Field Camera 3* onboard *HST* (WFC3; Kimble et al. 2008) on 2010 January 13 (GO-11591, PI: J.-P. Kneib). We use observations in the F110W and F160W passbands, with exposure times of 17 912 and 17 863 s, respectively, for the strong-lensing analysis.

¹<https://github.com/davidharvey1986/pyHST>

2.2 Ground-based observations

2.2.1 Imaging data

Multicolour imaging in the B , V , R_c , I_c , and z' passbands was obtained with the 8.3 m Subaru telescope's wide-field Suprime-Cam camera for 1440, 2160, 3240, 1800, and 1620 s, respectively. Observations were performed on 2006 December 21 (z'), 2001 December 11 (R_c, I_c), and 2009 January 23 (B, V). Near-UV imaging in the u^* passband was obtained with the 3.6 m CFHT's MegaPrime camera for 6162 s on 2006 November 27 (ID: 06BH34, PI: H. Ebeling). Near-infrared observations in the J and K_s passbands were performed with CFHT's Wide-field InfraRed Camera (WIRCam) on 2008 November 8 and 2007 October 25, respectively (IDs: 08BH63 and 07BH98, PI: C.-J. Ma). All observations were dithered to facilitate the removal of cosmic rays and minimize the impact of pixel defects and chip gaps; all data were reduced using standard procedures (Donovan 2007).

We use these data to measure photometric redshifts and thereby identify galaxies within, in front of, or behind the cluster. In order to allow a robust estimate of the spectral energy distribution (SED) to be obtained for all objects within the field of view, data from different passbands were seeing matched using the technique described in Kartaltepe et al. (2008). The object catalogue was then created with the SEXTRACTOR photometry package (Bertin & Arnouts 1996) in 'dual-image mode', with the R_c -band image as the reference detection image. Photometric redshifts for galaxies with magnitude $R_c < 24$ were subsequently computed using the adaptive SED-fitting code LE PHARE (Arnouts et al. 1999; Ilbert et al. 2009). For more details of this procedure, see Ma et al. (2008).

2.2.2 Spectroscopic observations

MS 0451 was also observed with the Multi-Unit Spectroscopic Explorer (MUSE; Bacon et al. 2010) at the VLT on 2016 January 10–11 (ID: 096.A-0105(A), PI: J.-P. Kneib), in WFM-NOAO-N mode and good seeing of approximately 0.8 arcsec. The MUSE observations consist of two pointings of three exposures, slightly shifted to account for systematic variations in the detector response, and cover a field of view of ~ 2.2 arcmin². These data are used for the strong-lensing analysis. They were reduced using version 1.6.4 of the MUSE standard pipeline (Weilbacher et al. 2012, 2014), which applies bias and illumination corrections; performs geometrical, astrometric, and flux calibrations; and then combines the individual exposures for each pointing into a single data cube. The sky residuals within each data cube were subtracted using the Zurich Atmosphere Purge algorithm (Soto et al. 2016), which masks sources identified by SEXTRACTOR (Bertin & Arnouts 1996), and then uses principal component analysis to model the sky background.

The spectroscopic redshifts of galaxies used in this work were compiled from the literature and complemented by redshifts measured by us, based on spectroscopic observations obtained in 2004 September with Gemini-North/GMOS on Mauna Kea. The latter used a 1 arcsec slit, the 800 l/mm grating, and a spectral range from approximately 4200 to 7000 Å. The resulting data were reduced using standard IRAF procedures.

2.3 XMM-Newton X-ray observations

MS 0451 was observed by XMM-Newton (observation ID: 0205670101, PI: D. Worrall) on 2004 September 16–17 for a total of 44 ks. We reduced the XMM-Newton/EPIC data using the XMMSAS

v16.1 software package and a pipeline developed in the framework of the XMM-Newton Cluster Outskirts Project (X-COP, Eckert et al. 2017). After performing the standard data reduction steps to extract calibrated event files, we used the XMMSAS tools MOS-FILTER and PN-FILTER to automatically define good-time intervals (unaffected by soft proton flares) of 24 (MOS1), 24 (MOS2), and 19 ks (PN). For more details of this procedure, see Ghirardini et al. (2019, section 2 and fig. 1). These data are used to measure the properties of the baryonic ICM.

An independent, ~ 50 ks Chandra observation provides a high-resolution X-ray view of the cluster core but is not used by us here, since the covered area does not match the extended HST mosaic that is the focus of this paper.

3 METHOD: GRAVITATIONAL LENSING ANALYSIS

3.1 Weak-lensing theory

Weak gravitational lensing is caused by gravitational fields (here created by the massive cluster MS 0451) that deflect light rays emitted by background galaxies, distorting their apparent size and shape. Projecting the cluster's 3D mass distribution along the line of sight yields a 2D surface density $\Sigma(\mathbf{R})$, where $\mathbf{R} = (x, y)$ is a position in the plane of the sky. The cluster's gravitational field causes an isotropic magnification of background galaxies by a factor known as 'convergence'

$$\kappa(\mathbf{R}) = \Sigma(\mathbf{R})/\Sigma_c, \quad (1)$$

where

$$\Sigma_c = (c^2 D_s)/(4\pi G D_1 D_{ls}) \quad (2)$$

is the critical surface mass density for lensing, and D_1 , D_s , D_{ls} are the angular diameter distances from the observer to the lens, from the observer to the source, and from the lens to the source, respectively.

While κ is difficult to observe, a related quantity is more readily measurable. $\Sigma(\mathbf{R})$ also induces a shear distortion

$$\boldsymbol{\gamma} = \gamma_1 + i\gamma_2 = |\boldsymbol{\gamma}|e^{2i\phi}, \quad (3)$$

whose real (imaginary) component is reflected in the apparent elongation of background galaxies along (at 45° to) an arbitrarily oriented real axis. A combination of shear and convergence known as 'reduced shear' $\mathbf{g} \equiv \boldsymbol{\gamma}/(1 - \kappa)$ can be measured from the ellipticities of background galaxies

$$\boldsymbol{\epsilon} = \boldsymbol{\epsilon}_{\text{int}} + G\mathbf{g}, \quad (4)$$

where $\boldsymbol{\epsilon}_{\text{int}}$ is a galaxy's intrinsic ellipticity, and G is a 'shear susceptibility' factor (see Section 3.2.2). Although the unknown intrinsic shapes of galaxies are a dominant source of noise, no bias is introduced if the lensed background galaxies are randomly oriented, $\langle \boldsymbol{\epsilon}_{\text{int}} \rangle = 0$.

3.2 Weak-lensing measurements

3.2.1 Source detection

We identify galaxies in the HST/ACS imaging mosaic from source properties determined with the SEXTRACTOR photometry package (Bertin & Arnouts 1996). To maximize sensitivity to distant (small and faint) galaxies that contain most of the lensing signal, we adopt the 'Hot-Cold' technique (Rix et al. 2004), i.e. we first create a

source catalogue using a ‘cold’ configuration, designed to detect only the brightest objects, and then a second one with a ‘hot’ configuration, optimized to detect faint objects. We then merge the two catalogues, removing any ‘hot’ sources that already have ‘cold’ detections. We also remove any objects located near saturated stars or saturated pixels, using polygonal mask regions defined by hand. Using `SEXTRACTOR` parameters, we assign each source an S/N ratio of detection, defined as $S/N \equiv \text{FLUX_AUTO}/\text{FLUXERR_AUTO}$, and classify objects as galaxies, stars, or spurious features based on their overall brightness (`MAG_AUTO`) and peak surface brightness (`MU_MAX`). The resulting *HST/ACS* catalogue contains 57 281 galaxies.

3.2.2 Weak-lensing shape measurements

We measure the shapes of galaxies detected in the *HST/ACS* images with the `PYRRG` (Harvey et al. 2019) implementation of the shear measurement method of Rhodes, Refregier & Groth (2000, hereafter `RRG`) which was designed to correct the small, diffraction-limited point spread function (PSF) of space-based observatories and has been calibrated on simulated data containing a known shear (Leauthaud et al. 2007).

We model the PSF from stellar images in each exposure. The size and ellipticity of the PSF varies over time, as thermal expansion and contraction of the telescope change the distance between the primary and secondary mirrors and hence the focus. We determine the effective focus of each exposure by comparing the ellipticity of observed stars with models created with the `TINYTIM` ray-tracing software (`TINYTIM`, Rhodes et al. 2007) and stacked to mimic the drizzling of multiple exposures. We interpolate the moments of the net PSF’s shape using a polynomial fitting function.

We then determine the shapes of galaxies to extract the needed weak-lensing information. We measure the second- and fourth-order moments of each galaxy,

$$I_{ij} = \frac{\sum \omega(x, y) x_i x_j I(x, y)}{\sum \omega(x, y) I(x, y)}, \quad (5)$$

$$I_{ijkl} = \frac{\sum \omega(x, y) x_i x_j x_k x_l I(x, y)}{\sum \omega(x, y) I(x, y)}, \quad (6)$$

where I is the intensity recorded in a pixel, ω is a Gaussian weight function included to suppress noise, and the sum is taken over all pixels. Each measured moment is corrected for convolution with the telescope’s PSF. We subsequently calculate each galaxy’s size

$$d = \sqrt{\frac{I_{xx} + I_{yy}}{2}} \quad (7)$$

and ellipticity $\epsilon \equiv \sqrt{\epsilon_1^2 + \epsilon_2^2}$, using the definitions

$$\epsilon_1 = \frac{I_{xx} - I_{yy}}{I_{xx} + I_{yy}}, \text{ and} \quad (8)$$

$$\epsilon_2 = \frac{2I_{xy}}{I_{xx} + I_{yy}}. \quad (9)$$

Applying equation (4), we finally obtain a shear estimator

$$\tilde{g} = C \frac{\epsilon}{G} \quad (10)$$

from each galaxy. Here, the ‘shear susceptibility factor’ G is measured from the global distribution of ϵ and fourth-order moments (`RRG`). The calibration factor C , defined by $1/C = 0.86_{-0.05}^{+0.07}$, is empirically measured from mock *HST* images in the same band and of the same depth (Leauthaud et al. 2007). We note that shapes of very small or faint galaxies are difficult to measure and may

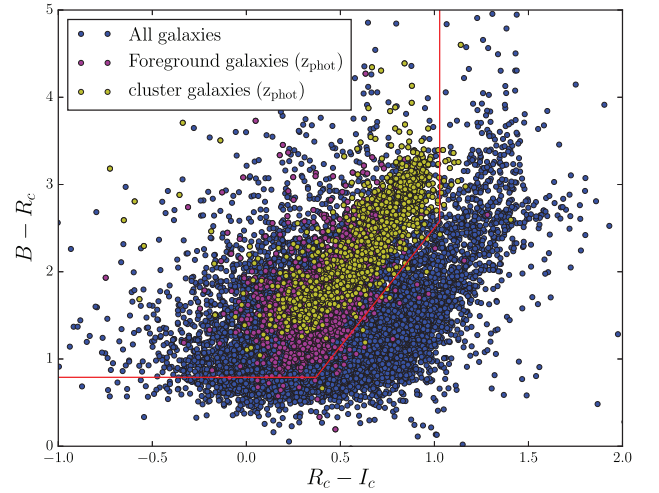


Figure 1. Colour–colour diagram ($B - R_C$ versus $R_C - I_C$) for objects within the *HST/ACS* mosaic of MS 0451. Blue dots represent all objects; magenta and yellow dots are galaxies classified as foreground and cluster galaxies, respectively, based on photometric redshifts. The red solid lines delineate the B , R_C , and I_C colour cuts applied to minimize contamination of the catalogue by unlensed objects.

be biased. As in the calibration tests, we exclude galaxies with size $d < 0.11$ arcsec, $S/N < 4.5$, or unphysical values of $\epsilon > 1$ (which can arise after PSF correction in the presence of noise; for a discussion of this effect, see Jauzac et al. 2012).

3.2.3 Identification of galaxies behind the cluster

The *HST/ACS* galaxy catalogue contains not only background galaxies but also foreground galaxies and cluster members that are not gravitationally lensed by the cluster. These unlensed galaxies dilute the shear signal. We use multicolour ground-based imaging to identify and eliminate them from our catalogue.

Robust photometric redshifts (see Section 2.2.1) can be assigned to the 13 per cent of galaxies in the *HST/ACS* catalogue that are brighter than $R_C = 24$. Based on these redshifts, we remove as likely cluster members all galaxies with $0.48 < z_{\text{phot}} < 0.61$. For galaxies with spectroscopic redshifts, we use the more stringent criterion $0.522 < z_{\text{spec}} < 0.566$ to eliminate cluster members.

For an additional 16 per cent of galaxies, we obtain multicolour photometry in at least the B , R_C , and I_C bands. After testing several criteria chosen elsewhere in the literature to identify foreground and cluster member galaxies (e.g. cuts in $B - R_C$ and $R_C - I_C$, or $B - V$ and $u - B$; see Medezinski et al. 2010, 2018b; Jauzac et al. 2012) we adopt cuts that retain only those galaxies with $(B - R_C) < 0.79$, $(R_C - I_C) > 1.03$, or $(B - R_C) < 2.72(R_C - I_C) - 0.216$ (Fig. 1). After these colour cuts are applied, the photometric redshift distribution of $R_C < 24$ galaxies suggests a remaining contamination from foreground galaxies and cluster members of ~ 4 per cent (Fig. 2, top panel), which is smaller than our statistical error. We shall refer to the combined 30 per cent of galaxies with photometric information as the ‘bright sample’.

From the remaining 70 per cent of galaxies without ground-based photometric information, we next discard the 6 per cent of galaxies that are brighter than $F814W < 24$. These are mainly foreground or cluster member galaxies: in the bright sample, 80 per cent of foreground galaxies and 89 per cent of cluster members have $F814W < 24$, and their combined magnitude distribution peaks

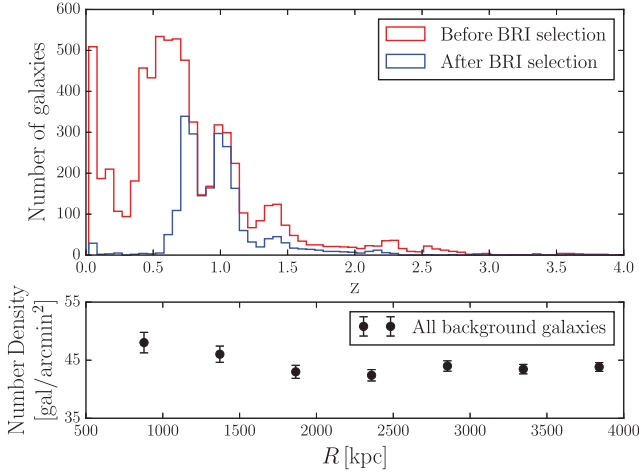


Figure 2. Identification of background galaxies. *Top:* Redshift distribution of all galaxies with spectroscopic or photometric redshifts (red histogram). The blue histogram shows the redshift distribution of galaxies classified as background sources based on B , R_C , and I_C colour–colour selection. *Bottom:* Number density of *all* background galaxies in the final weak-lensing catalogue (including fainter galaxies without observed colours), as a function of their projected distance from the cluster centre.

at $F814W \sim 23$. For the remaining ‘faint sample’ of galaxies, we assign nominal redshifts drawn at random from a distribution $N(z > 0.54) \propto (e^{-z/z_0})^\beta$, with $\beta = 1.8$ and median redshift $z_0 = 0.71$ (Natarajan & Kneib 1997; Gilmore & Natarajan 2009).

