Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **467,** 3801–3826 (2017)
Advance Access publication 2017 February 7

doi:10.1093/mnras/stx326

CLUMP-3D: three-dimensional lensing and multi-probe analysis of MACS J1206.2−0847, a remarkably regular cluster

Mauro Sereno,[1,2]★ Stefano Ettori,[1,3] Massimo Meneghetti,[1,3] Jack Sayers,[4] Keiichi Umetsu,[5] Julian Merten,[6] I-Non Chiu[5] and Adi Zitrin[7]

[1]*INAF – Osservatorio Astronomico di Bologna, via Ranzani 1, I-40127 Bologna, Italy*
[2]*Dipartimento di Fisica e Astronomia, Università di Bologna, viale Berti Pichat 6/2, I-40127 Bologna, Italy*
[3]*INFN, Sezione di Bologna, viale Berti Pichat 6/2, I-40127 Bologna, Italia*
[4]*Division of Physics, Mathematics and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA*
[5]*Institute of Astronomy and Astrophysics, Academia Sinica, PO Box 23-141, Taipei 10617, Taiwan*
[6]*Department of Physics, University of Oxford, Keble Road, Oxford OX1 3RH, UK*
[7]*Physics Department, Ben-Gurion University of the Negev, PO Box 653, Be'er-Sheva 84105, Israel*

Accepted 2017 February 3. Received 2017 February 2; in original form 2016 November 12

ABSTRACT

Multi-wavelength techniques can probe the distribution and the physical properties of baryons and dark matter in galaxy clusters from the inner regions out to the peripheries. We present a full three-dimensional analysis combining strong and weak lensing, X-ray surface brightness and temperature, and the Sunyaev–Zel'dovich effect. The method is applied to MACS J1206.2−0847, a remarkably regular, face-on, massive, $M_{200} = (1.1 \pm 0.2) \times 10^{15}\,\mathrm{M}_\odot\,h^{-1}$, cluster at $z = 0.44$. The measured concentration, $c_{200} = 6.3 \pm 1.2$, and the triaxial shape are common to haloes formed in a Λ cold dark matter scenario. The gas has settled in and follows the shape of the gravitational potential, which is evidence of pressure equilibrium via the shape theorem. There is no evidence for significant non-thermal pressure and the equilibrium is hydrostatic.

Key words: gravitational lensing: weak – galaxies: clusters: general – galaxies: clusters: individual: MACS J1206.2−0847 – galaxies: clusters: intracluster medium.

1 INTRODUCTION

Clusters of galaxies correspond to the densest regions to undergo gravitational relaxation in the hierarchical scenario of structure formation, where the Universe is dominated early by cold dark matter and later by dark energy in the form of a cosmological constant. Small density fluctuations rise and grow in the early Universe under the influence of gravity to create the massive, dark matter-dominated structures we observe today. Clusters are the most massive and latest objects to near viral equilibrium.

This prominent role makes clusters de facto laboratories to test cosmology, astrophysics and fundamental physics (Voit 2005). The accurate measurement of their mass and intrinsic properties is crucial to astrophysical investigation (Meneghetti et al. 2010; Postman et al. 2012; Rasia et al. 2012; Limousin et al. 2013).

We want a complete picture of the cluster properties and, in the spirit of the scientific method, we want to compare observations to theory. However, we face the problem that theoretical models and numerical simulations of formation and evolution of cosmic structure are naturally expressed in terms of mass and three-dimensional properties, whereas we have access to only partial information. We can easily measure only projected quantities.

Clusters of galaxy can be observed at very different wavelengths to provide independent data sets, from X-ray surface brightness (SB) and spectral observations of the intracluster medium (ICM), to gravitational lensing (GL) observations of the total mass distribution to the Sunyaev–Zel'dovich effect (SZe) in the radio band.

Weak lensing (WL) analyses are in principle independent of the equilibrium state but can measure only the projected mass. To infer the true mass, we have to deproject the lensing maps assuming a cluster shape, which we actually ignore. The assumption of spherical symmetry can introduce biases in the measurement of mass and concentration (Oguri et al. 2005; Sereno & Zitrin 2012). An unbiased analysis has to take into account shape and orientation (Gavazzi 2005; Limousin et al. 2013).

X-ray observations open a window on the cluster thermodynamics, but hypotheses on the hydrostatic equilibrium (HE) are needed to further the analysis and measure the mass. This is not optimal. By assuming HE, we can strongly bias the mass measurement (Rasia et al. 2012).

Simplifying assumptions can bias the results. This is critical also in the context of large present and future surveys (Laureijs et al. 2011; Pierre et al. 2016; Planck Collaboration XXVII 2016b),

★ E-mail: mauro.sereno@oabo.inaf.it

when reliable masses of selected clusters are used to calibrate large samples through scaling relations (Sereno & Ettori 2016).

Furthermore, when strong working hypotheses are enforced a priori, we cannot investigate any more fundamental cluster properties. The cluster shape shows how matter aggregates from large-scale perturbations (West 1994; Jing & Suto 2002). Assessing the equilibrium status is crucial to determine evolution and mechanisms of interaction of baryons and dark matter (Lee & Suto 2003; Kazantzidis et al. 2004).

In the era of precision astronomy, we have to use all the information and analyse it coherently. The distribution of three-dimensional shapes of a population of astronomical objects can be obtained by inverting statistical samples of projected maps, see Hubble (1926), Noerdlinger (1979), Binggeli (1980), Binney & de Vaucouleurs (1981), Fasano & Vio (1991), de Theije, Katgert & van Kampen (1995), Mohr et al. (1995), Basilakos, Plionis & Maddox (2000), Cooray (2000), Thakur & Chakraborty (2001), Alam & Ryden (2002), Ryden (1996), Plionis, Basilakos & Tovmassian (2004), Paz et al. (2006) and Kawahara (2010).

Individual clusters can be investigated with multi-probe approaches, see Zaroubi et al. (1998), Reblinsky (2000), Doré et al. (2001), Fox & Pen (2002), Puchwein & Bartelmann (2006), Mahdavi & Chang (2011) and Morandi et al. (2012). Joint X-ray and SZ analyses can probe the gas physics in detail, e.g. the occurrence and mass distribution of infalling gas clumps (Eckert et al. 2016; Tchernin et al. 2016).

Cluster maps can be deprojected combining different constraints of the cluster potential. However, despite the growing interest, this topic is still in its infancy. See Limousin et al. (2013) for a recent review on the asphericity of galaxy clusters.

CLUster Multi-Probes in Three Dimensions (CLUMP-3D) is a project to get the unbiased intrinsic properties of galaxy clusters. By exploiting rich data sets ranging from X-ray, to optical, to radio wavelengths, the mass and concentration can be determined together with the intrinsic shape and equilibrium status of the cluster as required by precision astronomy. The inversion problem is tackled with a Bayesian inference method.

This project builds on a series of methods developed by the same authors. De Filippis et al. (2005) and Sereno et al. (2006) first studied a sample of 25 clusters with X-ray and SZe data to find signs of a quite general triaxial morphology. The method was later generalized in a Bayesian framework in Sereno, Ettori & Baldi (2012). The triaxial analysis of strong lensing (SL) and WL was introduced in Sereno & Umetsu (2011), and a method combining lensing, X-ray and SZe was presented in Sereno et al. (2013). The latest application to Abell 1689 (Umetsu et al. 2015) showed that the cluster is elongated and not in equilibrium. The triaxial analysis reduces the apparent tension with theoretical predictions.

As a test case, we consider MACS J1206.2−0847 (hereafter MACS1206), an X-ray luminous cluster at $z = 0.439$ originally discovered in the Massive Cluster Survey (MACS; Ebeling, Edge & Henry 2001; Ebeling et al. 2009). MACS1206 was included in the CLASH (Cluster Lensing And Supernova survey with Hubble; Postman et al. 2012) sample on the basis of being massive and relatively relaxed. Zitrin et al. (2012b) carried out a detailed SL analysis of the cluster exploiting CLASH *HST* (*Hubble Space Telescope*) imaging and Very Large Telescope (VLT)/VIsible Multi Object Spectrograph (VIMOS) spectroscopic observations. Based on an SL analysis, Eichner et al. (2013) found evidence for tidally stripped haloes of the cluster galaxies.

Umetsu et al. (2012) performed an accurate mass reconstruction of the cluster from a combined WL distortion, magnification and SL analysis of wide-field Subaru BVR_cI_cz' and *HST* imaging. Alternative lensing analyses of the full CLASH sample were later presented in Merten et al. (2015) and Umetsu et al. (2016).

Biviano et al. (2013) exploited a rich data set of ∼600 spectroscopic redshifts, obtained as part of the VLT/VIMOS programme, to constrain the mass, the velocity-anisotropy and the pseudo-phase-space density profiles using the projected phase-space distribution. The overall agreement among different studies further suggests that the cluster is in a relaxed dynamical state.

The deep optical coverage is completed by X-ray data and ancillary measurements of the SZe collected with Bolocam, operating at 140 GHz at the Caltech Submillimeter Observatory (Czakon et al. 2015). This makes MACS1206 an ideal target for a detailed lensing and multi-wavelength analysis.

The paper is as follows. Sections 2 and 3 are devoted to the triaxial parametric modelling of the matter and gas distribution, respectively. In Section 4, we list the observational constraints. In Section 5, we present the data sets used for the analysis. The Bayesian inference method is introduced in Section 6. Results are presented and discussed in Section 7. A check for systematics is performed in Section 8. Section 9 is devoted to the conclusions. Appendix A summarizes the basics of the projection of an ellipsoidal volume density. Appendix B details how we approximated the shape of the gravitational potential.

1.1 Notations and conventions

Throughout the paper, the framework cosmological model is the concordance flat Λ cold dark matter (ΛCDM) universe with matter density parameter $\Omega_M = 1 - \Omega_\Lambda = 0.3$, Hubble constant $H_0 = 100\,h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ with $h = 0.7$ and power spectrum amplitude $\sigma_8 = 0.82$. $H(z)$ is the redshift-dependent Hubble parameter and $E_z \equiv H(z)/H_0$.

The ellipsoid is our reference geometric shape for haloes and distributions. We may further distinguish triaxial ellipsoids with three axes of different lengths, oblate or prolate spheroids, or spherical haloes.

O_Δ denotes a global property of the cluster measured within a region that encloses a mean overdensity of Δ times the critical density at the cluster redshift, $\rho_{cr} = 3H(z)^2/(8\pi G)$. For an ellipsoidal halo, this region is the ellipsoid of semi-major axis ζ_Δ and volume $(4\pi/3)q_{mat,1}q_{mat,2}\zeta_\Delta^3$, where $q_{mat,1}$ and $q_{mat,2}$ are the axial ratios. $O_{sph,\,\Delta}$ indicates that the quantity is computed in a spherical region of radius $r_{sph,\,\Delta}$.

The alternative subscript Δm indicates that the overdensity region is computed with respect to the mean cosmological matter density.

Throughout the paper, 'log' is the logarithm to base 10 and 'ln' is the natural logarithm.

The parameters of our model are listed in Table 1. Coordinates and derived quantities are listed in Table 2. The ellipticity ϵ of the cluster refers to its projected two-dimensional shape in the plane of the sky; the line-of-sight elongation $e_\parallel$ quantifies the extent of the cluster along the line of sight, see Appendix A.

Typical values and dispersions of the parameter distributions are usually computed as biweighted estimators (Beers, Flynn & Gebhardt 1990) of the marginalized posterior distributions.