Our final weak-lensing catalogue (combining the ‘bright’ and ‘faint’ samples) contains 21 232 background galaxies, at a density of 44 galaxies arcmin^{-2} . Before cuts, the galaxy density shows an excess of ~ 35 galaxies arcmin^{-2} within 1 Mpc of the cluster centre; our selection process removes this excess, leaving an approximately constant density throughout the field (Fig. 2, bottom panel), as expected for an uncorrelated population of background galaxies.

Of our final sample of background galaxies, 10 per cent, 11 per cent, and 79 per cent are selected via cuts in redshift, colour, and magnitude, respectively.

3.3 Strong-lensing constraints

For this analysis, we adopt the best-fitting strong-lensing mass model from Jauzac et al. (2020). We here only give a summary of the strong-lensing mass model, and refer the reader to Jauzac et al. (2020) for more details. The cluster core is modelled using two cluster-scale haloes and 144 galaxy-scale haloes associated with cluster galaxies. All potentials are modelled using pseudo-isothermal elliptical mass distributions (PIEMDs; Kassiola & Kovner 1993; Limousin, Kneib & Natarajan 2005; Elíasdóttir et al. 2007) that are described by seven parameters: position (x, y) , ellipticity e , position angle θ , core radius r_{core} , truncation radius r_{cut} , and velocity dispersion σ .

The 2D surface mass density of each PIEMD is described by

$$\Sigma(R) = \frac{\sigma^2}{2G} \frac{r_{\text{cut}}}{r_{\text{cut}} - r_{\text{core}}} \left(\frac{1}{\sqrt{R^2 + r_{\text{core}}^2}} - \frac{1}{\sqrt{R^2 + r_{\text{cut}}^2}} \right), \quad (11)$$

where the projected radius $R^2 = x^2/(1 + e_\Sigma)^2 + y^2/(1 - e_\Sigma)^2$ is defined by an ellipticity $e_\Sigma \equiv (a - b)/(a + b)$ with semimajor axis a and semiminor axis b (Kassiola & Kovner 1993; Natarajan & Kneib 1997). Note that LENSTOOL reports ellipticity $e \equiv (a^2 - b^2)/(a^2 + b^2)$ and internally converts e into e_Σ during optimization.

Table 1. Best-fitting parameters for the two cluster-scale haloes included in the strong-lensing mass model of MS 0451. Coordinates are given in arcseconds relative to the location of the BCG ($\alpha = 73.545202$, $\delta = -3.014386$). Since the truncation radius of the larger halo is far larger than the effective radius of the strong-lensing regime, it was frozen at 1 Mpc. For more details see Jauzac et al. (2020).

Parameter	Main halo	Second halo
R.A.	$-7.5^{+0.9}_{-1.2}$	$22.3^{+3.1}_{-0.1}$
Dec.	$-2.6^{+0.6}_{-0.7}$	$19.5^{+4.8}_{-0.1}$
σ (km s^{-1})	1001^{+30}_{-25}	810^{+210}_{-670}
e	$0.63^{+0.04}_{-0.03}$	$0.18^{+0.12}_{-0.06}$
θ (deg)	32.2 ± 0.5	147^{+9}_{-16}
r_{core} (kpc)	120^{+10}_{-6}	332^{+60}_{-30}
r_{cut} (kpc)	[1000]	680^{+200}_{-570}

Best-fitting parameters for the two cluster-scale components are listed in Table 1.

Seven cluster galaxies acting as small-scale perturbers of some of the multiple images are independently modelled as individual PIEMDs. The rest of the cluster galaxy population is modelled using scaling relations; to limit the number of free parameters, positions, ellipticities, and position angles of all galaxies are fixed to the respective values of the observed stellar component. The galaxies’ velocity dispersions are scaled from the observed stellar luminosity according to the Faber & Jackson (1976) relation, which describes well the mass in early-type cluster galaxies (Wuyts et al. 2004; Jullo et al. 2007).

The strong-lensing mass model is constrained by 16 systems of multiple images (47 images in total). These include well-known lensed objects such as a submillimetre arc at $z \sim 2.9$ (Borys et al. 2004), five other submillimetre systems (MacKenzie et al. 2014), a triply imaged galaxy (Takata et al. 2003), and six new systems identified with VLT/MUSE, including a quintuple image at $z = 6.7$ in the poorly constrained northern region. The latter has a redshift from VLT/XShooter observations and was previously studied by Knudsen et al. (2016). Five of these systems are spectroscopically confirmed, two of them newly identified by Jauzac et al. (2020) using MUSE observations. The quintuple-image system in combination with the two new systems identified through MUSE observations in the northern region motivated the addition of a second cluster-scale halo in the strong-lensing mass model. Without this second large-scale halo, the geometry of the $z = 6.7$ system cannot be recovered and the root-mean-square (rms) distance between the observed and predicted locations of the multiple images of other systems is unacceptably high at > 1.5 arcsec. Two close groups of cluster galaxies were identified in this region. Adding a third cluster-scale mass halo did not significantly improve the model.

The resulting best-fitting strong-lensing mass model has an rms separation of 0.6 arcsec. The best-fitting parameters of the two main cluster haloes are given in Table 1. Note that the coordinates of the haloes are given in arcseconds relative to the cluster centre, here the BCG ($\alpha = 73.545202$, $\delta = -3.014386$).

3.4 Lensing 2D mass map

We reconstruct the projected (2D) mass distribution using version 7.1 of LENSTOOL² (Jullo et al. 2007; Jullo & Kneib 2009; Jauzac

²Available at <https://projets.lam.fr/projects/lenstool/wiki>.

et al. 2012), whose performance has been quantified on mock *HST* data by Tam et al. (2020). Specifically, we compute the mean mass map from 1700 Markov chain Monte Carlo (MCMC) samples from the posterior. We find consistent but noisier results for the model parameters if we use the Kaiser & Squires (1993) direct-inversion method (see Appendix A).

3.4.1 Mass model

We set the mass distribution in the cluster core to our strong-lensing model (see Section 3.3), which consists of two cluster-scale haloes separated by 237 kpc and seven individually optimized galaxy-scale components.

To extend our analysis from the cluster centre to ~ 3 Mpc, we add a total of 1277 galaxy-scale haloes at the locations of cluster member galaxies, identified via spectroscopic and photometric redshifts over the entire field of view covered by the *HST* mosaic. Each of these is modelled as a PIEMD potential with fixed parameters $r_{\text{core}} = 0.15$ kpc and $r_{\text{cut}} = 58$ kpc, and a velocity dispersion σ that is scaled relative to an $m_K^* = 18.7$ galaxy with $\sigma^* = 163.1$ km s $^{-1}$ using the Faber & Jackson (1976) relation. Throughout the mosaic area outside the strong-lensing region, we then add a grid of masses, whose normalization is allowed to vary and whose resolution is adapted to the local *K*-band luminosity. Following the procedure illustrated in fig. 1 of Jullo & Kneib (2009), we create the multiresolution grid by first drawing a large hexagon over the entire field of view, and then splitting it into six equilateral triangles. If a single pixel inside any of these triangles exceeds a pre-defined threshold in surface brightness, we split that triangle into four smaller triangles. This refinement continues for six levels of recursion, until the brightest parts of the cluster are sampled at the highest resolution, corresponding to a triangle side of 16 arcsec. Once this tessellation process is complete, we place a circular ($e = 0$) PIEMD halo (equation 11) at the centre of each triangle, with a core radius r_c equal to the side length of the triangle, a truncation radius $r_t = 3r_c$, and a velocity dispersion that is left free to vary. To avoid superseding the strong-lens model, we prevent the mass grid from extending into the multiple image region, defined as an ellipse aligned with the cluster core ($a = 44$ arcsec, $b = 34$ arcsec, $\theta = 30^\circ$ counterclockwise with respect to the east–west axis, centred on $\alpha = 73.545202^\circ$, $\delta = -3.0143863^\circ$). We also exclude shear measurements from this region. Our final grid (Fig. 3) for MS 0451 model includes 5570 individual pseudo-isothermal mass distributions.

We optimize masses in the grid using the Bayesian MASSINF algorithm and the Gibbs approach to maximize the likelihood

$$\mathcal{L}_\gamma = \frac{1}{Z_L} \exp\left(-\frac{\chi^2}{2}\right), \quad (12)$$

where the goodness-of-fit statistic is

$$\chi^2 = \sum_{i=1}^M \sum_{j=1}^2 \frac{(\gamma_{j,i} - \gamma_{j,i}^{\text{model}}(\mathbf{R}_i))^2}{\sigma_\gamma^2} \quad (13)$$

(following Schneider, King & Erben 2000),³ where M is the number of background sources and the normalization is

$$Z_L = \prod_{i=1}^M \sqrt{2\pi} \sigma_{\gamma_i}. \quad (14)$$

³Note that LENSTOOL takes inputs in the form of ellipticity $e = (a^2 - b^2)/(a^2 + b^2)$ instead of shear (Jullo et al. 2014), so we use $\gamma^{\text{model}} = 2e^{\text{model}}$.

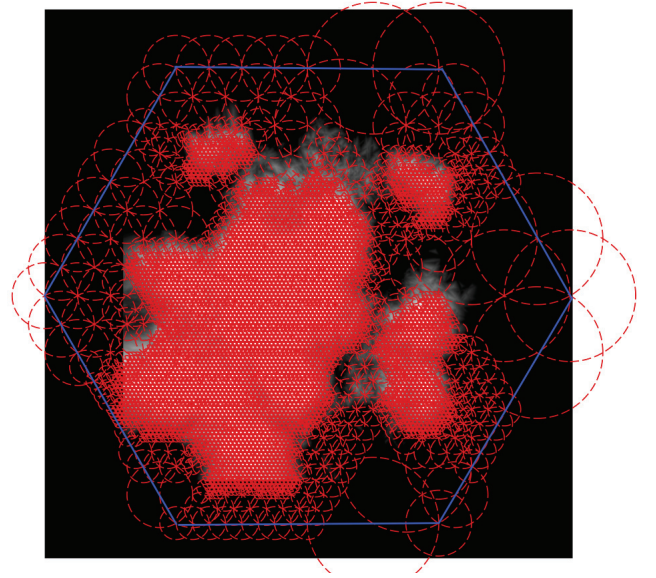


Figure 3. The multiscale grid that determines the maximum spatial resolution of the LENSTOOL mass reconstruction. One RBF is placed at the centre of each circle, with core radius r_c equal to the radius of the circle, and a free mass normalization. The grid is determined from (and shown overlaid upon) the cluster’s *K*-band emission. The blue hexagon covers an area slightly larger than the *HST* field of view.

Equations (13) and (14) involve the statistical uncertainty on each shear measurement, σ_γ . We estimate this for every galaxy as

$$\sigma_\gamma^2 = \sigma_{\gamma,\text{intrinsic}}^2 + \sigma_{\gamma,\text{measurement}}^2, \quad (15)$$

where intrinsic shape noise is constant $\sigma_{\gamma,\text{intrinsic}} = 0.27$ (Leauthaud et al. 2007) and $\sigma_{\gamma,\text{measurement}}$ is derived from the second- and fourth-order moments weighted by the sum of the variance and absolute value of the sky background (Harvey et al. 2019). At each step of the iteration, the 2 per cent most discrepant masses are adjusted.

3.4.2 Uncertainty

We estimate the noise in each pixel of the mass map via bootstrap resampling. To this end, we first select the two $2 \text{ Mpc} \times 2 \text{ Mpc}$ patches of the sky⁴ outside the cluster core that contain the smallest grid cells (and hence the highest *K*-band luminosity peaks) where substructures and filaments are most likely to be located. We choose a random orientation for each shear measurement (ϕ in equation 3) in these two patches of sky, then reconstruct a new mass map using LENSTOOL. Inside an aperture of $r < 480$ kpc, the mean noise level of 100 random realizations is found to be $\langle M \rangle = 2.08 \times 10^{13} M_\odot$, which is non-zero because of LENSTOOL’s positive-definite mass prior, and its rms uncertainty is $\sigma_M = 1.64 \times 10^{13} M_\odot$. We use the latter to normalize the S/N ratios of substructures detected in Section 5.4.

3.5 The lensing-derived 1D density profile

We calculate the cluster’s 1D radial density, $\Sigma(r)$, by (azimuthally) averaging the 2D mass distribution in logarithmically spaced annuli

⁴The two patches of sky used to estimate the level of noise in the weak-lensing map are centred at ($\alpha = 73.644053$, $\delta = -3.012986$) and ($\alpha = 73.426295$, $\delta = -3.089766$).

between 80 kpc and 4 Mpc. To enable a statistically rigorous analysis of this key characteristic, we calculate the full covariance matrix $C_{i,j}$ between measurements in each bin (see Section 3.5.2).

3.5.1 Model comparison

We compare the mean density profile to five models obtained from cosmological simulations: NFW (Navarro, Frenk & White 1996, 1997), generalized NFW (gNFW; Zhao 1996), Einasto (Einasto 1965), Burkert (Burkert 1995), and DK14 (Diemer & Kravtsov 2014). A mathematical definition and description of each halo model is given in Appendix B. We optimize the free parameters of each model using EMCEE (Foreman-Mackey et al. 2013) with a likelihood function

$$\log \mathcal{L} = -\frac{1}{2} \sum_{i,j=1}^{N_{\text{bin}}} (\Sigma_i - \hat{\Sigma}_i) C_{i,j}^{-1} (\Sigma_j - \hat{\Sigma}_j) - \frac{1}{2} N_{\text{bin}} \log(2\pi |C|), \quad (16)$$

where $\hat{\Sigma}$ is the model, $N_{\text{bin}} = 20$ is the total number of radial bins, and $|C|$ is the determinant of the covariance matrix. When fitting the NFW, gNFW, Einasto, and Burkert models, we adopt flat uniform priors for $M_{200} \in [5, 30] \times 10^{14} M_{\odot}$, and $c_{200} \in [1, 10]$. We also adopt a flat prior for the gNFW and Einasto shape parameters, $\alpha \in [0, 3]$ and $\alpha_E \in [0.02, 0.5]$, respectively. For the Burkert model, we use a flat prior for the core radius, $r_{\text{core}} \in [100, 800]$ kpc. For the DK14 model, following More et al. (2016) and Baxter et al. (2017), we use the priors for ρ_s , r_s , r_t , $\log(\alpha)$, $\log(\beta)$, $\log(\gamma)$, and s_e that are listed in table 2 of Chang et al. (2018). Because the location of MS 0451 is so well known from strong-lensing constraints, we omit their miscentring term.

To compare the goodness of fit for models with different numbers of free parameters, we calculate the Bayesian Information Criterion

$$\text{BIC} = -2 \log \mathcal{L} + k \log N_{\text{bin}}, \quad (17)$$

the Akaike Information Criterion

$$\text{AIC} = -2 \log \mathcal{L} + 2k, \quad (18)$$

and the corrected Akaike Information Criterion

$$\text{AICc} = \text{AIC} + \frac{2k(k+1)}{(N_{\text{bin}} - k - 1)}, \quad (19)$$

where k is the number of free parameters. These three information criteria include penalty terms for adding free parameters that make a model more complex. The AIC has a larger penalty term than the BIC; the AICc approaches AIC as N_{bin} increases, but is more robust for small N_{bin} . For all three criteria, lower values indicate a preferred model.

3.5.2 Covariance matrix

When fitting a parametric density profile to the azimuthally averaged mass maps, a first estimate of the uncertainty on the density at a given radius can be obtained by looking at the spread of densities at that radius in the MCMC samples generated by LENSTOOL (or bootstrap sampling, as described in Section 3.4.2). However, the non-local mapping between observable shear and reconstructed mass leads to correlations between adjacent pixels. To fully account for these dependences, we calculate the covariance matrix between radial bins i and j

$$C_{\text{SL+WL}(i,j)} = \frac{1}{N} \sum_{l=1}^N (\Sigma_{l,i} - \langle \Sigma_i \rangle) (\Sigma_{l,j} - \langle \Sigma_j \rangle), \quad (20)$$

where N is the number of MCMC samples generated by LENSTOOL, $\Sigma_{l,i}$ is the surface mass density of the l th sample in the i th spatial bin, and $\langle \Sigma_i \rangle$ is the mean surface mass density of MCMC samples in the i th spatial bin.

Inside the cluster, the dominant sources of statistical noise are the finite number and unknown intrinsic (unlensed) shapes of the background galaxies used in the weak-lensing analysis. Close to the cluster core, our default joint strong- and weak-lensing analysis underestimates the noise, because we fixed the strong-lensing potentials. To account for this effect in the covariance matrix, we reconstruct a separate mass map using only weak-lensing information, i.e. we apply the mass grid and shear measurements in the core region too. The resulting covariance matrix C_{WL} is valid out to $R \sim 1$ Mpc. We then define a combined covariance matrix

$$C_{\text{shape}(i,j)} = \begin{cases} C_{\text{SL+WL}(i,j)} + C_{\text{WL}(i,j)} & \text{for } R_i \text{ and } R_j < 1 \text{ Mpc} \\ C_{\text{SL+WL}(i,j)} & \text{otherwise.} \end{cases} \quad (21)$$

Note that this procedure overestimates $C_{i,j}$ in bins close to the $R \sim 1$ Mpc transition region. However, this effect is negligible in our measurement.

In the outskirts of a cluster, the statistical uncertainty is dominated by LSS projected on to the lens plane. While the specific realization of LSS along the line of sight to MS 0451 is unknown, we can statistically account for its contribution to the covariance matrix $C_{\text{LSS}(i,j)}$ by analysing mock observations of many simulated clusters. In our companion paper (Tam et al. 2020), we generate realizations of LSS along 1000 random lines of sight through the BAHAMAS simulation (McCarthy et al. 2017). We then integrate the 3D mass along the line of sight, weighted by the lensing sensitivity function $\beta(z_1, z_s) = D_{ls}/D_s$ with $\langle z_s \rangle = 0.97$, and interpreting it as a mass distribution in a single lens plane at $z_1 = 0.55$. For each LSS realization, we calculate an effective radial density profile, $\kappa_{\text{LSS}}(R)$, with the same radial binning as applied to our data, which allows us to calculate the full covariance matrix, $C_{\text{LSS}(i,j)}$, describing LSS at different radii.

Finally, we combine the two components of the covariance matrix across the full range of scales,

$$C_{(i,j)} = C_{\text{shape}(i,j)} + C_{\text{LSS}(i,j)}. \quad (22)$$

3.6 Lensing-derived halo shape

We measure the shape of the MS 0451 mass distribution by fitting our reconstructed 2D mass map with elliptical NFW models (eNFW; Oguri et al. 2010). We define the centre of the eNFW halo to be the position of BCG ($\alpha = 73.545202$, $\delta = -3.014386$), and then optimize⁵ its four free parameters (with the allowed range for the parameters within the optimization: $M_{200} \in [0.5, 3] \times 10^{15} M_{\odot}$, $c_{200} \in [1, 10]$, position angle $\phi \in [0, 180]^\circ$, and axial ratio $q = a/b \in [0.1, 0.9]$) to minimize the absolute difference between the observed and modelled mass maps, integrated inside a circular aperture. To measure variations in the cluster shape as a function of radius, we repeat this fit inside circular apertures of varying radii. We perform this fit on every mass model created in LENSTOOL's MCMC chains, and measure the mean and rms values for each free parameter, marginalizing over all others.

⁵We use the L-BFGS-B algorithm (Byrd et al. 1995) from PYTHON's SCIPY.MINIMIZE package <https://docs.scipy.org/doc/scipy/reference/generate/scipy.optimize.minimize.html>.

4 METHOD: X-RAY ANALYSIS

4.1 X-ray imaging analysis

We extract *XMM–Newton* images and exposure maps in the [0.7–1.2] keV band from the cleaned event files. To predict the spatial and spectral distribution of the particle-induced background, we analyse a collection of observations performed with a closed filter wheel (CFW). We compute model particle-background images by applying a scaling factor to the CFW data to match the count rates observed in the unexposed corners of the three EPIC cameras. The images, exposure maps, and background maps for the three detectors are then summed to maximize the S/N ratio.

To determine the thermodynamic properties of the cluster, we extract spectra for the three EPIC detectors in seven annular regions from the centre of the source to its outskirts (radial range 0–4 arcmin). We also extract the spectra of a region located well outside the cluster to estimate the properties of the local X-ray background. The redistribution matrices and effective area files are computed locally to model the telescope transmission and detector response. For each region, the spectra of the three detectors are fitted jointly in XSPEC (Arnaud, Dorman & Gordon 1999) with a model including the source (described as a single-temperature thin-plasma APEC model, Smith et al. 2001, absorbed by the Galactic equivalent neutral hydrogen density N_{H}), the local three-component X-ray background as fitted in the background region, and a phenomenological model tuned to reproduce the spectral shape and intensity of the particle background. The best-fitting parameters of the APEC model (temperature, emission measure, and metal abundance) as a function of radius are obtained by minimizing the C statistic.

4.2 X-ray 1D surface brightness profile

To measure the 1D surface brightness profile of the cluster, we use the azimuthal median technique (Eckert et al. 2015), which allows us to excise contributions from infalling substructures and asymmetries. To this end, we employ Voronoi tessellation to create an adaptively binned surface brightness map of the cluster. For each annulus, we then determine the median value of the distribution of surface brightness values. Uncertainties are estimated by performing 10^4 bootstrap resamplings of the distribution and computing the rms deviation of the measured medians. We measure the local background outside the cluster, where the brightness profile is flat, and subtract it from the source profile. Finally, gas-density profiles are determined by deprojecting the surface brightness profile, assuming spherical symmetry.

We estimate the mass profile of the cluster from the gas-density and temperature profiles, assuming HSE (see Pratt et al. 2019, for a review), i.e. that the pressure gradient balances the gravitational force:

$$\frac{dP_{\text{gas}}}{dr} = -\rho_{\text{gas}} \frac{GM(<r)}{r^2}. \quad (23)$$

The profile of the gravitational mass can thus be inferred from the gas-pressure and density profiles. To solve equation (23), we use the *backwards* approach introduced by Ettori et al. (2019) that adopts a parametric model for the mass profile (here, NFW) and combines it with the density profile computed through the multiscale decomposition technique to predict the pressure (and hence, temperature) as a function of radius. The model temperature profile is projected along the line of sight and corrected for multitemperature structure using scaling as described

in Mazzotta et al. (2004). The projected temperature profile is then compared to the data, and the parameters of the mass model (i.e. mass and concentration) are optimized using MCMC. The integration constant, which describes the overall pressure level at the edge, is left free while fitting and determined on the fly. For more details on the mass reconstruction technique and a careful assessment of systematic effects and uncertainties, see Ettori et al. (2019).

The cumulative gas-mass profile is computed by integrating the gas-density profile over the cluster volume, assuming spherical symmetry:

$$M_{\text{gas}}(<r) = \int_0^r 4\pi r'^2 \rho_{\text{gas}}(r') dr'. \quad (24)$$

Here, $\rho_{\text{gas}} = \mu m_p (n_e + n_{\text{H}})$, with n_e and $n_{\text{H}} = n_e/1.17$ being the number density of electrons and protons, respectively; $\mu = 0.61$ is the mean molecular weight, and m_p is the proton mass. Our procedure directly computes the hydrostatic gas fraction $f_{\text{gas, HSE}}(r) = M_{\text{gas}}(r)/M_{\text{HSE}}(r)$, which traces the virialization state of the gas (Eckert et al. 2019).

5 RESULTS

The mass distribution around MS 0451 (shown in Fig. 4) has a core that is elongated along an axis from south-east to north-west and surrounded by lower mass substructures. An alternative reconstruction using direct inversion following Kaiser & Squires (1993) finds consistent features (Appendix A). However, our primary LENSTOOL method achieves higher spatial resolution in regions containing cluster member galaxies and is more efficient at suppressing noise in regions without them.

The elongated core, which contains two distinct mass peaks, is consistent with an analysis of CFHT/Megacam ground-based weak-lensing measurements (Martinet et al. 2016, and shown in Fig. 5 with magenta contours provided via private communication by N. Martinet). We confirm the existence of several nearby substructures, but our higher S/N data do not show them connected into a filament running south-west to north-east, as hypothesized by Martinet et al. (2016). X-ray emission is detected out to $R = 1.7$ Mpc (Fig. 5).

5.1 Total mass and density profile

Our combined weak- and strong-lensing reconstruction smoothly extends the surface mass density profile outside the multiple-image region (Fig. 6). We measure a projected mass $M(R < 195 \text{ kpc}) = (1.85 \pm 0.87) \times 10^{14} M_{\odot}$, consistent with previous strong-lensing measurements of $1.73 \times 10^{14} M_{\odot}$ (Berciano Alba et al. 2010) and $1.8 \times 10^{14} M_{\odot}$ (Zitrin et al. 2011). At larger radii, our analysis is sensitive for the first time to additional infalling or projected substructures; compared to previous models, based solely on strong-lensing features, we detect excess mass at $R > 3$ Mpc.

Theoretically motivated models to fit the 1D lensing signal (Fig. 6) are described in Appendix B, and their best-fitting parameters are listed in Table 2. For the best-fitting NFW model, we measure a mass $M_{200c} = (1.65 \pm 0.24) \times 10^{15} M_{\odot}$ inside $R_{200c} = 1.99 \pm 0.06$ Mpc, or $M_{500} = (1.13 \pm 0.16) \times 10^{15} M_{\odot}$, and concentration $c_{200} = 3.79 \pm 0.36$. Within the statistical uncertainty, this result is consistent with the ground-based weak-lensing measurement of $M_{200} = (1.44^{+0.33}_{-0.26}) \times 10^{15} M_{\odot}$ for fixed $c_{200} = 4$ (Foëx et al. 2012). The Burkert profile is disfavoured by the BIC and AIC.

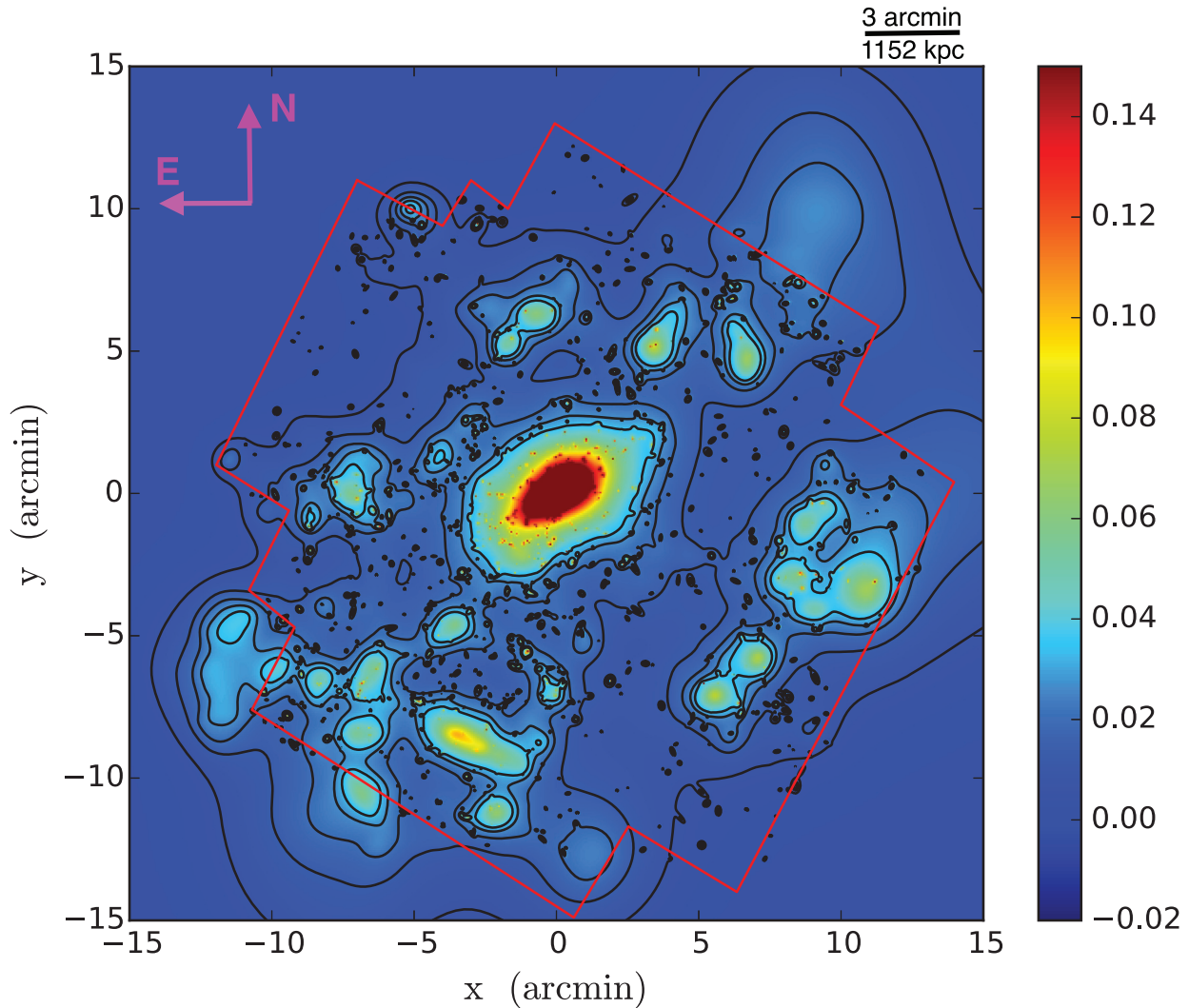


Figure 4. The projected distribution of mass around MS 0451, inferred from our LENTSTOOL strong- and weak-lensing reconstruction and centred on the BCG ($\alpha = 73.545202$, $\delta = -3.0143863$). Colours indicate the projected convergence, κ . Black contours mark the S/N ratio in steps of $1\sigma_{\Sigma}$, measured from bootstrap resampling (see Section 3.4.2). The red polygon delineates the field of view of the *HST*/ACS imaging mosaic.