2 TRIAXIAL MATTER DISTRIBUTION

The spherical cow is a humorous metaphor but it is very far from an overly simplified model. If we cannot tell the head from the tail, if we do not know the ground, if we are not even sure the cow is

Table 1. List of the parameters of the regression scheme. Units and description are in columns 2 and 3, respectively. The default priors used in the regression scheme are listed in column 4. $\mathcal{U}$ is the uniform prior; δ is the Dirac delta function for parameters set to fixed values. In column 5, we refer to the section where the parameter is introduced and in column 6 we refer to the main equations involving the parameters. Units of $\zeta_c\,\mathrm{kpc}\,h^{-1}$ mean that if e.g. the truncation radius is x times the core radius, then $\zeta_t = x$.

Symbol	Units	Description	Default prior	Section	Equation
		Total matter distribution			
M_{200}	$10^{15}\,\mathrm{M}_\odot\,h^{-1}$	Total mass within the ellipsoid of mean density $200\rho_{cr}$	$\mathcal{U}(10^{-1}, 10)$	2.2	13
c_{200}		Concentration parameter	$\mathcal{U}(0.1, 20)$	2.2	
		Matter shape and orientation			
$q_{mat,1}$		Minor-to-major axial ratio of the total matter distribution	$\mathcal{U}(0.1, 1)$	2.1	
$q_{mat,2}$		Minor-to-major axial ratio of the total matter distribution	$\mathcal{U}(q_{mat,1}, 1)$	2.1	
$\cos\vartheta$		Cosine of the inclination angle of the ellipsoid major axis	$\mathcal{U}(0, 1)$	2.1	
φ		Second Euler angle	$\mathcal{U}(-\pi/2, \pi/2)$	2.1	
ψ		Third Euler angle	$\mathcal{U}(-\pi/2, \pi/2)$	2.1	
		Gas shape			
$q_{ICM,\,1}$		Minor-to-major axial ratio of the ICM distribution	$\mathcal{U}(q_{mat,1}, 1)$	3.1	
$q_{ICM,\,2}$		Intermediate-to-major axial ratio of the ICM distribution	$\mathcal{U}(q_{ICM,1}, 1)$	3.1	
		Gas distribution			
n_0	cm^{-3}	Central scale density of the distribution of electrons	$\mathcal{U}(10^{-6}, 10)$	3.2	19
ζ_c	$\mathrm{kpc}\,h^{-1}$	Ellipsoidal core radius of the gas distribution	$\mathcal{U}(0, 10^4)$	3.2	19
ζ_t	$\zeta_c\mathrm{kpc}\,h^{-1}$	Ellipsoidal truncation radius of the gas distribution	$\mathcal{U}(0, 10)$	3.2	19
β		Slope of the gas distribution	$\mathcal{U}(0, 3)$	3.2	19
η		Inner slope of the gas distribution	$\mathcal{U}(0, 3)$	3.2	19
γ_{ICM}		Outer slope of the gas distribution	$\mathcal{U}(0, 3)$	3.2	19
		Gas temperature			
T_0	keV	Typical temperature of the gas	$\mathcal{U}(10^{-2}, 10^2)$	3.2	20, 22
ζ_{cT}	$\zeta_c\mathrm{kpc}\,h^{-1}$	Ellipsoidal truncation radius of the temperature profile	$\mathcal{U}(0, 10)$	3.2	20, 22
a_T		Intermediate slope of the temperature profile	$\delta(0)$	3.2	20, 22
b_T		Steepness of the temperature profile	$\delta(2)$	3.2	20, 22
c_T		Outer slope of the temperature profile	$\mathcal{U}(0, 3)$	3.2	20, 22
T_{cc}	keV	Temperature of the cool core	$\mathcal{U}(10^{-2}, T_0)$	3.2	20, 21
ζ_{cc}	$\zeta_c\mathrm{kpc}\,h^{-1}$	Ellipsoidal radius of the cool core	$\mathcal{U}(0, 10)$	3.2	20, 21
α_{cc}		Steepness of the cool core	$\delta(1.9)$	3.2	20, 21

a cow, make it spherical and we can still study the system without committing cancerous errors. But if we have some tools to tell the head from the side and we are sure that the cow is a cow, we can make it ellipsoidal to have a better insight.

For very irregular systems, the spherical approximation is still the better option. But if the galaxy cluster is well shaped, with the ellipsoidal model we can determine the properties of the system unbiased by shape and orientation.

The ellipsoidal model can be an improvement but it is not the final deal. Even in regular clusters, the eccentricity and the orientation of the matter distribution can change with the radius (Suto et al. 2016; Vega-Ferrero, Yepes & Gottlöber 2016). The gas distribution changes with radius too. Axial ratios and orientation of the ellipsoidal distribution have to be meant as effective.

2.1 Shape and orientation

The main assumption of our triaxial modelling is that the total mass distribution of galaxy clusters is approximately ellipsoidal. This is the natural extension of the spherical modelling. The matter isodensities are approximated as a family of concentric, coaxial ellipsoids. We assume that the ellipsoids are self-similar, i.e. ellipsoids are concentric and share the same axial ratio and orientation. The halo shape is determined by the axial ratios, which we denote as $q_{mat,1}$ (minor-to-major axial ratio) and $q_{mat,2}$ (intermediate-to-major axial ratio). The eccentricity is $e_i = \sqrt{1 - q_i^2}$.

Cosmological simulations showed that self-similarity is not strict (Kazantzidis et al. 2004). At $z = 0$, inner regions of simulated haloes are less spherical than outer regions (Suto et al. 2016; Vega-Ferrero et al. 2016). The radial dependence gradually changes with time (Suto et al. 2016). At $z = 1$, the axial ratios $q_{mat,1}$ increases towards the inner regions. On the other hand, the axial ratio always steeply decreases in the outskirts due to filamentary structure around the haloes.

The precise assessment of radial variations depends on how axial ratios are measured. Several methods have been proposed and some of them can be well defined only in simulations. This may not be the case of actual observed clusters. Furthermore, radial variations can be below the accuracy reached by present-day observational campaigns.

The self-similar ellipsoidal distribution with fixed axial ratios can then provide a good description of galaxy clusters (Jing & Suto 2002; Bonamigo et al. 2015). Measured axial ratios have to be intended as effective radially weighted averages.

2.1.1 *Flat distribution*

As reference prior distribution, we considered a nearly flat distribution covering the range $q_{min} \leq q_{mat,1} \leq 1$ and $q_{mat,1} \leq q_{mat,2} \leq 1$ (Sereno & Umetsu 2011). We assume that the marginalized probability $p(q_{mat,1})$ and the conditional probability $p(q_{mat,2}|q_{mat,1})$ are constant. In formulae,

$$p(q_{mat,1}) = 1/(1 - q_{min}) \tag{1}$$

for $q_{min} < q_{mat,1} \leq 1$ and zero otherwise, and

$$p(q_{mat,2}|q_{mat,1}) = 1/(1 - q_{mat,1}) \tag{2}$$

Table 2. List of coordinates and derived quantities. Symbols and descriptions are in columns 1 and 2, respectively. Quantities may refer to the total matter (mat), to the gas (ICM) and the gravitational potential (Φ). In this paper, the component will be indicated with a subscript, e.g. q_1 is the generic axial ratio whereas $q_{\rm mat,1}$, $q_{\rm ICM,\,1}$ and $q_{\Phi,\,1}$ refer to the total matter, the gas and the gravitational potential. In column 3, we refer to the section where the parameter is introduced and in column 4 we refer to the main equations involving the parameters.

Symbol	Description	Section	Equation
	3D coordinates		
ζ	Ellipsoidal radial coordinate	2.1	14
r	Spherical radial coordinate	2.1	
	2D coordinates		
ξ	Elliptical radial coordinate	2.1	
θ_ξ	Elliptical angular coordinate	2.1	
R	Circular radial coordinate	2.1	
θ_R	Circular angular coordinate	2.1	
	3D shape parameters		
$\mathcal{T}$	Triaxial parameter	3.1	16
e_i	Eccentricity of the ith axis	2.1	
	3D radii		
ζ_Δ	Semi-major axis of the ellipsoid of average density equal to Δ times the critical density $\rho_{\rm cr}$	1	
$\zeta_{\Delta{\rm m}}$	Semi-major axis of the ellipsoid of average density equal to Δ times the cosmological matter density	1	
$r_{\rm sph,\,\Delta}$	Radius of the sphere of average density equal to Δ times the cosmological critical density	1	
	2D projected parameters		
$q_\perp$	Axial ratio of the projected ellipse	Appendix A	A1
ϵ	Projected ellipticity	Appendix A	A5
θ_ϵ	Orientation angle of the projected ellipse	Appendix A	A6
	2D NFW parameters		
$\kappa_{\rm s}$	Scale convergence	4.1	24
$\xi_\perp$	Projected elliptical scale radius	4.1	24
	Line-of-sight parameters		
f	Line-of-sight projection	Appendix A	A10
$e_\parallel$	Elongation	Appendix A	A11

for $q_{\rm mat,1} \leq q_{\rm mat,2} \leq 1$ and zero otherwise. We fixed $q_{\rm min} = 0.1$. The flat distribution is compatible with very triaxial clusters ($q_{\rm mat,1} \lesssim q_{\rm mat,2} \ll 1$), which are preferentially excluded by N-body simulations. Here and in the following, p denotes the probability density, which can be larger than 1.

2.1.2 *N-body prior*

A population of ellipsoidal, coaligned, triaxial clusters fits well the relaxed clusters of galaxies produced in N-body simulations (Jing & Suto 2002). The distributions of the minor-to-major and intermediate-to-major axial ratios can be well described by simple functional forms. Most of the dependences on mass and redshift can be expressed in terms of the peak height, ν, from the spherical collapse theory (Bonamigo et al. 2015).

As a shape prior based on N-body simulations, we consider the results of Bonamigo et al. (2015), who analysed relaxed haloes from the Millennium XXL simulation and provided statistically significant predictions in the mass range above $3 \times 10^{14}\,{\rm M}_\odot\,h^{-1}$ at two redshifts ($z = 0$ and 1). Unrelaxed clusters were removed by selecting only haloes for which the offset between the most bound particle and the centre of mass of the particles enclosed by the ellipsoid was less than 5 per cent of their virial radius.

Bonamigo et al. (2015) found that the minor-to-major axial ratio, after rescaling in terms of the peak height, $\tilde{q}_1 = q_1 \nu^{0.255}$, follows a lognormal distribution,

$$p(\ln \tilde{q}_1) \sim \mathcal{N}(\mu = -0.49, \sigma = 0.20), \tag{3}$$

where $\mathcal{N}$ is the Gaussian distribution; the conditional probability for the rescaled ratio $\tilde{q}_2 = (q_2 - q_1)/(1 - q_1)$ can be written as a beta distribution,

$$P(\tilde{q}_2|q_1) \sim \mathcal{B}(\alpha, \beta), \tag{4}$$

where

$$\beta = 1.389 q_1^{-1.685} \tag{5}$$

$$\alpha = \beta/\left[1/(0.633 q_1 - 0.007) - 1\right]. \tag{6}$$

Alternatively, we also consider the priors based on Jing & Suto (2002), where the distribution of q_1 is approximated as

$$p(q_1) \sim \mathcal{N}(q_\mu / r_{q_1}, \sigma_{\rm s} = 0.113), \tag{7}$$

with $q_\mu = 0.54$ and

$$r_{q_1} = (M_{\rm vir}/M_*)^{0.07\Omega_{\rm M}(z)^{0.7}}, \tag{8}$$

with M_* the characteristic non-linear mass at redshift z and $M_{\rm vir}$ the virial mass. The conditional probability of q_2 can be expressed as

$$p(q_1/q_2|q_1) = \frac{3}{2(1 - r_{\rm min})}\left[1 - \frac{2q_1/q_2 - 1 - r_{\rm min}}{1 - r_{\rm min}}\right], \tag{9}$$

for $q_1/q_2 \geq r_{\rm min} \equiv \max[q_1, 0.5]$, whereas is null otherwise.