For the best-fitting gNFW and Einasto models, we find masses and concentrations slightly lower than for NFW. However, their BIC and AIC differ by less than 2 from the NFW ones. Thus, we conclude that our data are unable to distinguish between these three models with statistical significance (as quantified by Burnham & Anderson 2002). We therefore adopt the NFW model as our default in the following analysis.

We note that the additional complexity of the DK14 model captures a splashback-like feature at $R \sim 2$ Mpc (see Appendix C). However, the BIC and AIC both disfavour the DK14 model, and the mentioned feature might be caused by noise or the projection of unrelated LSS along the line of sight.

5.2 Halo shape

We measure the projected shape of MS 0451 by fitting the 2D mass distribution inside a circular aperture with an eNFW model. This approach yields results that are consistent with the previous 1D fit (Section 5.1): for the region inside $R < 3.24$ Mpc, we obtain

$M_{200c} = (1.57 \pm 0.14) \times 10^{15} M_{\odot}$ and $c_{200} = 3.7 \pm 0.4$. The best-fitting axial ratio $q = b/a$ varies as a function of radius, from $q = 0.48 \pm 0.01$ within $R < 649$ kpc to $q = 0.57 \pm 0.03$ within $R < 3.24$ Mpc. The cited statistical uncertainty may be an underestimate because we have neglected correlations between adjacent pixels in our error model of the mass map and, in the cluster core, because of our use of fixed strong-lensing potentials during MCMC parameter search. The axial ratio is consistent with simulations of general clusters (Jing & Suto 2002; Suto et al. 2016; Tam et al. 2020), but smaller than the value of $q = 0.72$ (649 kpc $< R < 974$ kpc) measured from ground-based lensing observations by Soucail et al. (2015). This discrepancy might be explained by the large smoothing kernel used by Soucail and coworkers to reconstruct the mass distribution, which artificially circularizes the data. Indeed, our results more closely resemble those from lensing analyses of large cluster samples, including Oguri et al. (2010), who found $\langle q \rangle = 0.54 \pm 0.04$ for 18 X-ray luminous clusters at $0.15 < z < 0.3$, and Umetsu et al. (2018), who found $\langle q \rangle = 0.67 \pm 0.07$ for the CLASH sample of 20 massive clusters.

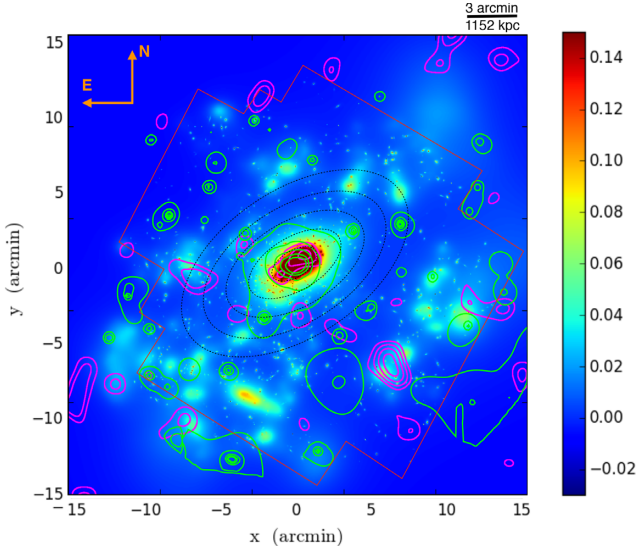


Figure 5. Alternative probes of the mass distribution around MS 0451, overlaid for ease of reference on the colour image from Fig. 4. Magenta contours show weak-lensing measurements from ground-based observations (private communication N. Martinet), starting at $3\sigma_\kappa$ and in steps of $1\sigma_\kappa$, the rms uncertainty on convergence. Green contours show the X-ray surface brightness as recorded by *XMM-Newton*. Black ellipses show the shape of the eNFW model that best fits our LENSTOOL reconstruction within circular apertures of different radii (defined by the semimajor axis).

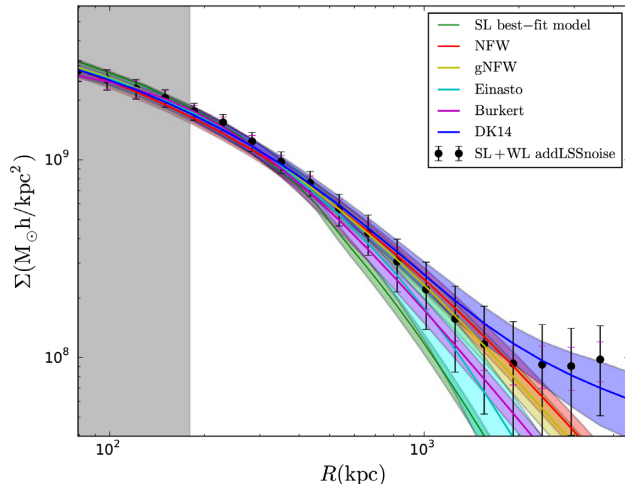


Figure 6. Azimuthally averaged 1D profile of mass in MS 0451 (black data points), from our combined strong- and weak-lensing analysis (Fig. 4). The double error bars show the statistical uncertainty caused solely by the galaxies’ intrinsic shapes (inner error bar) and the uncertainties when line-of-sight substructures are also taken into account (outer). The green curve shows the best-fitting model using only strong-lensing information (Jauzac et al. 2020), extrapolated beyond the multiple-image region (grey shaded area). Solid lines in other colours and their respective shaded areas show the mean and 68 per cent confidence intervals from fits to various models.

At all radii, we find that MS 0451 is elongated roughly along a north-west to south-east axis (Fig 5), with a mean orientation $\sim 31.9^\circ$ counterclockwise from East. The ~ 10 per cent variation in this angle between the inner ($R < 640$ kpc) and outer halo ($R < 3.24$ Mpc) agrees well with typical clusters in both simulations (Despali et al. 2017) and observations (Harvey et al. 2019).

5.3 Baryonic components

5.3.1 Distribution of baryons

To measure the cluster’s electron-density profile, we apply the non-parametric ‘onion peeling’ algorithm (Kriss, Cioffi & Canizares 1983) and the multiscale decomposition technique (Eckert et al. 2016) to the X-ray data (Fig. 7). Both methods assume spherical symmetry, and both yield consistent results. We find the distribution of baryons to be different from the one found by our lensing analysis, in that it shows a constant-density core, flatter than both our lensing results and the distribution of gas in a typical massive cluster from the X-COP low-redshift sample (Ghirardini et al. 2019).

To measure the cluster’s temperature profile, we fit a single-temperature plasma-emission model to the X-ray spectra extracted in six concentric annuli spanning the radial range 0–1.5 Mpc. We find that the temperature of the X-ray emitting gas decreases from ~ 9 in the core to ~ 6 keV in the outskirts, consistent with the ‘universal’ thermodynamic profile of X-COP clusters.

In a separate analysis of the X-ray data assuming HSE and spherical symmetry, we measure a hydrostatic mass of $M_{500,\text{HSE}} = (1.06 \pm 0.35) \times 10^{15} M_\odot$, and a concentration $c_{200,\text{HSE}} = 2.35^{+0.89}_{-0.70}$. The concentration is again lower than the one we obtain in our lensing analysis. Extrapolating the best-fitting model to large radii yields a total mass of $M_{200c,\text{HSE}} = (1.75 \pm 0.75) \times 10^{15} M_\odot$, which inevitably has large uncertainties because the X-ray emission at these radii is faint.

We note that the assumption of HSE may not be appropriate for this cluster. Deeper X-ray imaging and/or constraints on the SZ signal are required to quantify the level of non-thermal pressure support.

The radial entropy profile of the intracluster gas (Fig. 8), obtained by combining the measured spectroscopic temperature with the gas density, is consistent with the 3D entropy model recovered from the *backwards NFW fit* under the assumption of HSE. We find a strong entropy excess in the cluster core, compared with the entropy of the fully relaxed gas calculated from the gravitational-collapse model (Voit 2005). This large entropy excess confirms that MS 0451 does not contain a cool core.

5.3.2 Baryonic-mass fraction

To measure the gas-mass fraction f_{gas} , we first integrate the non-parametric gas profiles (which do not assume HSE) and obtain a total gas mass of $M_{\text{gas},500} = (1.29 \pm 0.15) \times 10^{14} M_\odot$ inside a sphere of radius $R_{500,\text{HSE}} = 1.28 \pm 0.14$ Mpc. Division by the total mass M_{500} of the NFW model, which best fits the lensing data inside a sphere of radius $R_{500} = (1.30 \pm 0.06)$ Mpc, yields $f_{\text{gas},500} = (11.6 \pm 2.1)$ per cent, in good agreement with the result of $f_{\text{gas},500,\text{HSE}} = (12.2 \pm 4.3)$ per cent from our analysis assuming HSE.

To measure the stellar-mass fraction, we use the ratio of stellar mass to light of quiescent galaxies

$$\log_{10}(M_*/L_K) = a z + b, \quad (25)$$

where $a = -0.18 \pm 0.04$ and $b = +0.07 \pm 0.04$ (Arnouts et al. 2007), assuming a Salpeter (1955) initial mass function (IMF).⁶

⁶To convert from Salpeter to a Chabrier (2003) IMF, we adjust the stellar masses by 0.25 dex and find $f_{*,500} = (1.6 \pm 0.24)$ per cent, similar to $f_* \sim 1.5$ per cent measured in the wide-field *HST* COSMOS survey (Leauthaud et al. 2012).

Table 2. Marginalized posterior constraints on cluster model parameters, and the differences between their information criteria and those of the best-fitting NFW model. The information criteria of an NFW are $\text{BIC}_{\text{NFW}} = 13934.50$, $\text{AIC}_{\text{NFW}} = 13932.51$, and $\text{AIC}_{\text{cNFW}} = 13933.22$. Lower values indicate preferred models.

Models	M_{200c} ($10^{14} M_{\odot}$)	c_{200}	Shape parameter	ΔBIC	ΔAIC	ΔAICc
NFW	16.51 ± 2.44	3.79 ± 0.36	–	0	0	0
gNFW	16.10 ± 1.94	4.47 ± 0.47	$\alpha = 0.57 \pm 0.20$	0.51	–0.48	0.02
Einasto	14.32 ± 2.67	4.26 ± 0.50	$\alpha_E = 0.42 \pm 0.11$	0.02	–1.00	–0.51
Burkert	13.62 ± 1.62		$r_{\text{core}} = (230 \pm 20) \text{ kpc}$	7.60	6.60	7.10
DK14 ^a				9.60	4.60	12.90
X-ray	17.47 ± 7.61	$2.35^{+0.89}_{-0.70}$	–			

^aParameters of the DK14 model are excluded from this table for clarity. These are listed in Table C1.

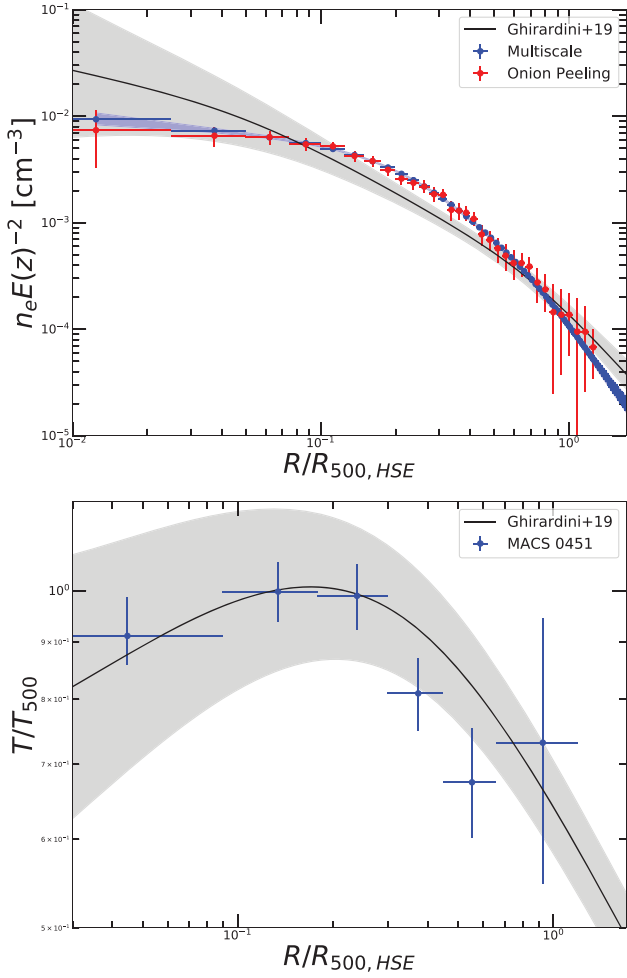


Figure 7. Thermodynamic profiles of MS 0451’s ICM, scaled according to the self-similar model (Kaiser 1986). *Top*: Deprojected electron-density profile of the cluster computed using the onion peeling (red), and multiscale decomposition (blue) methods. *Bottom*: Spectroscopic-temperature profile of the cluster (blue). In both panels, the black curve and grey shaded areas show the mean profile and 1σ scatter of the X-COP sample of massive clusters at low redshift (Ghirardini et al. 2019) for comparison.