2.1.3 Random orientation

The orientation of the halo is established by three Euler's angles, ϑ, φ and ψ, with ϑ quantifying the inclination of the major axis with respect to the line of sight.

A priori, we considered a population of randomly oriented clusters with

$$p(\cos\vartheta) = 1 \tag{10}$$

for $0 \leq \cos\vartheta \leq 1$,

$$p(\varphi) = \frac{1}{\pi} \tag{11}$$

for $-\pi/2 \leq \varphi \leq \pi/2$, and

$$p(\psi) = \frac{1}{\pi} \tag{12}$$

for $-\pi/2 \leq \psi \leq \pi/2$.

2.2 Density profile

The Navarro–Frenk–White (NFW) density profile embodies the most relevant features of matter haloes (Navarro, Frenk & White 1996, 1997),

$$\rho_{\rm NFW} = \frac{\rho_{\rm s}}{(\zeta/\zeta_{\rm s})(1+\zeta/\zeta_{\rm s})^2}, \tag{13}$$

where ζ is the ellipsoidal radius and $\zeta_{\rm s}$ is the scale radius. In the coordinate frame oriented along the principal axes of the ellipse

$$\zeta^2 = \frac{x_1^2}{q_1^2} + \frac{x_2^2}{q_2^2} + x_3^2. \tag{14}$$

The NFW density profile can be described by two parameters. M_{200} is the mass within the ellipsoid,

$$M_{200} \equiv (800\pi/3)\rho_{\rm cr}\, q_{\rm mat,1} q_{\rm mat,2} \zeta_{200}^3. \tag{15}$$

The concentration is $c_{200} \equiv \zeta_{200}/\zeta_{\rm s}$. Ellipsoidal mass and concentration follow the same relations as found in numerical *N*-body simulations for spherically averaged haloes (Corless, King & Clowe 2009).

3 TRIAXIAL GAS DISTRIBUTION

3.1 Shape

In our modelling, the gas distribution is ellipsoidal and coaligned with the matter. This is supported by the observed morphologies of galaxy clusters in the X-ray band, which are nearly elliptical (Kawahara 2010).

In the ideal case of a halo in perfect HE, the assumptions of ellipsoidal and fixed matter distribution and of ellipsoidal and fixed gas distribution are mutually exclusive. If the matter is ellipsoidal, the potential that originates from it cannot be ellipsoidal. The gas distribution traces the gravitational potential, and turns rounder in the outer regions. On the other hand, given an ellipsoidal gas density, the gravitational potential is ellipsoidal too and can turn unphysical for extreme axial ratios, with negative density regions or unlikely configurations.

However, the variation of eccentricity of the potential of an ellipsoidal mass distribution in the radial range covered by observations is small (Lee & Suto 2003, 2004) and the ellipsoidal approximation for the gas is suitable in the inner regions or when small eccentricities are considered (Buote & Humphrey 2012).

In Appendix B, we detail how the shape of the gravitational potential of an ellipsoidal halo can be approximated. The ratio of eccentricities of potential (e_Φ) to matter ($e_{\rm mat}$) is nearly constant up to the length-scale, with $e_{\Phi,\,i}/e_{{\rm mat},\,i} \simeq 0.7$ for $i = 1, 2$. Furthermore, the variation in projected ellipticity of the potential is usually smaller than the observational error on the measured ellipticity of the X-ray SB map.

Physical processes perturb the only hypothetical perfect equilibrium. Radiative cooling, turbulence or feedback mechanisms strongly affect the gas shape, which can show a distinctly oblate configuration towards the central regions compared to the underlying dark matter potential shape (Lau et al. 2011). Outside the core, radiative processes can make the ICM distribution rounder. The effect of filamentary accretion or merging events can have dramatic effects in the outskirts.

These mechanisms can be effective. In a conservative scheme, the overall triaxiality of the gas cannot be strictly related to the underlying shape of the dark matter potential. Total matter distribution and gas have to be modelled independently. We then explore different scenarios to relate gas and total matter distributions.

In the less informative one ($q_{\rm ICM} \geq q_{\rm mat}$), we assume that total matter and gas are shaped as coaligned ellipsoids with fixed, but different eccentricity. The angles ϑ, φ and ψ set the orientation of both distributions. In this scheme, the gas is rounder than the total matter, i.e. $q_{{\rm ICM},\,1} \geq q_{\rm mat,1}$ and $q_{{\rm ICM},\,2} \geq q_{\rm mat,2}$. The prior on $q_{{\rm ICM},\,1}$ is then similar to that of $q_{\rm mat,2}$ in the case of the flat distribution for the matter axial ratios.

In the second scheme ($\mathcal{T}_{\rm mat} = \mathcal{T}_{\rm ICM}$), the matter and gas distributions share the same triaxiality parameter,

$$\mathcal{T} = (1-q_2^2)/(1-q_1^2). \tag{16}$$

If two distributions have the same triaxiality, the misalignment angle between the orientations in the plane of the sky is zero (Romanowsky & Kochanek 1998). The axial ratios of the gas distribution, $q_{{\rm ICM},\,i}$, can be expressed in terms of the corresponding axial ratios of the matter distributions as

$$q_{{\rm ICM},i} = \sqrt{1-(e_{\rm ICM}/e_{\rm mat})^2(1-q_{{\rm mat},i}^2)}. \tag{17}$$

Being $\mathcal{T}_{\rm mat} = \mathcal{T}_{\rm ICM}$, then $e_{{\rm ICM},\,1}/e_{\rm mat,1} = e_{{\rm ICM},\,2}/e_{\rm mat,2} = e_{\rm ICM}/e_{\rm mat}$. The above assumptions limit the number of free axial ratios to three: $q_{\rm mat,1}$ and $q_{\rm mat,2}$ for the matter and $q_{{\rm ICM},\,1}$ for the gas. $q_{{\rm ICM},\,2}$ is determined by $\mathcal{T}_{\rm mat}$ and $q_{{\rm ICM},\,1}$,

$$q_{\rm ICM,2} = \sqrt{1-\frac{1-q_{\rm ICM,1}^2}{\mathcal{T}_{\rm mat}^2}}. \tag{18}$$

Under these hypotheses, total matter and gas have different projected ellipticities and elongations but share the same orientation in the plane of the sky, θ_ϵ. This is in agreement with what observed in MACS1206, where the centroid and the orientation of the SB map coincide with those of the projected mass distribution inferred from lensing, see Section 7.6.

In the third scheme ($q_{\rm ICM} = q_\Phi$), we assume that the gas follows the potential. This scheme formally conflicts with our assumption that both total matter and gas are distributed as ellipsoids with fixed eccentricity. In fact, the potential of an ellipsoidal matter distribution is not ellipsoidal. However, in most cases, the isopotential surfaces can be still well approximated as ellipsoids whose axial ratios vary slightly with the radius, see Appendix B. We can then consider the axial ratios of the potential at a typical radius and assume that the gas distribution is ellipsoidal with those axial ratios at each radius.

In the following, we measure the effective axial ratios of the potential at $\zeta = \zeta_{200}/3 \sim \zeta_{2500}$, the radius better probed by X-ray observations. Alternatively, we can also consider $\zeta = 2\zeta_{200}/3 \sim \zeta_{500}$.

3.2 Thermodynamics

We describe the thermodynamical properties of the ICM in terms of the distribution of the gas and its temperature.

The electronic density was modelled with the parametric profile (Vikhlinin et al. 2006; Ettori et al. 2009),

$$n_{\rm e} = n_0 \left(\frac{\zeta}{\zeta_{\rm c}}\right)^{-\eta} \left[1 + \left(\frac{\zeta}{\zeta_{\rm c}}\right)^2\right]^{-3\beta/2+\eta/2} \left[1 + \left(\frac{\zeta}{\zeta_{\rm t}}\right)^3\right]^{-\frac{\gamma_{\rm ICM}}{3}}, \tag{19}$$

where n_0 is the central electron density, $\zeta_{\rm c}$ is the core elliptical radius, $\zeta_{\rm t}(>\zeta_{\rm c})$ is the tidal radius, β is the slope in the intermediate regions, and η and $\gamma_{\rm ICM}$ are the inner and outer slope, respectively.

For the intrinsic temperature, we used (Vikhlinin et al. 2006; Baldi et al. 2012)

$$T_{\rm 3D}(\zeta) = T_0 t_{\rm cc}(\zeta) t_{\rm out}(\zeta), \tag{20}$$

where T_0 sets the temperature scale, $t_{\rm cc}(r)$ describes the temperature decline in the cluster cool core,

$$t_{\rm cc}(\zeta) = \frac{x + T_{\rm cc}/T_0}{1+x}, \quad x = (\zeta/\zeta_{\rm cc})^{\alpha_{\rm cc}}, \tag{21}$$

and $t_{\rm out}(r)$ parametrizes the temperature profile outside the central cool region,

$$t_{\rm out} = \frac{(\zeta/\zeta_{cT})^{-a_T}}{[1 + (\zeta/\zeta_{cT})^{b_T}]^{c_T/b_T}}; \tag{22}$$

the radius ζ_{cT} is the transition radius. Some of the parameters describing the profile at the truncation are degenerate. We fixed $b_T = 2$.

4 OBSERVATIONAL CONSTRAINTS

The intrinsic properties of the cluster can be obtained by deprojection of the observed maps.

4.1 Lensing

Lensing analysis can provide the projected mass density of the cluster,

$$\Sigma_{\rm mat} = \int_{\parallel} \rho_{\rm mat}\, {\rm d}l, \tag{23}$$

where the subscript $\parallel$ denotes integration along the line of sight.

Ellipsoidal 3D haloes project as elliptical 2D profiles, see Appendix A. For the NFW halo, the convergence κ, i.e. the surface mass density in units of the critical density for lensing, $\Sigma_{\rm cr} = c^2 D_{\rm s}/(4\pi G\, D_{\rm d}\, D_{\rm ds})$, where $D_{\rm s}$, $D_{\rm d}$ and $D_{\rm ds}$ are the source, the lens and the lens–source angular diameter distances, respectively, can be written as

$$\kappa_{\rm NFW}(x) = \frac{2\kappa_{\rm s}}{1-x^2}\left[\frac{1}{\sqrt{1-x^2}}\,{\rm arccosh}\left(\frac{1}{x}\right) - 1\right], \tag{24}$$

where x is the dimensionless elliptical radius, $x \equiv \xi/\xi_{\perp}$. The elliptical isodensities are characterized by the ellipticity $\epsilon_{\rm mat}$ and the direction angle $\theta_{\epsilon,\,\rm mat}$. The central strength $\kappa_{\rm s}$ and the projected scale radius are related to mass and concentration and depend on shape and orientation parameters too.