Applying this relation to all 1277 cluster member galaxies in the *HST*/ACS mosaic yields a mean value of $\langle M_*/L_K \rangle = 0.94 \pm 0.003$ and a total stellar mass of $M_{*,500} = (3.37 \pm 0.03) \times 10^{13} M_{\odot}$, where the uncertainty is the error of the mean. Note that, although

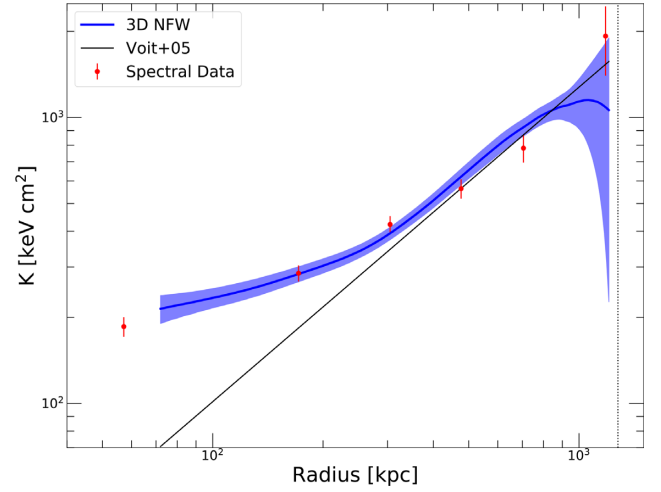


Figure 8. Radial profile of the gas entropy. The red data points are obtained from the measured spectroscopic temperature and the gas density. The blue curve is the model optimized using the *backwards fit* method. For comparison, the black curve shows the gas entropy predicted by the Voit (2005) gravitational-collapse model.

the integration is performed over a cylinder of radius R_{500} rather than a sphere, the result is largely unaffected since the stellar mass is extremely centrally concentrated. We thus adopt a stellar-mass fraction of $f_{*,500} = (3.0 \pm 0.4)$ per cent.

Combining these measurements, we obtain a total baryonic-mass fraction, $f_{b,500} \equiv f_{*,500} + f_{\text{gas},500} = (14.6 \pm 1.4)$ per cent. This value is consistent with the mean cosmic baryon fraction of $f_b = (14 \pm 2)$ per cent measured from the outskirts of clusters at $z < 0.16$ (Mantz et al. 2014), and also with $f_b = (15.6 \pm 0.3)$ per cent from the cosmic microwave background (Planck Collaboration XIII 2016).

5.4 Group-scale substructures

To study the low-density environment of LSS surrounding MS 0451, we subtract the strong-lensing potentials from the LENSTOOL convergence map (Fig 9). Outside the main halo, we detect 14 weak-lensing peaks with $S/N > 3$ integrated within circular apertures of radius $R = 480$ kpc. To determine whether these 14 overdensities are at the redshift of the cluster, we assess the redshift distribution of all galaxies within those apertures that have spectroscopic or (mainly) photometric redshifts (Appendix D). The redshift distribution of galaxies along the line of sight to Substructures 1, 2, 3, 4, 5, and

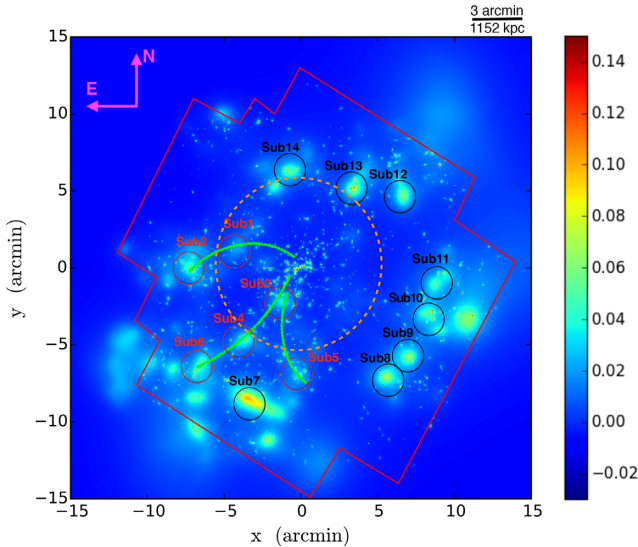


Figure 9. The low-density environment surrounding MS 0451. The colour image shows lensing convergence with SL potentials subtracted: all the remaining signal was constrained by the potential grid and cluster member galaxies. The dashed orange circle has radius $R_{200c} = 1.99$ Mpc. Smaller circles (with radius 480 kpc) mark substructures with a projected mass $> 3\Sigma_M$ inside that aperture; red circles have optical counterparts at the cluster redshift. Green lines suggest the extent and direction of possible large-scale filaments.

6 peaks in the range $0.48 < z < 0.61$. We thus infer that these structures are part of the extended cluster, while all others are projections of structures at other redshifts along our line of sight. The total masses and stellar masses of the six substructures likely to be associated with MS 0451 are listed in Table 3.

Previous ground-based weak-lensing analyses identified only Substructures 1 and 2 (Martinet et al. 2016), or Substructure 2 at a modest 2σ significance (Soucail et al. 2015). Our identification of 12 significant new structures demonstrates the unique ability of space-based imaging to detect weak-lensing signals in low-density environments.

The mismatch between the structures uncovered by our lensing and X-ray analyses is puzzling. Although the depth of the existing *XMM-Newton* data should be sufficient to detect $\sim 10^{14} M_\odot$ haloes, we detect faint X-ray emission only from Substructure 6. Brighter – but misaligned – X-ray emission is seen near Substructures 2 and 5, and between Substructures 3 and 4. The discrepancy might be caused by selection biases in our analyses. On the lensing side, the high mass of Substructure 3 ($M_{\text{tot}} = (1.3 \pm 0.3) \times 10^{14} M_\odot$ inside a 480 kpc aperture) might be erroneous and caused by the structure’s proximity to the cluster core. If the main cluster is imperfectly modelled and subtracted, its residual projected mass could artificially boost the lensing signal of Substructure 3. Indeed, all substructures are closer to the cluster’s major axis than to its minor axis, and hence the lensing signal from all of them could be biased high. Conversely, proximity to the cluster core also results in a high X-ray background, which lowers the S/N ratio of the X-ray emission and thus raises the detection threshold. An alternative, physical explanation for the low X-ray emission from these substructures could be that, within $R_{200m} = (2.51 \pm 0.14)$ Mpc, they probably also lie within the 3D splashback radius of MS 0451 (More, Diemer & Kravtsov 2015) and thus may have already passed through pericentre. Ram-pressure stripping during the passage through the main halo could

have removed much of their hot gas and thus reduced their X-ray luminosity.

5.5 Filaments

5.5.1 Alignments of substructures

Based on the distribution of substructures around MS 0451, we propose that three filaments are connected to the cluster core (shown as green lines in Fig. 9). The first of these possible filaments extends East of the cluster, encompassing Substructures 1 and 2 and containing mean convergence $\langle \kappa \rangle = 0.022 \pm 0.006$. The second points south-east, encompassing Substructures 3, 4, and 6, with $\langle \kappa \rangle = 0.033 \pm 0.007$. The third, finally, turns South, from Substructure 3 to Substructure 5 and also has mean convergence $\langle \kappa \rangle = 0.033 \pm 0.007$. For each of these three candidate filaments, the density contrast exceeds the threshold value of $\kappa = 0.005$ defined in our companion paper (Tam et al. 2020) to identify filaments, and each has a mean *excess convergence* greater than 0.02, even after subtracting the smooth, cluster-scale mass distribution.

All three possible filaments point in a similar direction, close to the main cluster’s south-east/north-west major axis. We detect no substructures in the opposite direction along the same axis (with the possible exception of an unconvincing feature just outside the *HST* mosaic to the north-west). This is strikingly different from the typical distribution of mass in cosmological simulations, which usually show a symmetry of infalling material along both directions of a cluster’s major axis, as the system grows and becomes increasingly elongated as the result of gradual, continuous accretion along filaments.

5.5.2 Aperture multipole moments

Extended structures can also be identified through measurements of aperture multipole moments (AMMs) of the 2D mass distribution (Schneider & Bartelmann 1997), defined as

$$Q^{(n)}(\mathbf{R}) = \int_0^\infty \int_0^{2\pi} |\mathbf{R}' - \mathbf{R}|^{n+1} e^{n i \phi} U(|\mathbf{R}' - \mathbf{R}|) \kappa(\mathbf{R}') d\mathbf{R}' d\phi, \quad (26)$$

where n is the order of the multipole, (R, ϕ) are polar coordinates, and $U(R)$ is a radially symmetric weight function with characteristic scale $R_{\text{max}}^{(n)}$. In tests using mock observations of 10 massive simulated clusters with $M_{200} \sim 10^{15} M_\odot$ at $z = 0.55$ (Tam et al. 2020), we developed a combination of AMMs that highlights the signal from extended filaments,

$$Q \equiv \alpha_0 Q^{(0)} + \alpha_1 Q^{(1)} + \alpha_2 Q^{(2)}, \quad (27)$$

with optimized constants $\alpha_0 = -\alpha_1 = 0.7$, $\alpha_2 = 1$, and $R_{\text{max}}^{(0)} = 1'$, $R_{\text{max}}^{(1)} = R_{\text{max}}^{(2)} = 2'$. This choice of constants enables the detection of narrow filaments with a purity of greater than 75 per cent and completeness in excess of 40 per cent. The quadrupole term, $Q^{(2)}$, is sensitive to linearly extended mass distributions. The dipole term $Q^{(1)}$ fills in the rings which are added around isolated substructures, and the monopole term $Q^{(0)}$ suppresses signal between structures. To avoid the massive cluster haloes that would dominate the contrast in the low-density filaments, we apply this filter to the convergence map of MS 0451 after subtracting the strong-lensing component. The resulting Q map is shown in Fig. 10. We quantify the level of noise by defining σ_Q as the standard deviation of all pixels in the Q map.

Table 3. Confirmed substructure detections in MS 0451 containing cluster member galaxies with a detection S/N of at least 3. Columns show the location of each mass peak, the mean redshift of member galaxies within a 480 kpc aperture, the lensing and stellar masses integrated within the same aperture, and the S/N ratio of detection, using the mean noise level of the mass map (Section 3.4.2).

Substructure	RA	Dec.	$\langle z \rangle$	$M_{\text{tot}} (10^{13} M_{\odot})$	$M_{\text{stellar}} (10^{11} M_{\odot})$	Detection S/N
Sub1	4:54:26.917	−2:59:39.894	0.62	6.17 ± 2.70	6.59 ± 1.79	3.76
Sub2	4:54:39.389	−3:00:32.808	0.58	8.34 ± 3.30	12.70 ± 2.26	5.09
Sub3	4:54:15.278	−3:03:11.620	0.61	13.50 ± 2.90	31.25 ± 3.04	8.23
Sub4	4:54:26.088	−3:05:37.949	0.63	7.17 ± 3.23	5.83 ± 1.56	4.37
Sub5	4:54:11.745	−3:07:30.042	0.55	6.12 ± 2.96	10.55 ± 2.13	3.73
Sub6	4:54:37.972	−3:07:33.134	0.56	8.42 ± 3.37	8.05 ± 2.28	5.13

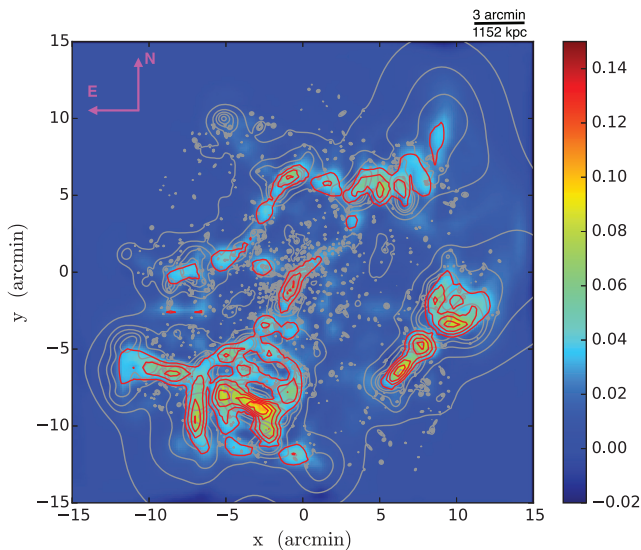


Figure 10. The low-density environment around MS 0451 (Fig. 9), filtered using AMMs to search for extended, filamentary structures. Red contours indicate the S/N ratio, starting at 2 and increasing in steps of 1. Grey contours correspond to the unfiltered mass distribution.

Although the S/N ratios for Q are low, the three possible filaments proposed in Section 5.5 are also highlighted by the AMM filter, with S/N ratios of ~ 2 – 3 . Additional extended structures may exist at other redshifts. In particular, Substructures 8, 9, and 11 might form a linked system at $z \sim 0.7$ behind the cluster (Appendix D).

While the results presented in Fig. 10 are promising, the limiting, single-orbit depth of the available *HST* observations prevents us from drawing firm conclusions from this measurement. Deeper space-based observations obtained by future surveys will increase the density of detected background galaxies and reduce the noise level of this technique.

6 DISCUSSION: INFERRED DYNAMICAL STATE

According to N -body and hydrodynamical simulations (e.g. Ritchie & Thomas 2002; Poole et al. 2006; Nelson et al. 2011), mergers between two clusters (classified by their masses as the primary and secondary component) proceed in five distinct stages: pre-interaction, first core–core interaction, apocentric passage, secondary core accretion, and relaxation. After the first core–core interaction, the gas of the two merging haloes (including the

remaining part of the secondary component’s cool core) moves outward. After the two cores reach maximum separation, the secondary core falls back towards the primary core and is accreted. Finally, the system evolves into a single merger remnant.

Three lines of evidence suggest that MS 0451 is in a post-collision state, approximately 2–7 Gyr after the first core passage of two progenitors that are now approaching second apocentre:

(i) While we observe a bimodal (Section 3.3) and elongated (Section 5.2) distribution of dark matter, we find a spherical distribution of gas with almost constant central density (Section 5.3.1). Such a contrasting configuration is seen in simulations of major mergers 2–7 Gyr after first core–core interaction when, ‘following the merger, the resultant system settles into virial equilibrium sooner than into HSE’ (Poole et al. 2006). This period, in which the gas has not yet had time to fully relax and settle into the gravitational potential of the combined halo, represents the second infall phase before the system’s final relaxation. These findings apply to a wide range of initial conditions regarding the progenitor mass ratio ($M_{\text{primary}}:M_{\text{secondary}} = 1:1, 3:1, 10:1$) and the ratio between the secondary’s transverse velocity and the primary’s circular velocity, a quantity that affects impact parameters ($v_t/v_c = 0.0, 0.15, 0.45$).

(ii) In simulations, merging increases the entropy of gas in the cluster core, leading to a large core radius and low concentration (Ritchie & Thomas 2002), exactly as seen in MS 0451 (Fig. 8).

(iii) The cluster is connected to its large-scale environment through a number of substructures (Section 5.4) and possible filaments (Section 5.5). The distinctly asymmetric distribution of these features differs starkly from that of simulated clusters that grow through smooth, continuous accretion without strong directional preferences, and whose major axes are aligned with filaments in opposite directions.

The dynamical history inferred from our analysis of MS 0451 provides a possible explanation for the star formation history observed in this system. In a comparative study of massive clusters, Moran et al. (2007a, b) used optical and near-UV spectroscopy of passive spirals within 1.5 Mpc of the cluster core to conclude that the star formation in MS 0451 was abruptly quenched at a redshift of $z = 2$, i.e. ~ 5 Gyr before the redshift of observation, consistent with our estimate of the temporal evolution of the merger event. Moran et al. (2007b) ascribe the sudden cessation of star formation to ram-pressure stripping by a particularly dense ICM, as evidenced by MS 0451’s bright, extended X-ray emission observed today. The merger scenario proposed by us lends strong support to this explanation, by adding a contemporaneous second component of intracluster gas, moving at high relative velocity through the cluster core and thus dramatically increasing the ram pressure (cf. Fujita

et al. 1999; Kronberger et al. 2008; Vijayaraghavan & Ricker 2013). Beyond 1.5 Mpc, the passive spirals in MS 0451 show a ‘starvation-like’ gradual cessation of star formation, consistent with secular pre-processing in infalling groups. From their classification of galaxy morphologies in the *HST* imaging data, Moran et al. (2007b) also concluded that passive spirals are all but absent inside the central 600 kpc, having evolved into S0 galaxies. This finding too could be the result of ongoing, enhanced ram-pressure stripping by the current ICM as galaxies fall towards the cluster core. Alternatively, the lack of passive spirals within the inner regions of MS 0451 could be a residual indication of the merging subhalo’s trajectory, as it takes ~ 400 Myr for a galaxy to travel ~ 600 kpc across MS 0451. Our inferences about MS 0451’s merger dynamics thus complement and support the conclusions of previous studies of its star formation history.

7 CONCLUSIONS

We present the first combined strong and weak gravitational lensing analysis of the massive galaxy cluster MS 0451, exploiting the largest mosaic of *HST* imaging around any massive cluster: 41 ACS pointings covering an area of $\sim 20 \times 20$ arcmin² ($\sim 6 \times 6$ Mpc²). The strong-lensing model exploits 16 multiple-image systems, and our weak-lensing analysis uses a catalogue of 20 138 background galaxies (~ 44 arcmin⁻²). We combine these constraints using the LENSTOOL multiscale grid technique.

The reconstructed mass distribution of MS 0451 reveals a bimodal cluster core, elongated along the south-east to north-west axis and surrounded by six substructures, as well as eight weak-lensing peaks created by mass concentrations at other redshifts projected along our line of sight. We find a total mass of the system of $M_{200} = (1.65 \pm 0.24) \times 10^{15} M_{\odot}$ with an NFW concentration of $c_{200} = 3.79 \pm 0.36$ (the gNFW and Einasto models yield similar results, while a Burkert model is disfavoured). Our mass map is consistent with that of the most recent ground-based weak-lensing analysis (Soucail et al. 2015) but resolves three times more substructures at equivalent significance of detection. The mass distribution of MS 0451 becomes more circular at large radii, parametrized by an axial ratio that decreases from $q = b/a = 0.48 \pm 0.01$ within a projected radius of $R = 640$ kpc to $q = 0.57 \pm 0.03$ inside $R = 3.2$ Mpc. A flattening of MS 0451’s density profile at $R \approx 2$ Mpc is well fit by the splashback feature in the DK14 model. However, this model’s additional complexity negates the improved fit according to Bayesian Information Criteria; the aforementioned flattening may thus just be noise or due to LSS projected from other redshifts.

In our X-ray analysis, we measure a baryonic-mass fraction of $f_{b,500} = (14.6 \pm 1.4)$ per cent for MS 0451, consistent with the cosmic baryon fraction (Planck Collaboration XIII 2016), and a total mass of $M_{200} = (1.75 \pm 0.75) \times 10^{15} M_{\odot}$, in good agreement with the lensing estimate. We note though that the assumption of HSE underlying the X-ray mass measurement is unlikely to be valid, given that the cluster’s gas is distributed very differently than the dark matter. The distribution of gas is circularly symmetric, with a constant-density core and low concentration, $c_{200,HSE} = 2.35^{+0.89}_{-0.70}$. We also find a strong excess of gas entropy in the cluster’s central 300 kpc.

Similarly contrasting distributions of gas and dark matter are seen in simulations of post-merger clusters (Poole et al. 2006). The matter distribution observed in MS 0451 suggests that the cluster underwent a major merger ~ 2 – 7 Gyr ago, and that the two dark matter haloes in the centre are now approaching second

apocentre. This merger would have quenched star formation, as ram-pressure from the dense ICM stripped cold gas from cluster member galaxies. Thus the evidence from gravitational lensing, X-ray emission, optical photometry, and spectroscopy all point to a consistent dynamical history.

We find tentative evidence of three filaments extending from the cluster. The distribution of substructures and a noisy measurement of AMMs indicate that all three point in similar directions, between East and South. Interestingly, their distribution is asymmetric, with no counterparts to the North or West. AMMs appear to be a promising method to detect extended filaments. However, our measurements based on single-orbit *HST* data are dominated by shape noise, and deeper space-based observations will be necessary to robustly test this method.

In the next decade, wide-field, space-based surveys at high resolution are planned as part of the *Euclid* and the *Nancy Grace Roman Space Telescope* missions, as the Vera C. Rubin Observatory becomes operational on the ground. MS 0451 will be an ideal target for future studies to characterize infalling substructures along filaments, the timing of star formation processes after a major merger, and the late-time evolution of cluster collisions.

ACKNOWLEDGEMENTS

We would like to thank the referee for giving useful comments and improving our manuscript. We are grateful for constructive discussions of this work with A. Edge, A. Niemiec, and J. Nightingale. We thank N. Martinet for sharing unpublished results from his analysis, allowing a comparison between our two studies. SIT is supported by Van Mildert College Trust PhD Scholarships. MJJ is supported by the United Kingdom Research and Innovation (UKRI) Future Leaders Fellowship ‘Using Cosmic Beasts to uncover the Nature of Dark Matter’ [grant number MR/S017216/1]. RM is supported by a Royal Society University Research Fellowship. DH acknowledges support by the ITP Delta foundation. AR is supported by the European Research Council (ERC-StG-716532-PUNCA) and STFC (ST/N001494/1).

Facilities: This paper uses data from observations GO-9836 with the NASA/ESA *Hubble Space Telescope*, obtained at the Space Telescope Science Institute, which is operated by AURA Inc, under NASA contract NAS 5-26555. All data are available from the telescopes’ archives. This paper uses data from observations (ID: 0205670101) obtained with *XMM-Newton*, an ESA science mission with instruments and contributions directly funded by ESA Member States and NASA. This paper uses data from observations (ID: 06BH34, PI: H. Ebeling) obtained with MegaPrime/MegaCam, a joint project of Canada–France–Hawaii Telescope (CFHT) and CEA/DAPNIA, at the CFHT which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l’Univers of the Centre National de la Recherche Scientifique (CNRS) in France, and the University of Hawaii. This work is based in part on data products produced at TERAPIX, and the Canadian Astronomy Data Centre as part of the CFHT Legacy Survey, a collaborative project of NRC and CNRS. This paper uses data from observations (IDs: 08BH63 and 07BH98, PI: C.-J. Ma) obtained with WIRCam, a joint project of CFHT, the Academia Sinica Institute of Astronomy and Astrophysics (ASIAA) in Taiwan, the Korea Astronomy and Space Science Institute (KASI) in Korea, Canada, France, and the CFHT that is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l’Univers of the Centre National de la Recherche Scientifique of France, and the University of Hawaii.

This work is also based on data collected at Subaru Telescope, which is operated by the National Astronomical Observatory of Japan. This project was also supported by the Science and Technology Facilities Council [grant number ST/P000541/1]. This work used the DiRAC Durham supercomputing facility, which is managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). Its equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, and ST/R002371/1, Durham University, and STFC operations grant ST/R000832/1. DiRAC is part of the UK National e-Infrastructure.

DATA AVAILABILITY

The data underlying this article are available from the corresponding author on reasonable request.

REFERENCES

- Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, *MNRAS*, 426, 2046
- Aragón-Calvo M. A., van de Weygaert R., Jones B. J. T., van der Hulst J. M., 2007, *ApJ*, 655, L5
- Arnaud K., Dorman B., Gordon C., 1999, Astrophysics Source Code Library, record ascl:9910.005
- Arnouts S., Cristiani S., Moscardini L., Matarrese S., Lucchin F., Fontana A., Giallongo E., 1999, *MNRAS*, 310, 540
- Arnouts S. et al., 2007, *A&A*, 476, 137
- Bacon R. et al., 2010, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 7735, Ground-Based and Airborne Instrumentation for Astronomy III. SPIE, Bellingham, p. 773508
- Bahcall N. A., Cen R., 1993, *ApJ*, 407, L49
- Bartelmann M., Maturi M., 2017, *Scholarpedia*, 12, 32440
- Baxter E. et al., 2017, *ApJ*, 841, 18
- Beers T. C., Flynn K., Gebhardt K., 1990, *AJ*, 100, 32
- Berciano Alba A., Koopmans L. V. E., Garrett M. A., Wucknitz O., Limousin M., 2010, *A&A*, 509, A54
- Bertin E., Arnouts S., 1996, *A&AS*, 117, 393
- Bond J. R., Kofman L., Pogosyan D., 1996, *Nature*, 380, 603
- Borys C. et al., 2004, *MNRAS*, 352, 759
- Bradač M. et al., 2006, *ApJ*, 652, 937
- Bradač M. et al., 2008, *ApJ*, 681, 187
- Broadhurst T., Takada M., Umetsu K., Kong X., Arimoto N., Chiba M., Futamase T., 2005, *ApJ*, 619, L143
- Burkert A., 1995, *ApJ*, 447, L25
- Burnham K., Anderson D., 2002, A practical information-theoretic approach. Springer, New York
- Byrd R., Lu P., Nocedal J., Zhu C., 1995, *SIAM J. Sci. Comput.*, 16, 1190
- Chabrier G., 2003, *PASP*, 115, 763
- Chang C. et al., 2018, *ApJ*, 864, 83
- Chiu I.-N., Umetsu K., Sereno M., Ettori S., Meneghetti M., Merten J., Sayers J., Zitrin A., 2018, *ApJ*, 860, 126
- Clampitt J., Jain B., 2016, *MNRAS*, 457, 4135
- Clowe D. et al., 2006, *A&A*, 451, 395
- Contigiani O., Hoekstra H., Bahé Y. M., 2019, *MNRAS*, 485, 408
- De Filippis E., Sereno M., Bautz M. W., Longo G., 2005, *ApJ*, 625, 108
- de Haan T. et al., 2016, *ApJ*, 832, 95
- Despali G., Giocoli C., Bonamigo M., Limousin M., Tormen G., 2017, *MNRAS*, 466, 181
- Diemer B., 2018, *ApJS*, 239, 35
- Diemer B., Kravtsov A. V., 2014, *ApJ*, 789, 1 (DK14)
- Dietrich J. P., Werner N., Clowe D., Finoguenov A., Kitching T., Miller L., Simionescu A., 2012, *Nature*, 487, 202
- Donahue M., Gaskin J. A., Patel S. K., Joy M., Clowe D., Hughes J. P., 2003, *ApJ*, 598, 190
- Donovan D. A. K., 2007, PhD thesis, University of Hawai'i at Manoa
- Dutton A. A., Macciò A. V., 2014, *MNRAS*, 441, 3359
- Ebeling H., Barrett E., Donovan D., Ma C.-J., Edge A. C., van Speybroeck L., 2007, *ApJ*, 661, L33
- Ebeling H., Edge A. C., Henry J. P., 2001, *ApJ*, 553, 668
- Eckert D., Ettori S., Pointecouteau E., Molendi S., Paltani S., Tchernin C., 2017, *Astron. Nachr.*, 338, 293
- Eckert D., Roncarelli M., Ettori S., Molendi S., Vazza F., Gastaldello F., Rossetti M., 2015, *MNRAS*, 447, 2198
- Eckert D. et al., 2016, *A&A*, 592, A12
- Eckert D. et al., 2019, *A&A*, 621, A40
- Einasto J., 1965, Tr. Astrofiz. Inst. Alma-Ata, 5, 87
- Elíasdóttir Á. et al., 2007, preprint ([arXiv:0710.5636](https://arxiv.org/abs/0710.5636))
- Ettori S. et al., 2019, *A&A*, 621, A39
- Evans A. K. D., Bridle S., 2009, *ApJ*, 695, 1446
- Faber S. M., Jackson R. E., 1976, *ApJ*, 204, 668
- Ford H. C. et al., 1996, in Bely P. Y., Breckinridge J. B., eds, Proc. SPIE Conf. Ser. Vol. 2807, Space Telescopes and Instruments IV. SPIE, Bellingham, p. 184
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
- Foëx G., Soucail G., Pointecouteau E., Arnaud M., Limousin M., Pratt G. W., 2012, *A&A*, 546, A106
- Fujita Y., Takizawa M., Nagashima M., Enoki M., 1999, *PASJ*, 51, L1
- Gavazzi R., Mellier Y., Fort B., Cuillandre J.-C., Dantel-Fort M., 2004, *A&A*, 422, 407
- 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
- Ghirardini V. et al., 2019, *A&A*, 621, A41
- Gilmore J., Natarajan P., 2009, *MNRAS*, 396, 354
- Gioia I. M., Maccacaro T., Schild R. E., Wolter A., Stocke J. T., Morris S. L., Henry J. P., 1990, *ApJS*, 72, 567
- Graham A. W., Merritt D., Moore B., Diemand J., Terzic B., 2006, *Astron. J.*, 132, 2685
- Gray M. E., Taylor A. N., Meisenheimer K., Dye S., Wolf C., Thommes E., 2002, *ApJ*, 568, 141
- Harvey D., Massey R., Kitching T., Taylor A., Tittley E., 2015, *Science*, 347, 1462
- Harvey D., Tam S.-I., Jauzac M., Massey R., Rhodes J., 2019, preprint ([arXiv:1911.06333](https://arxiv.org/abs/1911.06333))
- Herbonnet R. et al., 2019, preprint ([arXiv:1912.04414](https://arxiv.org/abs/1912.04414))
- Hoekstra H., 2013, preprint ([arXiv:1312.5981](https://arxiv.org/abs/1312.5981))
- Hoffman S., Avila R., 2018, STScI ISR, 2017-02
- Ilbert O. et al., 2009, *ApJ*, 690, 1236
- Jauzac M., Klein B., Kneib J.-P., Richard J., Rexroth M., Schäfer C., Verdier A., 2020, preprint ([arXiv:2006.10700](https://arxiv.org/abs/2006.10700))
- Jauzac M. et al., 2012, *MNRAS*, 426, 3369
- Jauzac M. et al., 2015, *MNRAS*, 446, 4132
- Jauzac M. et al., 2016, *MNRAS*, 463, 3876
- Jauzac M. et al., 2018, *MNRAS*, 481, 2901
- Jeltema T. E., Canizares C. R., Bautz M. W., Buote D. A., 2005, *ApJ*, 624, 606
- Jing Y. P., Suto Y., 2002, *ApJ*, 574, 538
- Jullo E., Kneib J.-P., 2009, *MNRAS*, 395, 1319
- Jullo E., Kneib J.-P., Limousin M., Elíasdóttir Á., Marshall P. J., Verdugo T., 2007, *New J. Phys.*, 9, 447
- Jullo E., Pires S., Jauzac M., Kneib J.-P., 2014, *MNRAS*, 437, 3969
- Jørgensen I., Chiboucas K., 2013, *AJ*, 145, 77
- Jørgensen I., Chiboucas K., Hiben P., Nielsen L. D., Takamiya M., 2018, *ApJS*, 235, 29
- Kaiser N., 1986, *MNRAS*, 222, 323
- Kaiser N., Squires G., 1993, *ApJ*, 404, 441
- Kaiser N., Wilson G., Luppino G., Kofman L., Gioia I., Metzger M., Dahle H., 1998, preprint ([arXiv:9809268](https://arxiv.org/abs/9809268))
- Kartaltepe J. S., Ebeling H., Ma C. J., Donovan D., 2008, *MNRAS*, 389, 1240
- Kassiola A., Kovner I., 1993, *ApJ*, 417, 450