Following Appendix A, explicit formulae between the intrinsic parameters, M_{200} and c_{200}, shape and orientation and measurable projected parameters, $\kappa_{\rm s}$ and $\xi_{\perp}$, can be written as, see also Sereno, Lubini & Jetzer (2010b),

$$\delta_c = \frac{1}{e_{\parallel}} \frac{\kappa_{\rm s}}{\xi_{\perp}} \frac{\Sigma_{\rm cr}}{\rho_{\rm cr}}, \tag{25}$$

where, as usual,

$$\delta_c = \frac{200}{3} \frac{c_{200}}{\ln(1+c_{200}) - c_{200}/(1+c_{200})} \tag{26}$$

and

$$M_{200} = \frac{4\pi}{3} 200 \rho_{\rm cr} c_{200}^3 (e_{\parallel}\sqrt{f}\xi_{\perp})^3. \tag{27}$$

The convergence map can be derived from lensing analyses (Merten et al. 2015; Umetsu et al. 2016). The corresponding χ^2 function can be expressed as (Oguri et al. 2005)

$$\chi^2_{\rm GL} = \sum_{m,n=1}^{N_{\rm GL}} \left[\kappa(\boldsymbol{\theta}_m) - \hat{\kappa}(\boldsymbol{\theta}_m)\right] \left(C_{\rm GL}^{-1}\right)_{mn} \left[\kappa(\boldsymbol{\theta}_n) - \hat{\kappa}(\boldsymbol{\theta}_n)\right], \tag{28}$$

where $\boldsymbol{\kappa} = \{\kappa(\boldsymbol{\theta}_m)\}_{m=1}^{N_{\rm GL}}$ is the convergence map from the lensing analysis, $C_{\rm GL}^{-1}$ is the inverse of the error covariance matrix and the hat symbol denotes a modelled quantity. Here, $\kappa(\boldsymbol{\theta}_m)$ can be seen as either the convergence at a particular position, e.g. the convergence in a pixel, or the mean convergence in a large area, e.g. the mean convergence in an annular bin. $N_{\rm GL}$ is the number of convergence measurements.

4.2 X-ray SB

The observed X-ray SB in an energy band due to bremsstrahlung and line radiation resulting from electron–ion collisions in the high-temperature plasma can be written as

$${\rm SB} = \frac{1}{4\pi(1+z)^3} \int_{\parallel} n_{\rm e}^2 \Lambda_{\rm eff}(T_{\rm e}, \mathcal{Z})\, {\rm d}l, \tag{29}$$

where $T_{\rm e}$ is the intrinsic temperature, $\mathcal{Z}$ is the metallicity and $\Lambda_{\rm eff}$ is the effective cooling function of the ICM in the cluster rest frame, which depends on the energy-dependent area of the instrument (Reese et al. 2010).

In the present paper, we performed a full 2D analysis of the X-ray map, and we determined ellipticity and orientation at once with the gas distribution. Photon number counts follow the Poisson distribution. The χ^2 analogue is the Cash statistic (Cash 1979),

$$\mathcal{C}_{\rm SB,2D} = 2 \sum_{i=1}^{N_{\rm SB,2D}} \left[\hat{N}_i - N_i \ln \hat{N}_i\right], \tag{30}$$

where $N_{\rm SB,\,2D}$ is the number of pixels, N_i is the number of observed photons in the ith pixel and $\hat{N}_i (= \hat{S}_i + \hat{B}_i)$ is the sum of source $\hat{S}_i$ and background $\hat{B}_i$ model amplitudes,

$$\hat{S}_i = t_{\rm exp} A_i {\rm SB}_i, \tag{31}$$

with $t_{\rm exp}$ the exposure time and A_i the pixel collecting area. The formalism in equations (30) and (31) assumes that the background is known and measured in regions well outside the cluster emission. This approximation holds as far as the background is smaller than the source amplitude, otherwise we should model the background emission with a second independent Poisson process.

In the outer regions where the signal is comparable to the background, it can be convenient to measure the mean signal in annular regions. Due to the large total number counts, the statistics is

approximately Gaussian and the χ^2 function for the averaged SB can be written as

$$\chi^2_{\rm SB,1D} = \sum_{i=1}^{N_{\rm SB,1D}} \left(\frac{S_{{\rm X},i} - \hat{S}_{{\rm X},i}}{\delta_{S,i}} \right)^2, \quad (32)$$

where $\hat{S}_{{\rm X},i}$ is the model prediction for the X-ray SB in the ith annulus, $\delta_{S,i}$ is the measurement uncertainty and $N_{\rm SB,\,1D}$ is the number of annular bins.

4.3 Observed temperature

The temperature $T_{\rm e}$ in equation (29) is the intrinsic temperature. The spectroscopic temperature $T_{\rm sp}$ measured by space observatories is well approximated by (Mazzotta et al. 2004)

$$T_{\rm sp} = \frac{\int W T_{\rm e} {\rm d}V}{\int W {\rm d}V}, \quad W = \frac{n^2}{T_{\rm e}^{3/4}}. \quad (33)$$

The χ^2 function for the temperature can be written as (Sereno et al. 2012)

$$\chi^2_T = \sum_{i=1}^{N_T} \left(\frac{T_{{\rm sp},i} - \hat{T}_{{\rm sp},i}}{\delta_{T,i}} \right)^2, \quad (34)$$

where $\hat{T}_{{\rm sp},i}$ is the model prediction for the corresponding observed spectroscopic temperatures $T_{{\rm sp},\,i}$ in the ith angular bin and δ_{T_i} is the measurement uncertainty.

4.4 The Sunyaev–Zel'dovich effect

The Sunyaev–Zel'dovich effect (SZe) is the distortion of the cosmic microwave background (CMB) spectrum due to inverse Compton scattering by the hot ICM energetic electrons (Sunyaev & Zeldovich 1970; Birkinshaw 1999). The amplitude of the signal can be expressed in terms of the Compton-y parameter, which is proportional to the integral of the electron pressure along the line of sight,

$$y \equiv \frac{\sigma_{\rm T} k_{\rm B}}{m_{\rm e} c^2} \int_{\parallel} n_{\rm e} T_{\rm e} {\rm d}l, \quad (35)$$

where $k_{\rm B}$ is the Boltzmann constant, $\sigma_{\rm T}$ is the Thompson cross-section, $m_{\rm e}$ is the electron mass and c is the speed of light in vacuum.

The measured temperature decrement $\Delta T_{\rm SZ}$ of the CMB for an isothermal plasma is given by

$$\Delta T_{\rm SZ} = f_{\rm SZ}(\nu, T) T_{\rm CMB}\, y, \quad (36)$$

where $T_{\rm CMB}$ is the temperature of the CMB and $f_{\rm SZ}(\nu, T)$ accounts for relativistic corrections at frequency ν.

As overall measure of the thermal energy content in a cluster we consider Y, i.e. the Compton-y parameter integrated over a cluster region,

$$Y_\Omega = \int_\Omega y(\theta)\, {\rm d}\theta, \quad (37)$$

where Ω is the angular area. Y is nearly independent of the model of gas distribution used for the analysis and it is a robust quantity for observational tests (Benson et al. 2004). In addition, integrating the Compton-y diminishes (though does not completely remove) effects resulting from the presence of strong entropy features in the central regions of clusters (McCarthy et al. 2003).

The χ^2 function can be written as

$$\chi^2_{\rm SZ} = \sum_{m,n=1}^{N_{\rm SZ}} \left[Y(\Omega_m) - \hat{Y}(\Omega_m) \right] \left(C^{-1}_{\rm SZ} \right)_{mn} \left[Y(\Omega_n) - \hat{Y}(\Omega_n) \right], \quad (38)$$

where $\mathbf{Y} = \{Y(\Omega_m)\}_{m=1}^{N_{\rm SZ}}$ is the set of integrated Compton parameters in the circular annuli, $C^{-1}_{\rm SZ}$ is the inverse of the uncertainty covariance matrix and the hat symbol denotes a modelled quantity.

5 DATA ANALYSIS

In this section, we present the data sets used for the analysis. The different data sets cover a large radial range, from the cluster core up to $\gtrsim$2 Mpc h^{-1}, see Fig. 1.

5.1 Chandra

In the X-ray band, we make use of the archived *Chandra* exposure of MACS1206 obtained in AO3 in Advanced CCD Imaging Spectrometer-I (ACIS-I) configuration, see Fig. 2. Using CIAO 4.8 software (Fruscione et al. 2006) and the calibration data base CALDB 4.7.1, we prepared a cleaned (by grade, status, bad pixels and time intervals affected from flares in the background count rate) events file for a total filtered exposure time of 22.9 ks. The background has been extracted locally over three circular regions with radius of 2 arcmin located at ∼6–7 arcmin from the cluster X-ray peak. Exposure-corrected images in the [0.7–2] keV band are produced.

The point sources identified with the tool `wavedetct` were masked and their regions filled with values of counts from surrounding background areas through the CIAO tool `dmfilth`.

We performed the 2D analysis of the circular region enclosing 80 per cent of the total source emission, with radius $\theta_{80\,\rm per\,cent} = 1.61$ arcmin $\sim 385\,{\rm kpc}\,h^{-1}$. Pixels were binned four by four, with a final resolution of 1.968 arcsec. After the excision of the inner region of radius 5 arcsec, we ended up with 7589 binned pixels. Outside the 80 per cent region, we extracted the SB profiles in 20 circular annuli up to ∼3.5 arcmin, see Fig. 3.

We combined the 2D and the 1D analysis of the SB by summing the Cash, see equation (30), and the χ^2, see equation (32), statistics,[1]

$$\chi^2_{\rm SB} = \mathcal{C}_{\rm SB,2D} + \chi^2_{\rm SB,1D}. \quad (39)$$

Spectra were accumulated in five circular annuli up to $\lesssim$2.8 arcmin and fitted in XSPEC software package (v.12.9; Arnaud 1996) with an absorbed thermal model represented by the components `tbabs`, with a fixed Galactic absorption $n_{\rm H} = 4.35 \times 10^{20}\,{\rm cm}^{-2}$ extrapolated from H I radio maps in Kalberla et al. (2005) and `apec`, with redshift fixed to the value of 0.439 and leaving free three parameters: normalization, temperature and metallicity.

The same emission model with metallicity fixed to the median value, $\mathcal{Z} = 0.29$, was used for regression. The temperature profile is shown in Fig. 4.

5.2 Bolocam

MACS1206 belongs to the Bolocam X-ray SZ sample, a set of 47 clusters with publicly available data from Bolocam[2] and *Chandra*.

Due to the filtering applied to the Bolocam data, the mean signal level of the maps is unconstrained by the Bolocam alone. Because

[1] The left-hand side of equation (39) is not properly a χ^2.
[2] http://irsa.ipac.caltech.edu/data/Planck/release_2/ancillary-data/bolocam/