- Keeton C. R., 2001, preprint ([arXiv:0102341](https://arxiv.org/abs/0102341))
- Kilbinger M., 2015, *Rep. Prog. Phys.*, 78, 086901
- Kimble R. A., MacKenty J. W., O’Connell R. W., Townsend J. A., 2008, in Oschmann J. M., Jr., de Graauw M. W. M., MacEwen H. A., eds, Proc. SPIE Conf. Ser. Vol. 7010, Space Telescopes and Instrumentation 2008: Optical, Infrared, and Millimeter. SPIE, Bellingham, p. 70101E
- Klypin A., Yepes G., Gottlob S., Prada F., Hess S., 2016, *MNRAS*, 457, 4340
- Kneib J.-P., Natarajan P., 2011, *A&AR*, 19, 47
- Knudsen K. K., Richard J., Kneib J.-P., Jauzac M., Clément B., Drouart G., Egami E., Lindroos L., 2016, *MNRAS*, 462, L6
- Kriss G. A., Cioffi D. F., Canizares C. R., 1983, *ApJ*, 272, 439
- Kronberger T., Kapferer W., Ferrari C., Unterguggenberger S., Schindler S., 2008, *A&A*, 481, 337
- Laureijs R. et al., 2011, preprint ([arXiv:1110.3193](https://arxiv.org/abs/1110.3193))
- Leauthaud A. et al., 2007, *ApJS*, 172, 219
- Leauthaud A. et al., 2012, *ApJ*, 744, 159
- Limousin M., Kneib J.-P., Natarajan P., 2005, *MNRAS*, 356, 309
- Ludlow A. D. et al., 2013, *MNRAS*, 432, 1103
- Luppino G. A., Gioia I. M., Hammer F., Le Fèvre O., Annis J. A., 1999, *A&AS*, 136, 117
- Ma C.-J., Ebeling H., Donovan D., Barrett E., 2008, *ApJ*, 684, 160
- MacKenzie T. P. et al., 2014, *MNRAS*, 445, 201
- Mantz A. B., Allen S. W., Morris R. G., Rapetti D. A., Applegate D. E., Kelly P. L., von der Linden A., Schmidt R. W., 2014, *MNRAS*, 440, 2077
- Martinet N. et al., 2016, *A&A*, 590, A69
- Massey R., Kitching T., Richard J., 2010, *Rep. Prog. Phys.*, 73, 086901
- Massey R. et al., 2007, *Nature*, 445, 286
- Massey R. et al., 2014, *MNRAS*, 439, 887
- Mazzotta P., Rasia E., Moscardini L., Tormen G., 2004, *MNRAS*, 354, 10
- McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, *MNRAS*, 465, 2936
- McClintock T. et al., 2019, *MNRAS*, 482, 1352
- Medezinski E., Broadhurst T., Umetsu K., Oguri M., Rephaeli Y., Benítez N., 2010, *MNRAS*, 405, 257
- Medezinski E. et al., 2018a, *PASJ*, 70, 28
- Medezinski E. et al., 2018b, *PASJ*, 70, 30
- Meneghetti M., Jain B., Bartelmann M., Dolag K., 2005, *MNRAS*, 362, 1301
- Merten J., Cacciato M., Meneghetti M., Mignone C., Bartelmann M., 2009, *A&A*, 500, 681
- Merten J. et al., 2011, *MNRAS*, 417, 333
- Miles N. D., Lim P.-L., Bellini A., Grogin N. A., 2018, STSci ISR, 2018-05
- Miyatake H. et al., 2019, *ApJ*, 875, 63
- Molnar S. M., Broadhurst T., 2018, *ApJ*, 862, 112
- Molnar S. M., Hughes J. P., Donahue M., Joy M., 2002, *ApJ*, 573, L91
- Moran S. M., Ellis R. S., Treu T., Smith G. P., Rich R. M., Smail I., 2007a, *ApJ*, 665, 1067
- Moran S. M., Loh B. L., Ellis R. S., Treu T., Bundy K., MacArthur L. A., 2007b, *ApJ*, 671, 1503
- More S., Diemer B., Kravtsov A. V., 2015, *ApJ*, 810, 36
- More S. et al., 2016, *ApJ*, 825, 39
- Natarajan P., Kneib J.-P., 1997, *MNRAS*, 287, 833
- Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Nelson K., Rudd D. H., Shaw L., Nagai D., 2012, *ApJ*, 751, 121
- Ogrea G. A. et al., 2015, *ApJ*, 812, 153
- Oguri M., Takada M., Okabe N., Smith G. P., 2010, *MNRAS*, 405, 2215
- Okabe N., Smith G. P., 2016, *MNRAS*, 461, 3794
- Planck Collaboration XIII, 2016, *A&A*, 594, A13
- Poole G. B., Fardal M. A., Babul A., McCarthy I. G., Quinn T., Wadsley J., 2006, *MNRAS*, 373, 881
- Pratt G. W., Arnaud M., Biviano A., Eckert D., Ettori S., Nagai D., Okabe N., Reiprich T. H., 2019, *Space Sci. Rev.*, 215, 25
- Rehmann R. L. et al., 2019, *MNRAS*, 486, 77
- Rhodes J., Refregier A., Groth E. J., 2000, *ApJ*, 536, 79 (RRG)
- Rhodes J. D. et al., 2007, *ApJS*, 172, 203
- Ritchie B. W., Thomas P. A., 2002, *MNRAS*, 329, 675
- Rix H.-W. et al., 2004, *ApJS*, 152, 163
- Robertson A., Harvey D., Massey R., Eke V., McCarthy I. G., Jauzac M., Li B., Schaye J., 2019, *MNRAS*, 488, 3646
- Romualdez L. J. et al., 2016, preprint ([arXiv:1608.02502](https://arxiv.org/abs/1608.02502))
- Romualdez L. J. et al., 2018, Ground-based and Airborne Instrumentation for Astronomy VII. Proc. SPIE 10702, 107020R
- Rozo E. et al., 2010, *ApJ*, 708, 645
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Sayers J. et al., 2019, *ApJ*, 880, 45
- Schaye J. et al., 2015, *MNRAS*, 446, 521
- Schneider P., Bartelmann M., 1997, *MNRAS*, 286, 696
- Schneider P., King L., Erben T., 2000, *A&A*, 353, 41
- Schwinn J., Jauzac M., Baugh C. M., Bartelmann M., Eckert D., Harvey D., Natarajan P., Massey R., 2017, *MNRAS*, 467, 2913
- Serenio M., Umetsu K., 2011, *MNRAS*, 416, 3187
- Shan H. et al., 2012, *ApJ*, 748, 56
- Shin T.-h., Clampitt J., Jain B., Bernstein G., Neil A., Rozo E., Rykoff E., 2018, *MNRAS*, 475, 2421
- Smith R. K., Brickhouse N. S., Liedahl D. A., Raymond J. C., 2001, *ApJ*, 556, L91
- Soto K. T., Lilly S. J., Bacon R., Richard J., Conseil S., 2016, *MNRAS*, 458, 3210
- Soucail G., Foëx G., Pointecouteau E., Arnaud M., Limousin M., 2015, *A&A*, 581, A31
- Spergel D. et al., 2013, preprint ([arXiv:1305.5422](https://arxiv.org/abs/1305.5422))
- Springel V. et al., 2005, *Nature*, 435, 629
- Starck J.-L., Pires S., Réfrégier A., 2006, *A&A*, 451, 1139
- Suto D., Kitayama T., Nishimichi T., Sasaki S., Suto Y., 2016, *PASJ*, 68, 97
- Takata T. et al., 2003, *PASJ*, 55, 789
- Tam S. I., Massey R., Jauzac M., Robertson A., 2020, preprint ([arXiv:2006.10155](https://arxiv.org/abs/2006.10155))
- Treu T., Ellis R. S., 2015, *Contemp. Phys.*, 56, 17
- Umetsu K., Diemer B., 2017, *ApJ*, 836, 231
- Umetsu K. et al., 2014, *ApJ*, 795, 163
- Umetsu K. et al., 2015, *ApJ*, 806, 207
- Umetsu K. et al., 2018, *ApJ*, 860, 104
- Umetsu K. et al., 2020, *ApJ*, 890, 148
- van Uitert E. et al., 2017, *MNRAS*, 467, 4131
- Vijayaraghavan R., Ricker P. M., 2013, *MNRAS*, 435, 2713
- Voit G. M., 2005, *Rev. Mod. Phys.*, 77, 207
- Warren M. S., Quinn P. J., Salmon J. K., Zurek W. H., 1992, *ApJ*, 399, 405
- Weilbacher P. M., Streicher O., Urrutia T., Jarno A., Pécontal-Rousset A., Bacon R., B’ohm P., 2012, in Radziwill N. M., Chiozzi G., eds, Proc. SPIE Conf. Ser. Vol. 8451, Software and Cyberinfrastructure for Astronomy II. SPIE, Bellingham, p. 84510B
- Weilbacher P. M., Streicher O., Urrutia T., Pécontal-Rousset A., Jarno A., Bacon R., 2014, in Manset N., Forshay P., eds, Astronomical Society of the Pacific Conference Series Vol. 485, Astronomical Data Analysis Software and Systems XXIII, p. 451
- White S. D. M., Rees M. J., 1978, *MNRAS*, 183, 341
- Wuyts S., van Dokkum P. G., Kelson D. D., Franx M., Illingworth G. D., 2004, *ApJ*, 605, 677
- Yess C., Shandarin S. F., 1996, *ApJ*, 465, 2
- Zhao H., 1996, *MNRAS*, 278, 488
- Zitrin A., Broadhurst T., Barkana R., Rephaeli Y., Benítez N., 2011, *MNRAS*, 410, 1939
- Zitrin A. et al., 2013, *ApJ*, 762, L30

APPENDIX A: MASS MAPPING WITH KS93 AND MRLENS

As an alternative to the LENSTOOL multiscale grid method (Section 3.4), we also explore the Kaiser & Squires (1993) direct-inversion method to map the lensing mass. Direct inversion converts

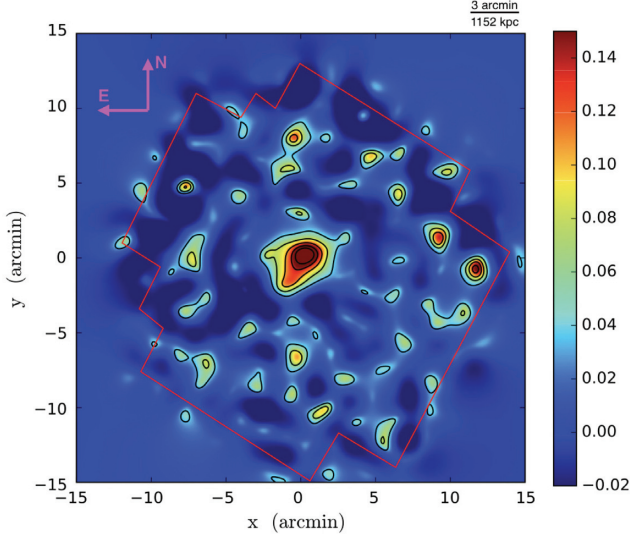


Figure A1. Convergence map of MS 0451 obtained with an alternative method that combines KS93 and MRLens. The field is centred at the location of the BCG, with a red polygon indicating the extent of the *HST* imaging mosaic. Black contours show statistical significance thresholds, starting at $1\sigma_\kappa$ and spaced linearly in units of $1\sigma_\kappa$. Overall the map is consistent with the results of our LENSTOOL method (Fig. 4), but noisier and of lower resolution.

an observed, binned shear map, $\boldsymbol{\gamma}(\mathbf{R})$, into a convergence map, $\kappa(\mathbf{R})$, via their Fourier transforms $\hat{\boldsymbol{\gamma}}(\mathbf{k})$ and $\hat{\kappa}(\mathbf{k})$:

$$\hat{\kappa}(\mathbf{k}) = \frac{1}{2} \left(\frac{k_1^2 - k_2^2}{k_1^2 + k_2^2} \right) \hat{\gamma}_1(\mathbf{k}) + \frac{1}{2} \left(\frac{k_1 k_2}{k_1^2 + k_2^2} \right) \hat{\gamma}_2(\mathbf{k}), \quad (\text{A1})$$

where $\mathbf{k} = (k_1, k_2)$ is the wavevector conjugate to \mathbf{R} . We implement this method using a top-hat window function of radius $R_f = 0.2$ arcmin, $S(\Delta R) = H(R_f - |\Delta R|)$, where $H(x)$ is the Heaviside function, (Merten et al. 2009; Umetsu et al. 2015) to (re)bin the weak-lensing shear catalogues $\boldsymbol{\gamma}$ into 340×340 pixel grids; we also use discrete Fourier transforms. For the region outside the *HST* field of view, we mitigate boundary effects by zero padding out to $105 \text{ arcmin} \times 105 \text{ arcmin}$. Note that this method does not make use of any strong-lensing information.

To denoise the convergence map, we apply the MRLens (multiresolution methods for gravitational lensing) software.⁷ The noisy 2D convergence map is decomposed into six wavelet scales and filtered in 10 iterations. We refer the reader to Starck, Pires & Réfrégier (2006) for a complete description of the MRLens algorithm.