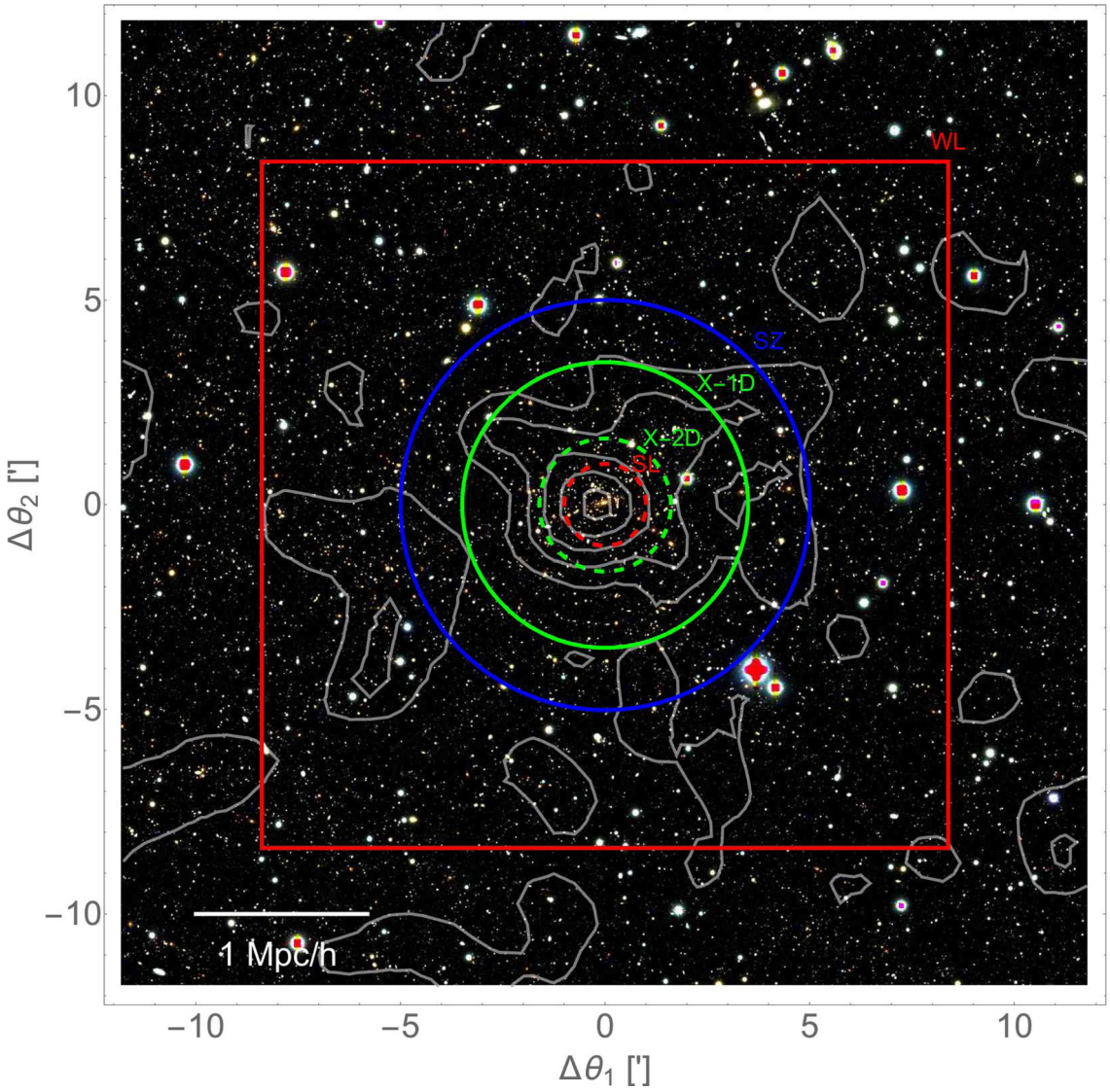

Figure 1. Subaru BVR_c composite colour images of the galaxy cluster MACS1206. The geometric forms centred in the optical centre enclose the regions exploited for inference by each probe: the dashed red circle at 1 arcmin and the red square of semi-size 2 Mpc $h^{-1} \sim 8.39$ arcmin for SL and WL; the dashed and full green circles at $\theta_{80\,\mathrm{per\,cent}} = 1.61$ and 3.49 arcmin for the 2D- and 1D-X-ray analysis, respectively; the blue circle at 5 arcmin for the SZe. The smoothed mass contours from the WL analysis of the Subaru observations are overlaid in white. The convergence levels (for a reference source redshift of $z_s = 20\,000$) go from $\kappa = 0.1$ to 0.5, with increments of 0.1. The image size is 24 arcmin $\times$ 24 arcmin. The horizontal bar represents 1 Mpc h^{-1} at the cluster redshift. North is top and east is left.

the maps are not in general large enough to include areas outside the cluster with negligible SZ emission, there is no way to directly determine this mean level. As a result, the mean signal level in the publicly available images was constrained based on a parametric model fit to the data, which effectively provided an extrapolation beyond the edges of the map to regions where the SZ signal is approximately zero, see Czakon et al. (2015) for more details. In general, this method results in statistical uncertainties on the mean signal level that are significant when computing aperture fluxes from the maps. Furthermore, because the parametric model must be extrapolated beyond the edge of the data, there is the potential for unquantified systematic errors if the cluster is not well described by the model in the extrapolation region.

For this analysis, external SZ measurements from the *Planck* all-sky survey were employed to obtain more precise mean signal estimates. Specifically, the parametric model used to set the mean signal level was jointly constrained by the Bolocam data and the publicly available *Planck* y-maps[3] according to the procedures described in Sayers et al. (2016). Furthermore, the same F-test procedure described in Czakon et al. (2015) was used to determine the minimal parametric model required by the joint data, which was a spherical model with a floating scale radius (the Bolocam data alone required a spherical model with a fixed scale radius). Because the *Planck* y-maps extend well beyond the edge of the cluster, this joint model can be constrained without extrapolation. Compared to the publicly available Bolocam data, the updated mean signal level is shifted upwards by approximately 30 μK_{CMB}, resulting in a 5 per cent reduction in the absolute value of the peak SZ signal (which is negative at the Bolocam wavelength).

[3] http://irsa.ipac.caltech.edu/data/Planck/release_2/docs/

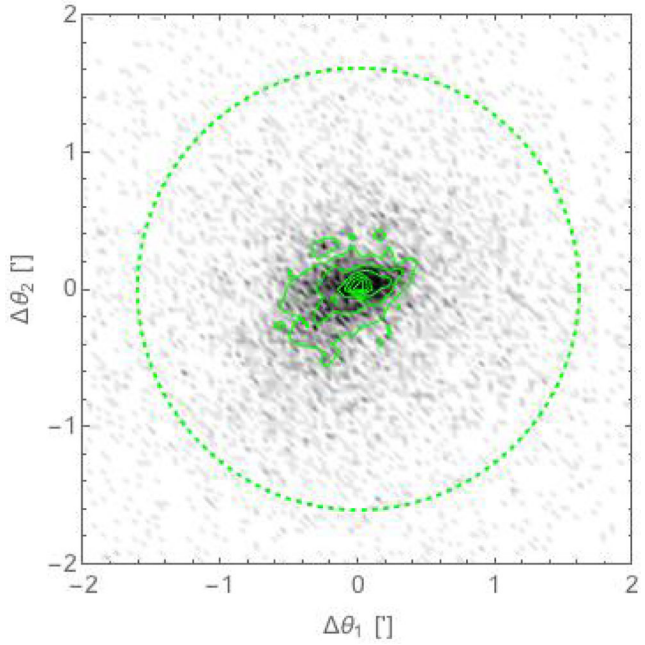

Figure 2. *Chandra* exposure-corrected image in the [0.7–2] keV band used for the spatial analysis. The dashed green circle of radius $\theta_{80\,\mathrm{per\,cent}} = 1.61$ arcmin encloses 80 per cent of the source light. Green contours are X-ray SB contours at arbitrary levels. The image is centred on the optical centre. North is top and east is left.

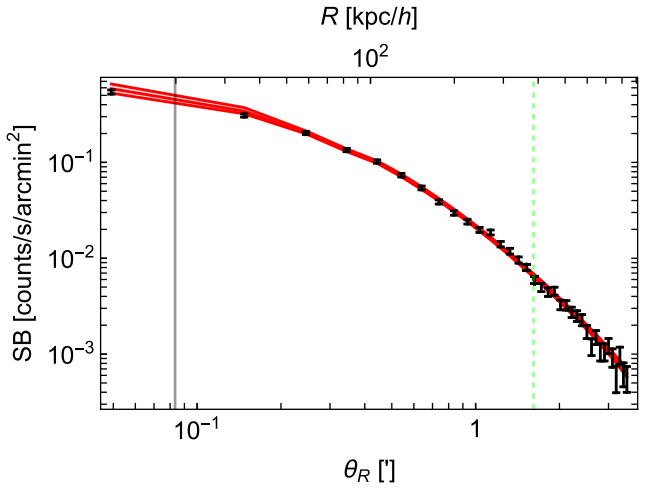

Figure 3. Radial profile along the projected spherical axis of the [0.7–2] keV SB measured by *Chandra* (black points with error bars) in circular annuli. The innermost vertical (black) line is at 5 arcsec, i.e. the minimum radius of the fitting region. The outermost (dashed green) line is at $R_{80\,\mathrm{per\,cent}}$, the radius of the circle including 80 per cent of the total cluster emission. The red line and the shadowed region plot the predicted median profile and the 68.3 per cent region obtained assuming a flat prior for the axial ratio $q_{\mathrm{mat},1}$ of the matter distribution and $q_{\mathrm{ICM}} \geq q$. The red line is not a fit to the plotted points, but it stems from the combined fit to all probes.

For our analysis, we use this updated unfiltered Bolocam map, which contains the SZ image after deconvolving the effects of the filtering due to the Bolocam data processing. As a result, this image provides a representation of the SZ signal from the cluster suitable for aperture photometry. The details of the reduction are given in Sayers et al. (2011) and Czakon et al. (2015). Since the

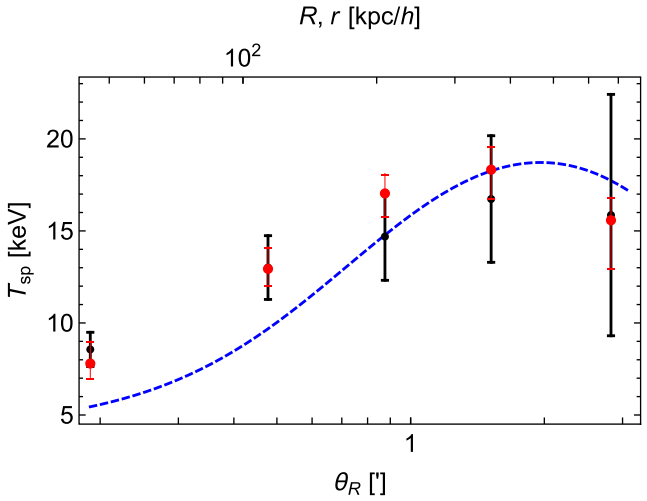

Figure 4. Radial profile of the projected temperature measured by *Chandra* (points with error bars) in circular annuli. Red points and bars denote the median spectroscopic-like temperatures and 68.3 per cent uncertainty obtained assuming a flat prior for the axial ratio $q_{\mathrm{mat},1}$ of the matter distribution and $q_{\mathrm{ICM}} \geq q$. The projected spectroscopic-like temperature is a function of the spherical radius r. The dashed blue curve represents the temperature of the gas distribution T_{e} as measured along the ellipsoidal radius ζ. The plotted lines stem from the combined fit to all probes.

deconvolution results in significant noise on large angular scales, the image is truncated to a size of 10 arcmin × 10 arcmin, see Fig. 5.

The aperture flux was computed in circular annuli based on all of the pixels with centres falling inside each given annulus. Due to the relatively coarse pixelization of the map, a geometric correction factor is then applied to the flux values. Specifically, the deficit or excess of area at the inner or the outer border with respect to the smooth circular area is computed, and Y is corrected by attributing to the area difference the median y value measured at the respective border. In this analysis, we measured the integrated Compton parameter in five equally spaced annular bins up to a maximum radius of 5 arcmin, see Fig. 6. The width of the annuli is 1 arcmin, comparable to the point spread function (PSF) full width at half-maximum (FWHM), in order to limit correlation effects.

If the integration radius is not significantly larger than the Bolocam PSF, some of the SZe emission within the aperture appears outside due to beam smearing and some of the emission from the outside is mapped inside (mostly at the inner border). As a result, the estimates of Y obtained from directly integrating the images can be biased. To estimate the boosting factor, we followed Czakon et al. (2015). The Compton flux is computed both before and after deconvolution. The convolution Gaussian kernel accounts for both the Bolocam PSF (58 arcsec FWHM) and the pointing accuracy (5 arcsec). The Bolocam measured value is then corrected by the ratio of the Y values determined from the unsmoothed and beam-smoothed maps.