Although the resulting mass map (Fig. A1) is noisier than the LENSTOOL map (Fig. 4), the cluster core is still detected at more than $4\sigma_\kappa$ significance, where σ_κ is the standard deviation of κ over all pixels within the *HST* field of view. The MS 0451 reconstruction also shows an elongation along the south-east to north-west direction, consistent with other analyses (Soucail et al. 2015; Martinet et al. 2016) that use only weak-lensing data. Several weak-lensing peaks are detected at lower statistical significance ($1-3\sigma_\kappa$) than with LENSTOOL, including all confirmed substructures. Overall, the level of noise and spurious peaks (mainly from line-of-sight projections) in this convergence map are similar to what is found in maps

⁷We implement MRLens using the 2017 June 26 version of software available from <https://www.cosmostat.org/software/mrlens>.

reconstructed from mock observations of simulated clusters (see fig. 3 in Tam et al. 2020).

APPENDIX B: RADIAL DENSITY PROFILE MODELS

In Section 3.5, we fit the observed mass distribution in MS 0451 using analytical models of the 3D density; from these we compute the projected surface density by integrating along the line of sight,

$$\Sigma(R) = 2 \int_R^\infty \frac{r\rho(r)}{\sqrt{r^2 - R^2}} dr. \quad (\text{B1})$$

We here briefly summarize the five models investigated in our work.

B1 NFW profile

Numerical simulations (Navarro et al. 1996, 1997) suggest that dark matter haloes can be described by a universal density profile with a two-parameter functional form of

$$\rho_{\text{NFW}} = \frac{\rho_s}{(r/r_s)(1 + (r/r_s))^2}, \quad (\text{B2})$$

where ρ_s and r_s are the characteristic density and radius, respectively. At $r = r_s$, the logarithmic density slope equals the isothermal value, $\text{dln}\rho/\text{dln}r = 2$. For any given cosmology and cluster redshift, this model can also be parametrized in terms of the concentration parameter $c_{200} \equiv R_{200}/r_s$ (here R_{200} is the radius at which the mean density is equal to 200 times the critical density of the Universe, ρ_c) and the halo mass, $M_{200} = (4\pi/3)200\rho_c R_{200}^3$.

B2 Generalized NFW (gNFW) profile

A generalized version of the NFW model, of the form

$$\rho_{\text{gNFW}} = \frac{\rho_s}{(r/r_s)^\alpha (1 + (r/r_s))^{3-\alpha}} \quad (\text{B3})$$

(Zhao 1996). This profile features a power-law-shaped central cusp, $\rho \propto r^{-\alpha}$ and reduces to an NFW model for $\alpha = 1$. Generalizing the approach taken for the NFW model, we describe the gNFW profile with a central slope α , a halo mass M_{200} , and a concentration $c_{200} = r_{200}/(2 - \alpha)r_s$. The radial dependence of the gNFW lensing signal was calculated by Keeton (2001).

B3 Einasto profile

Several *N*-body simulations (e.g. Graham et al. 2006; Dutton & Macciò 2014; Harvey et al. 2015; Klypin et al. 2016) have shown that cold dark matter (CDM) haloes can be best described by the Einasto density profile (Einasto 1965), which is given by

$$\rho_{\text{Einasto}} = \rho_s \exp \left\{ -\frac{2}{\alpha_E} \left[\left(\frac{r}{r_s} \right)^{\alpha_E} - 1 \right] \right\}, \quad (\text{B4})$$

where α_E is the shape parameter describing the steepness of the logarithmic slope. Alternatively, we can express the density profile as

$$\frac{\text{d log } \rho}{\text{d log } r} = -2 \left(\frac{r}{r_s} \right)^{\alpha_E}. \quad (\text{B5})$$

An Einasto profile with $\alpha_E \sim 0.18$ has a similar shape as an NFW profile at a given concentration (Ludlow et al. 2013).

B4 Burkert profile

The dark matter halo proposed by Burkert (1995) is described by a density profile

$$\rho_{\text{Burkert}} = \frac{\rho_{\text{core}}}{(1 + r/r_{\text{core}})(1 + r^2/r_{\text{core}}^2)}, \quad (\text{B6})$$

where ρ_{core} and r_{core} parametrize the density and size of a (constant density) core.

B5 Diemer & Kravtsov profile

The more flexible density profile proposed by DK14 was calibrated using a suite of Λ CDM simulations and is described by two components: (1) collapsed matter, modelled by a truncated Einasto profile (Einasto 1965), and (2) infalling material, modelled by a power-law function. The complete model is given by

$$\rho(r) = \rho_{\text{Einasto}}(r) \times f_{\text{trans}}(r) + \rho_{\text{infall}}(r) \quad (\text{B7})$$

$$f_{\text{trans}}(r) = \left[1 + \left(\frac{r}{r_t} \right)^\beta \right]^{-\gamma/\beta}, \quad (\text{B8})$$

$$\rho_{\text{infall}}(r) = \frac{\rho_m b_e}{\Delta_{\text{max}}^{-1} + \left(\frac{r}{5r_{200m}} \right)^{-s_e}}, \quad (\text{B9})$$

where $\Delta_{\text{max}} = 10^3$, and the transition term f_{trans} captures the steepening of the profile around a truncation radius r_t . The shape parameters γ and β determine the steepness of the profile and how quickly the slope changes, respectively. For the infalling material, the power-law profile (which decreases with radius, since $s_e > 0$) approaches the mean density of the Universe, ρ_m , at large radii. Δ_{max} is introduced to avoid a spurious contribution towards the centre of the cluster. We adapted the publicly available code COLOSSUS (Diemer 2018) to calculate the DK14 density profile.

APPENDIX C: SPLASHBACK RADIUS

N -body simulations by DK14 show that density profiles of dark matter haloes exhibit a sharp steepening close to the virial radius. This feature depends on the accretion history of the cluster and results from an absence of particles orbiting beyond the radius of second turnaround. It gives us a physically motivated definition for the boundary of dark matter haloes. Here, we investigate the splashback feature of MS 0451 by fitting its density profile with a DK14 profile (Section 5.1). The marginalized posterior constraints are listed in Table C1 where we employ the biweight estimators of Beers, Flynn & Gebhardt (1990) for the centre and dispersion of the marginalized posterior distributions (e.g. Sereno & Umetsu 2011; Umetsu et al. 2014; Chiu et al. 2018).

Table C1. Marginalized posterior constraints on the DK14 model.

Parameter	Constraint
ρ_s ($10^3 M_\odot \text{ kpc}^{-3}$)	598.32 ± 6.41
r_s (Mpc)	0.33 ± 0.12
r_t (Mpc)	1.76 ± 1.34
$\log \alpha$	-0.78 ± 0.39
$\log \beta$	0.87 ± 0.45
$\log \gamma$	0.70 ± 0.45
b_e	1.96 ± 1.00
s_e	1.75 ± 0.54

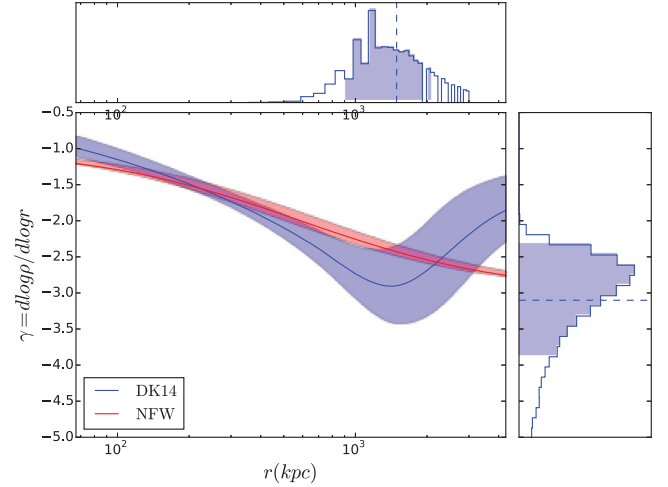


Figure C1. Radial gradient of the total mass distribution of MS 0451 from fitted NFW (red) and DK14 (blue) models. Solid lines show the mean inferred values; shaded regions show 68 per cent confidence intervals. The upper and right panels show the posterior probability distributions of the splashback radius r_{sp} and the gradient at the splashback radius $\gamma(r_{\text{sp}})$. Dashed lines and shaded regions indicate the mean and 68 per cent confidence intervals, respectively.

Following More, Diemer & Kravtsov 2015, we define the splashback radius r_{sp} as the location of a local minimum in the logarithmic slope of the density profile, $\gamma \equiv d \log \rho / d \log r$. Fig. C1 shows the mean and 68 per cent confidence intervals of γ inferred from the DK14 fit, together with the posterior probability distribution of the splashback radius, and the posterior probability distribution of $\gamma(r_{\text{sp}})$. The biweight central location of $\gamma(r_{\text{sp}})$ is -3.10 ± 0.74 at $r_{\text{sp}} = (1.49 \pm 0.57)$ Mpc. This is not significantly different from the value of $\gamma \sim -2.4$ for the best-fitting NFW model at this radius, and the Bayesian and Akaike Information Criteria disfavour the DK14 model because of its increased complexity. Furthermore, our best-fitting r_{sp} value is lower than predictions from cosmological simulations and other observational analysis (e.g. Contigiani, Hoekstra & Bahé 2018).

It is possible that a true splashback feature lies close to (or outside) the edge of the *HST* field of view, where ‘noise’ in the form of lensing signal from projected substructures exceeds the lensing signal of the cluster and is correlated between radial bins. Hence, similarly to the findings obtained by Umetsu & Diemer (2017) for a sample of 16 clusters, our measurements for MS 0451 place only a lower limit on the splashback radius, i.e. $r_{\text{sp}} > 1.49$ Mpc.

APPENDIX D: REDSHIFT DISTRIBUTION OF DETECTED WEAK-LENSING PEAKS

The summed probability density functions (PDFs) of photometric redshifts z_{phot} for all galaxies within $R = 480$ kpc of each substructure are shown in Fig. D1. For comparison, we also show the redshift distribution of galaxies observed in *HST* imaging in a blank region of sky (the COSMOS field), within the same passband and to the same depth, normalized to the number density of galaxies in our catalogue that have photometric redshifts (a higher fraction of COSMOS galaxies have photometric redshifts, particularly at high redshift, due to the deeper Subaru imaging). The 1σ scatter in the redshift distribution of COSMOS galaxies, calculated using

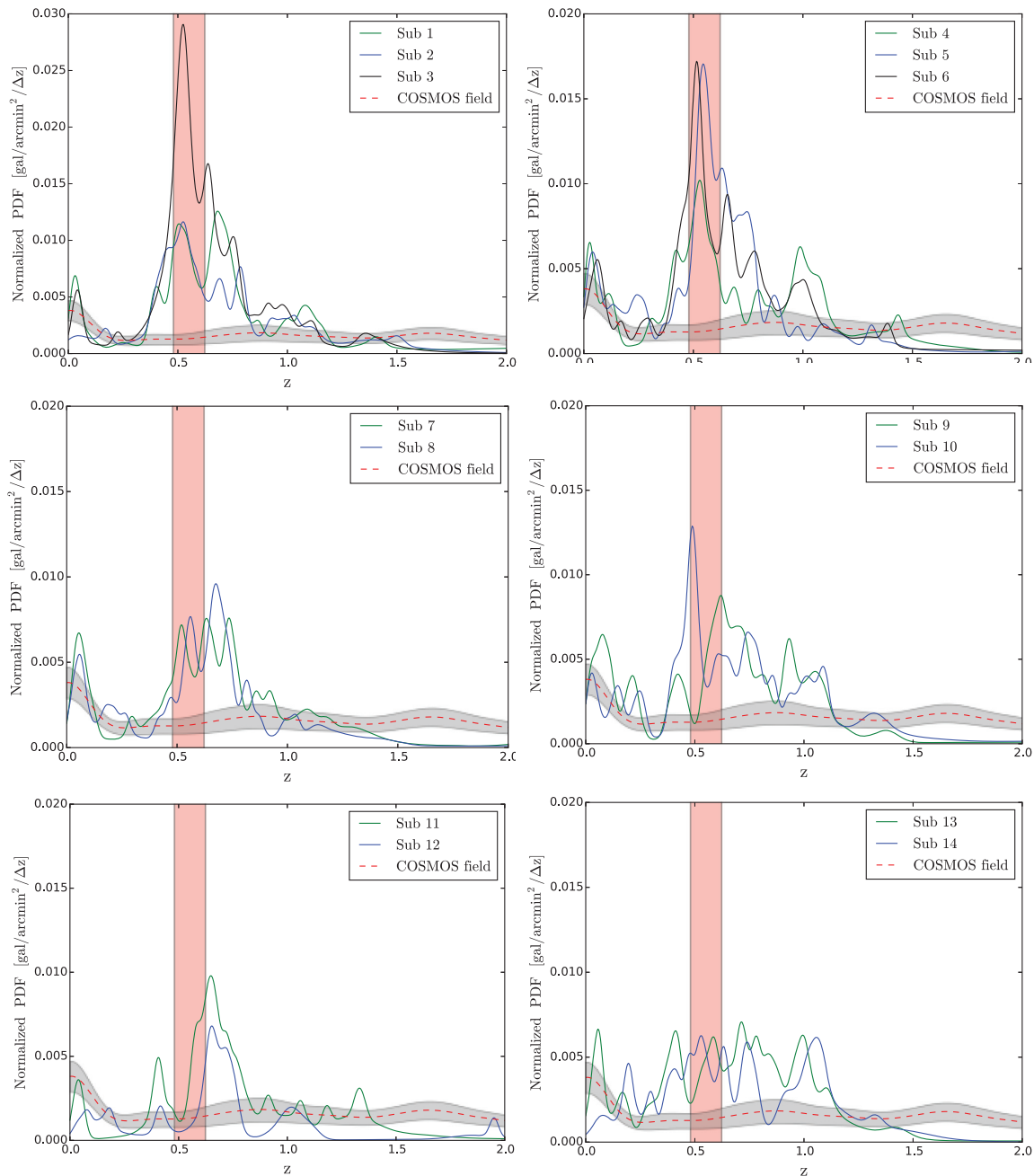


Figure D1. Normalized PDF of photometric redshifts z_{phot} for all galaxies within a circular aperture ($R = 480$ kpc) of each of the 14 detected weak-lensing peaks. Within the typical uncertainties of photometric redshifts, the galaxy overdensities at $0.48 < z_{\text{phot}} < 0.61$ (vertical red bands) are consistent with being at the redshift of MS 0451. For comparison, the red dashed line shows the redshift distribution of galaxies detected in comparable *HST* imaging of a blank patch of sky (the COSMOS field) and the associated 1σ scatter (grey band).

the bootstrap method, is unchanged by the higher precision of z_{phot} measurements that use many more colours.

Substructures 1, 2, 3, 4, 5, and 6 are dominated by galaxies at the same redshift as the main cluster, and are therefore likely to be physically associated. By contrast, Substructures 7–14 (marked by black circles in Fig. 9) are dominated by galaxies at different redshifts, or at a mixture of redshifts, and are therefore implau-

sible as parts of MS 0451. We note that Substructures 8, 9, and 11 might be a linked system behind the cluster ($z \sim 0.7$) and even appear as an extended mass distribution in the AMM map (Fig. 10).

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