For precise error estimates, we computed the aperture flux using identical apertures for the 1000 noise realizations of the Bolocam maps. These realizations fully encapsulate all of the noise statistics of the data, including instrumental, atmospheric and astronomical sources of noise, and all pixel–pixel correlations due to these noise sources. Furthermore, these realizations also include fluctuations based on the mean signal level, which have been updated according to the joint Bolocam/*Planck* fitting procedure described above. The distribution of these 1000 values was then used to measure the

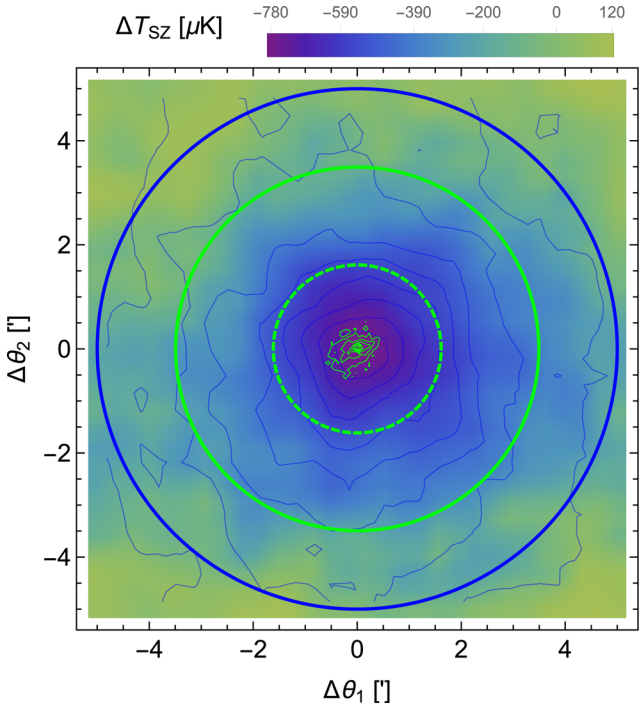

Figure 5. Bolocam/*Planck* deconvolved SZe decrement image of MACS1206. The image is 10 arcmin × 10 arcmin in size and centred on the optical cluster centre. For visualization purposes, the image was smoothed with a Gaussian kernel of standard deviation 1.0 arcmin. The colour bar indicates the temperature decrement. The blue contours go from $\Delta T_{\rm CMB} = -900$ to 0 μK, with increments of 100 μK. The blue circle at 5 arcmin encloses the region considered for inference. The dashed green circle of radius $\theta_{80\,\rm per\,cent} = 1.61$ arcmin encloses 80 per cent of the X-ray emission. Green contours are X-ray SB contours at arbitrary levels. North is top and east is left.

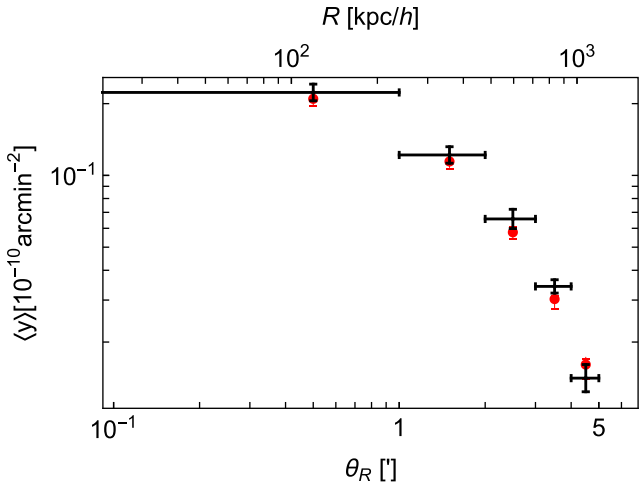

Figure 6. Mean Compton parameter in circular annuli. Black points denote the combined Bolocam/*Planck* data, and the horizontal black bars cover the radial range of the annulus. Red points and bars denote the median fitted values obtained assuming a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution and $q_{\rm ICM} \geq q$. The plotted lines stem from the combined fit to all probes.

uncertainty covariance matrix for the aperture fluxes within the annuli.

5.3 Lensing

The lensing features of MACS1206 were observed and studied in the context of the CLASH collaboration. We refer to Umetsu et al. (2012, 2016) and Merten et al. (2015) for a full presentation of the data and of the methods.

5.3.1 Strong lensing

For our analysis, we considered the joint analysis of the inner regions of MACS1206 based on *HST* data presented by Zitrin et al. (2015).[4] Constraints in the inner core come mostly from strong-lens modelling of multiple-image systems identified with deep *HST* imaging and VLT/VIMOS spectroscopy. MACS1206 hosts a well-known giant arc system at $z_{\rm s} = 1.03$ (Ebeling et al. 2009). Zitrin et al. (2012a) identified 12 more candidate multiple-image systems of distant sources, bringing the total known for this cluster to 50 multiply lensed images of 13 sources. The images cover fairly evenly the central region, 3 arcsec $\lesssim \theta \lesssim$ 1 arcmin, and span a wide redshift range of $1 \lesssim z_{\rm s} \lesssim 5.5$.

Additional constraints come from WL in the *HST* field, which exploits the high density of lensed background galaxies, ∼50 galaxies arcmin^{-2}.

The analysis was performed in two distinct parametrizations: the first one adopts light-traces-mass for both galaxies and dark matter (Zitrin et al. 2009) while the other assumes an analytical, elliptical NFW form for the dark matter halo components (Zitrin et al. 2013). As a reference, we considered the NFW parametrization, which conforms better to our modelling. It does not employ any external shear and it lets the dark matter (DM) haloes to be elliptical.

The very inner region of the cluster core is dominated by the brightest cluster galaxy (BCG) and by massive ellipticals, see Fig. 7. Whereas the average mass distribution and radial slope follow the general profile, local substructures can strongly impact the inferred local projected shape and orientation of the core. On the geometrical side, the analysis of the cluster core can deviate from the overall picture.

We did not consider the full fine-resolution 2D map but we applied a conservative azimuthal binning scheme. We computed the mean convergence in six equally spaced angular annuli centred in the optical centre. The innermost and the outermost radii were set to an angular scale of 5 arcsec and 1 arcmin, respectively.

This conservative approach for the SL analysis is also justified by the number of noise realizations, which is not big enough to properly compute the uncertainty covariance matrix of the 2D-SL map.

The formal statistical errors of the models are underestimated (Zitrin et al. 2015). The actual (and much larger) uncertainties that account for model-dependent systematics can be inferred by comparing the convergence profiles derived under the two distinct modellings from Zitrin et al. (2015). The projected mass enclosed within the critical curves of the CLASH clusters agrees typically within ∼15 per cent, which gives an empirical assessment of the true underlying errors.

[4] The mass models are available through the *HST* Archive at https://archive.stsci.edu/prepds/clash/

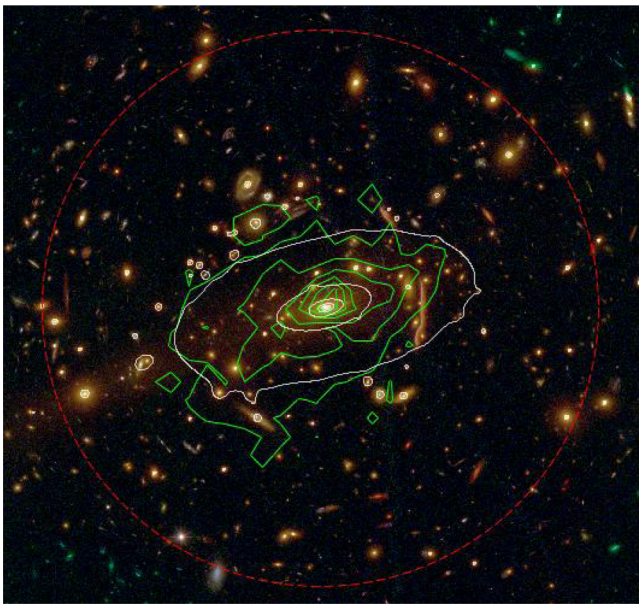

Figure 7. *HST* composite colour images centred of the core of the galaxy cluster MACS1206. The dashed red circle at 1 arcmin encloses the region considered for the SL analysis. The mass contours from the analysis of *HST* observations from Zitrin et al. (2015) based on the NFW parametrization are overlaid in white. The mass contours follow the convergence and go from $\kappa = 0.1$ to 1.0, with increments of 0.1 (for a reference source redshift of $z_s = 20\,000$). The green contours are the smoothed arbitrary levels of the X-ray SB. North is top and east is left.

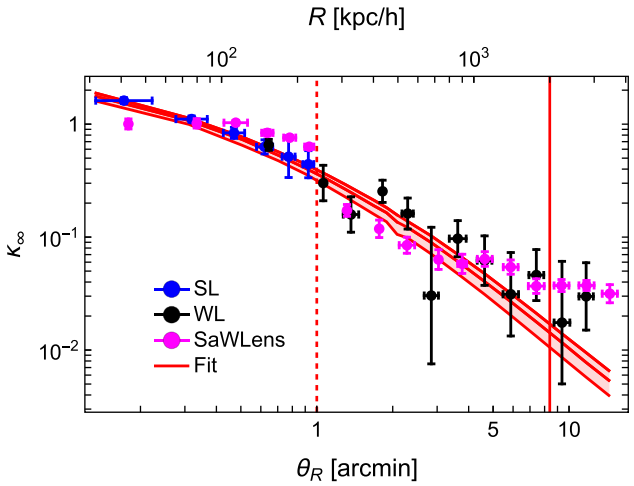

Figure 8. Radial profile of the convergence map in circular annuli. Convergence is rescaled to $z = 20\,000$. The black and blue points are obtained from the reference WL and SL analysis, respectively. The magenta points were obtained with SaWLens. The vertical lines denote the fitting regimes. The innermost is at 1 arcmin, i.e. the maximum radius for SL. The outermost is at 2 Mpc h^{-1}, which includes the considered WL region. The red profile and the shadowed region plot the median profile and the 68.3 per cent region around the median obtained assuming a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution and $q_{\rm ICM} \geq q$. The red line is not a fit to the plotted points, but it stems from the combined fit to all probes.

We computed the standard deviation of the differences of the convergences of the pixels in each annulus and added this error in quadrature to the formal statistical error. For this computation, original maps were rebinned with a final formal resolution of 0.65 arcsec.

Since the uncertainty budget is dominated by this second term, we considered only the diagonal part of the covariance matrix.

The convergence profile is plotted in Fig. 8. The SL- and WL-based convergences are in remarkable agreement in the overlapping region.

5.4 Subaru WL analysis

For our Subaru WL mapmaking, we used the non-parametric method of Umetsu et al. (2015), a two-dimensional extension of the Cluster Lensing Mass Inversion (CLUMI) code developed by Umetsu et al. (2011). We reconstructed the projected mass distribution around MACS1206 from a joint analysis of two-dimensional shear and azimuthally averaged magnification. The combined constraints allow us to break the mass-sheet degeneracy. Full details will be given in a forthcoming paper that will present a systematic two-dimensional WL analysis of 20 CLASH clusters (Umetsu et al., in preparation). Here, we briefly outline the methods and data used in our analysis.

5.4.1 Data

The WL analysis of the wide-field exploits Subaru data and photometry. MACS1206 was observed in deep BVR_cI_cz' with the wide-field camera Suprime-Cam at the prime focus of the 8.3 m Subaru telescope. The mosaicked images cover a field of approximately 36 arcmin × 34 arcmin (Umetsu et al. 2012).

In the present study, we used the WL shear and magnification data as obtained by Umetsu et al. (2014). We refer the reader to Umetsu et al. (2014, section 4) and Umetsu et al. (2016, sections 3.2–3.4) for a summary of our Subaru data and data analysis. We performed a new two-dimensional shear analysis using the shear catalogue presented in Umetsu et al. (2014). The magnification bias measurements used in this study are presented in Umetsu et al. (2014, fig. 2). The systematic uncertainty in the absolute mass calibration using the CLASH WL measurements was estimated to be ∼8 per cent (Umetsu et al. 2014).

5.4.2 Estimators and covariance matrix

For mapmaking, the lensing field was pixelized into a regular Cartesian grid of $N_{\rm pix} = 48^2 = 2304$ independent pixels with $\Delta\theta = 0.5$ arcmin spacing, covering the central 24 arcmin × 24 arcmin region. The magnification bias analysis in this region was performed by Umetsu et al. (2014).

The lensing signal is described by the vector $\boldsymbol{s} = \{\Sigma_m\}_{m=1}^{N_{\rm pix}}$ whose elements contain the cell-averaged values of the two-dimensional surface mass density field $\Sigma(\boldsymbol{\theta})$, sampled on the grid. Both the convergence, $\kappa(\boldsymbol{\theta})$, and the shear, $\gamma(\boldsymbol{\theta})$, fields can be written as linear combinations of $\boldsymbol{s}$ (Umetsu et al. 2015). In our approach, we combine the observed spatial distortion pattern $g(\boldsymbol{\theta})$ with the azimuthally averaged magnification measurements $n(\theta)$, which impose a set of azimuthally integrated constraints on the mass distributions.

The magnification is locally related to κ and the magnification constraints provide the otherwise unconstrained normalization of $\Sigma(\boldsymbol{\theta})$ over a set of concentric annuli where count measurements are available, thus effectively breaking the mass-sheet degeneracy. No assumption is made of azimuthal symmetry or isotropy of the underlying mass distribution.

The (complex) reduced shear field can be reconstructed from shape measurements of background galaxies. The weighted average $g_m \equiv g(\boldsymbol{\theta}_m)$ on the grid ($m = 1, 2, \ldots, N_{\rm pix}$) can be written as

$$g_m = \left[\sum_k S(\boldsymbol{\theta}_{(k)}, \boldsymbol{\theta}_m) w_{(k)} g_{(k)}\right] \left[\sum_k S(\boldsymbol{\theta}_{(k)}, \boldsymbol{\theta}_m) w_{(k)}\right]^{-1}, \quad (40)$$

where $g_{(k)}$ is an estimate of $g(\boldsymbol{\theta})$ for the kth galaxy at $\boldsymbol{\theta}_{(k)}$. The statistical weight $w_{(k)}$ can be written as $w_{(k)} = 1/(\sigma^2_{g(k)} + \alpha^2_g)$, with $\sigma^2_{g(k)}$ the error variance of $g_{(k)}$ and $\alpha^2_g (= 0.16)$ the softening variance (Umetsu et al. 2014), as typical of the mean variance $\overline{\sigma^2_g}$ found in Subaru observations (Umetsu et al. 2009; Okabe et al. 2010). $S(\boldsymbol{\theta}_{(k)}, \boldsymbol{\theta}_m)$ is the window function. We adopted a top-hat window of filtering radius $\theta_{\rm f} = 0.4$ arcmin (Umetsu et al. 2014).

The error covariance matrix for the weighted average g_m is

$$\mathrm{Cov}(g_{\alpha,m}, g_{\beta,n}) \equiv \delta_{\alpha\beta} \left(C_g\right)_{mn} = \frac{\delta_{\alpha\beta}}{2} \sigma_{g,m} \sigma_{g,n} \xi_H(|\boldsymbol{\theta}_m - \boldsymbol{\theta}_n|), \quad (41)$$

where the indices $\alpha, \beta \in \{1, 2\}$, $\sigma_{g,m} = \sigma_g(\boldsymbol{\theta}_m)$ denotes the error dispersion for $g_m = g_{1,m} + \mathrm{i} g_{2,m}$ ($m = 1, 2, \ldots, N_{\rm pix}$) and $\xi_H(x; \theta_{\rm f})$ is the auto-correlation of a pillbox of radius $\theta_{\rm f}$ normalized as $\xi_H(0; \theta_{\rm f}) = 1$ (Umetsu et al. 2015, equation 11).

We interpreted the observed WL signal, equation (40), accounting for the non-linear effect on the source-averaged reduced shear. The signal can be written as (Seitz & Schneider 1997)

$$\hat{g}(\boldsymbol{\theta}_m) = \frac{\gamma(\boldsymbol{\theta}_m)}{1 - f_{W,g}\, \kappa(\boldsymbol{\theta}_m)}, \quad (42)$$

where $f_{W,g} = \langle\beta^2\rangle_g / \langle\beta\rangle_g^2$ is a correction factor of the order unity estimated from the source-averaged lensing depths $\langle\beta\rangle_g = \langle D_{\rm ls}/D_{\rm os}\rangle_g$ and $\langle\beta^2\rangle_g = \langle(D_{\rm ls}/D_{\rm os})^2\rangle_g$ for the shear WL analysis. We used $\langle\beta\rangle_g = 0.54 \pm 0.03$ and $f_{W,g} = 1.06$ (Umetsu et al. 2014, table 3).

We excluded from our analysis the pixels within the Einstein radius $\theta_{\rm Ein}(z_{\rm s} = 2) = 26.8$ arcsec (Umetsu et al. 2016, table 1), where $\Sigma(\boldsymbol{\theta})$ can be close to or greater than the critical value $\Sigma_{\rm crit}$, as well as those containing no background galaxies with usable shape measurements. For distortion measurements, $g_1(\boldsymbol{\theta})$ and $g_2(\boldsymbol{\theta})$, a total of 2293 measurement pixels are usable, corresponding to 4586 constraints.

For magnification measurements, we used the azimuthally averaged source number counts $\{n_{\mu,i}\}_{i=1}^{N_{\rm bin}}$ and their total errors $\{\sigma_{\mu,i}\}_{i=1}^{N_{\rm bin}}$ as obtained by Umetsu et al. (2014) using their flux-limited sample of $BR_{\rm C}z'$-selected red background galaxies. For additional analysis details, we refer to Umetsu et al. (2016, section 3.4). The magnification constraints were measured in $N_{\rm bin} = 10$ log-spaced circular annuli centred on the cluster, spanning the range $[\theta_{\rm min}, \theta_{\rm max}] = [0.9\,\mathrm{arcmin}, 16\,\mathrm{arcmin}]$ with a constant logarithmic spacing, $\Delta\ln\theta \simeq 0.29$.

The theoretical azimuthally averaged source number counts is (Umetsu et al. 2015, 2016)

$$\hat{n}_{\mu,i} = \overline{n}_\mu \sum_m \mathcal{P}_{im} \mu^{2.5s-1}(\boldsymbol{\theta}_m), \quad (43)$$

where $\mu = [(1-\kappa)^2 - |\gamma|^2]^{-1}$ is the lensing magnification in the subcritical regime, $\overline{n}_\mu$ is the (unlensed) mean background surface number density of source galaxies and s is the logarithmic count slope evaluated at the fainter magnitude limit $m_{\rm lim}$, $s = [\mathrm{d}\log_{10} N(< m)/\mathrm{d}m]_{m=m_{\rm lim}}$; $\mathcal{P}_{im} = (\sum_m A_{mi})^{-1} A_{mi}$ is the projection matrix normalized in each annulus as $\sum_m \mathcal{P}_{im} = 1$. Here, A_{mi} represents the fraction of the area of the mth cell lying within the ith annular bin ($0 \le A_{mi} \le 1$).

The background sample used for lensing magnification differs from the galaxy population used for shape measurements. The mean lensing depth of the background sample is $\langle\beta\rangle_\mu = 0.51 \pm 0.03$ with $z_{\rm eff} = 1.04$ (Umetsu et al. 2014, table 4). The count normalization and slope parameters are $\overline{n}_\mu = (11.4 \pm 0.4)\,\mathrm{arcmin}^{-2}$ and $s = 0.13 \pm 0.05$ at the fainter magnitude limit $z' = 24.6$ ABmag (Umetsu et al. 2014, table 4) using the source counts in the outskirts at [10 arcmin, $\theta_{\rm max}$] (Umetsu et al. 2014).

5.4.3 *Joint reconstruction of the mass map*

We reconstructed the mass distribution of MACS1206 from a joint likelihood analysis of the shear and magnification measurements, $\{g_{1,m}, g_{2,m}\}_{m=1}^{N_{\rm pix}}$ and $\{n_{\mu,i}\}_{i=1}^{N_{\rm bin}}$. The model $\boldsymbol{m}$ is specified by $\boldsymbol{s}$ with $N_{\rm pix} = 2304$ signal parameters and a set of calibration parameters to marginalize over, $\boldsymbol{c} = \{\langle\beta\rangle_g, f_{W,g}, \langle\beta\rangle_\mu, \overline{n}_\mu, s\}$. We have a total of $N_{\rm data} = 4586 + 10 = 4596$ constraints, yielding $N_{\rm data} - N_{\rm pix} = N_{\rm dof} = 2292$ degrees of freedom.

The posterior probability of a model $\boldsymbol{m}$ is proportional to the product of the likelihood $\mathcal{L}(\boldsymbol{m})$ and the prior probability. We used Gaussian priors on the calibration nuisance parameters $\boldsymbol{c}$, given in terms of quadratic penalty terms with mean values and errors estimated from data (Umetsu et al. 2014) as stated above.

The joint likelihood function $\mathcal{L}(\boldsymbol{m}) = \mathcal{L}_g(\boldsymbol{m})\mathcal{L}_\mu(\boldsymbol{m})$ for combined WL data is given as the product of the likelihood functions for shear, $\mathcal{L}_g$, and magnification, $\mathcal{L}_\mu$. The shear log-likelihood function $l_g \equiv -2\mathcal{L}_g$ can be written as

$$l_g = \frac{1}{2} \sum_{m,n=1}^{N_{\rm pix}} \sum_{\alpha=1}^{2} [g_{\alpha,m} - \hat{g}_{\alpha,m}(\boldsymbol{m})] \left(\mathcal{W}_g\right)_{mn} \times [g_{\alpha,n} - \hat{g}_{\alpha,n}(\boldsymbol{m})] + \mathrm{const.}, \quad (44)$$

where $\hat{g}_{\alpha,m}(\boldsymbol{m})$ is the theoretical expectation for $g_{\alpha,m} = g_\alpha(\boldsymbol{\theta}_m)$; $\mathcal{W}_g$ is the shear weight matrix,

$$\left(\mathcal{W}_g\right)_{mn} = M_m M_n \left(C_g^{-1}\right)_{mn}, \quad (45)$$

where M_m is the mask weight, defined such that $M_m = 0$ if the mth cell is masked out and $M_m = 1$ otherwise.

The log-likelihood function for magnification bias data, $l_\mu \equiv -2\ln\mathcal{L}_\mu$, is given by

$$l_\mu = \frac{1}{2} \sum_{i=1}^{N_{\rm bin}} \frac{[n_{\mu,i} - \hat{n}_{\mu,i}(\boldsymbol{m})]^2}{\sigma^2_{\mu,i}} + \mathrm{const.}, \quad (46)$$

where $\hat{n}_{\mu,i}(\boldsymbol{m})$ is the theoretical prediction for the observed counts $n_{\mu,i}$, see equation (43). Following Umetsu et al. (2015), we used Monte Carlo integration to calculate the projection matrix $\mathcal{P}_{im}$ of size $N_{\rm bin} \times N_{\rm pix}$, which is needed to predict $\{n_{\mu,i}\}_{i=1}^{N_{\rm bin}}$ for a given $\boldsymbol{m} = (\boldsymbol{s}, \boldsymbol{c})$.

The negative log-posterior function $F(\boldsymbol{m})$ is expressed as the linear sum of the log-likelihood functions (l_g, l_μ) and the Gaussian prior terms. The *best-fitting* solution, $\hat{\boldsymbol{m}}$, corresponds to the global maximum of the joint posterior distribution. The resulting mass map is plotted in Fig. 9.

Uncertainties are estimated by evaluating the Fisher matrix at $\boldsymbol{m} = \hat{\boldsymbol{m}}$ as (Umetsu et al. 2015)

$$\mathcal{F}_{pp'} = \left\langle \frac{\partial^2 F(\boldsymbol{m})}{\partial m_p \partial m_{p'}} \right\rangle \Bigg|_{\boldsymbol{m}=\hat{\boldsymbol{m}}}, \quad (47)$$

where the angular brackets represent an ensemble average over all possible (an infinite number of) noise realizations, and the indices (p, p') run over all model parameters $\boldsymbol{m} = (\boldsymbol{s}, \boldsymbol{c})$.

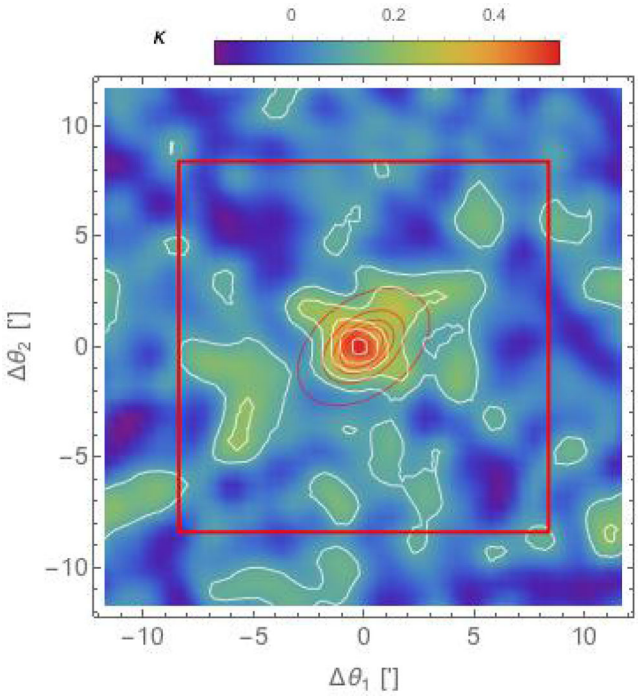

Figure 9. Two-dimensional WL mass map of MACS1206 reconstructed using wide-field multi-colour imaging observations. The image is 24 arcmin × 24 arcmin in size and centred on the optical cluster centre. For visualization purposes, the image was smoothed with a Gaussian kernel of standard deviation 1.5 arcmin. The colour bar indicates the lensing convergence κ for a reference source at $z_s = 20\,000$. The convergence isolevels go from $\kappa = 0.1$ to 1.0, with increments of 0.1. White and red contours follow the smoothed map and the predicted median NFW profile obtained assuming a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution and $q_{\rm ICM} \geq q$, respectively. The red square of semi-size $2\,{\rm Mpc}\,h^{-1} \sim 8.39$ arcmin encloses the WL region used for inference. North is top and east is left.

The posterior covariance matrix is then obtained as

$$\mathrm{Cov}(m_p, m_{p'}) \equiv \left(C^{\rm stat}\right)_{pp'} = \left(\mathcal{F}^{-1}\right)_{pp'}. \quad (48)$$

Additionally, we accounted for the uncorrelated large-scale structures projected along the line of sight. The cosmic noise covariance matrix $(C^{\rm lss})_{mn}$ for the pixelized surface mass density distribution $\boldsymbol{s} = \{\Sigma_m\}_{m=1}^{N_{\rm pix}}$ was computed by projecting the non-linear matter power spectrum of Smith et al. (2003) for the *Wilkinson Microwave Anisotropy Probe* seven-year cosmology (Komatsu et al. 2011).

The total covariance matrix for the pixelized surface mass density distribution is $(C)_{mn} = (C^{\rm stat})_{mn} + (C^{\rm lss})_{mn}$.

5.4.4 *Parametric analysis of the mass map*

We finally fitted the convergence map to constrain the parametric mass distribution. Following Merten et al. (2015), we set the fitting area to a square with physical scale of $2\,{\rm Mpc}\,h^{-1}$ centred in the optical centre.

We summed the χ^2 contributions from the fine-resolution grid dominated by the SL observations $\chi^2_{\rm SL}$ and the term from the WL observations by the Subaru telescope, $\chi^2_{\rm WL}$,

$$\chi^2_{\rm GL} = \chi^2_{\rm SL} + \chi^2_{\rm WL}. \quad (49)$$

5.5 Cluster centre

Misidentification of the cluster centre is a potential source of systematic errors for joint, multi-wavelength analyses. We fixed the cluster centre at the sky position of the BCG of RA = 12:06:12.15, Dec. = −08:48:03.4 (J2000). Miscentring effects are small in MACS1206. The BCG position agrees with the peak of X-ray emission (RA = 12:06:12.08, Dec. = −08:48:02.6) within 1.3 arcsec or a projected offset distance of $\sim 5\,{\rm kpc}\,h^{-1}$ at the cluster redshift, see Fig. 7. Furthermore, the BCG and the peak of the total mass distribution are offset by just 1 arcsec, well within the uncertainties (Umetsu et al. 2012), see Figs 1 and 7.

In the present work, we conservatively limited our analysis to radii greater than 5 arcsec ($\sim 20\,{\rm kpc}\,h^{-1}$), which slightly exceeds the location of the innermost SL constraint and is sufficiently large to avoid the BCG contribution. Our inner radial limit corresponds roughly to four times the offset between the BCG centre and the X-ray peak, beyond which smoothing from the cluster miscentring effects on the convergence profile is sufficiently negligible (Johnston et al. 2007; Umetsu et al. 2012).

6 INFERENCE

The problem of finding the volume density distribution of haloes whose projected isocontours are similar ellipses has no unique solution (Stark 1977).

In the generic terms, we have to constrain the five unknown geometrical intrinsic properties (two axial ratios and three orientation angles) of the ellipsoidal halo. However, we can only measure three observable quantities, i.e. the ellipticity ϵ, the orientation θ_ϵ and the elongation $e_\parallel$, see Appendix A. Ellipticity and orientation can be inferred from a single map whereas the elongation can be derived by combining data sets with a different dependence on density. For example, we can combine X-ray and SZ to directly infer the elongation of the gas distribution. For an isothermal plasma (De Filippis et al. 2005),

$$e_\parallel = \frac{1}{D_{\rm d}} \frac{\Delta T^2_{{\rm SZ},0}}{{\rm SB}_0} \frac{\Lambda}{T^2}. \quad (50)$$

No assumption is needed about HE but clumpiness and contamination from structures along the line of sight can bias the result.

The problem is then underconstrained even with an ideal multi-probe data set without noise (Sereno 2007).

However, as far as the effect on observations is considered, some intrinsic geometrical parameters are more equal than others (Orwell 1949; Sereno & Umetsu 2011; Sereno et al. 2012). In a triaxial analysis, ellipticity and elongation strongly depend on the minor-to-major axial ratio, $q_{\rm mat,1}$, and on the inclination angle, ϑ. The projected orientation angle is just ψ plus an arbitrary constant. That makes three main parameters for three observables.

Bayesian inference is suitable to cluster deprojection. Whereas some main parameters are essentially derived from the data, others can be more subject to priors. For example, we obtain similar results for the minor-to-major axial ratio and the halo orientation either using $q_{\rm mat,1}$ and $q_{\rm ICM,\,1}$ as free parameters and $q_{\rm mat,2}$ and $q_{\rm ICM,\,2}$ as functions of the other axial ratios or considering all four axial ratios as free.

To assess realistic probability distributions for the parameters, we exploited the Bayes' theorem, which states that

$$p(\mathbf{P}|\mathbf{d}) \propto \mathcal{L}(\mathbf{P}|\mathbf{d}) p(\mathbf{P}), \quad (51)$$

Table 3. Results of the regression under different priors for the mass and gas shape. For the mass, we considered either spherical, flat or *N*-body priors. For the gas, we considered $q_{\rm ICM} \geq q$ and, additionally, either $\mathcal{T}_{\rm mat} = \mathcal{T}_{\rm ICM}$ or $q_{\rm ICM} = q_\Phi$. The reference case, i.e. *q*-flat and $q_{\rm ICM} \geq q$, is listed in column 2. Parameters in brackets are held fixed. Units are as in Table 1. Typical values and dispersions are computed as biweighted estimators of the marginalized posterior distributions.

	q-flat $q_{\rm ICM} \geq q$	*q*-spherical	*q*-flat $\mathcal{T}_{\rm mat} = \mathcal{T}_{\rm ICM}$	*q*-flat $q_{\rm ICM} = q_\Phi$	*q*-*N*body $q_{\rm ICM} \geq q$
M_{200}	1.137 ± 0.229	1.057 ± 0.157	1.027 ± 0.362	0.951 ± 0.242	1.069 ± 0.238
c_{200}	6.277 ± 1.188	5.047 ± 0.730	6.682 ± 1.823	6.469 ± 1.537	6.051 ± 1.126
$q_{\rm mat,1}$	0.466 ± 0.119	[1]	0.406 ± 0.136	0.409 ± 0.111	0.439 ± 0.077
$q_{\rm mat,2}$	0.735 ± 0.176	[1]	0.759 ± 0.143	0.544 ± 0.096	0.586 ± 0.098
$\cos\vartheta$	0.297 ± 0.204	–	0.319 ± 0.217	0.220 ± 0.184	0.272 ± 0.194
φ	−0.438 ± 1.609	–	−0.075 ± 1.257	0.147 ± 1.058	0.154 ± 1.474
ψ	1.027 ± 0.123	–	0.972 ± 0.278	0.937 ± 0.061	0.931 ± 0.152
$q_{\rm ICM,\,1}$	0.587 ± 0.109	–	0.498 ± 0.119	0.734 ± 0.055	0.501 ± 0.071
$q_{\rm ICM,\,2}$	0.779 ± 0.057	–	0.783 ± 0.113	0.791 ± 0.052	0.784 ± 0.070
n_0	0.010 ± 0.001	–	0.011 ± 0.002	0.010 ± 0.001	0.011 ± 0.002
$\zeta_{\rm c}$	169.780 ± 12.02	–	167.300 ± 8.520	163.900 ± 3.466	166.430 ± 6.748
$\zeta_{\rm t}/\zeta_{\rm c}$	8.250 ± 1.635	–	8.125 ± 1.358	7.576 ± 1.563	7.921 ± 1.663
β	0.600 ± 0.025	–	0.591 ± 0.031	0.578 ± 0.030	0.592 ± 0.025
η	0.627 ± 0.047	–	0.628 ± 0.060	0.638 ± 0.069	0.631 ± 0.049
$\gamma_{\rm ICM}$	1.810 ± 0.632	–	2.322 ± 0.390	2.383 ± 0.506	2.335 ± 0.449
T_0	24.245 ± 2.646	–	21.607 ± 2.197	20.319 ± 3.337	21.727 ± 2.193
$\zeta_{cT}/\zeta_{\rm c}$	9.069 ± 1.034	–	8.670 ± 0.817	8.559 ± 1.137	8.667 ± 0.738
a_T	[0.0]	[0.0]	[0.0]	[0.0]	[0.0]
b_T	[2.0]	[2.0]	[2.0]	[2.0]	[2.0]
c_T	2.576 ± 0.313	–	2.779 ± 0.169	2.770 ± 0.220	2.782 ± 0.167
$T_{\rm cc}$	4.221 ± 1.328	–	4.046 ± 1.159	3.398 ± 1.076	3.871 ± 1.053
$\zeta_{\rm cc}/\zeta_{\rm c}$	1.140 ± 0.211	–	0.998 ± 0.147	0.967 ± 0.193	0.969 ± 0.144
$\alpha_{\rm cc}$	[1.9]	[1.9]	[1.9]	[1.9]	[1.9]

where $p(\mathbf{P}|\mathbf{d})$ is the posterior probability of the parameters $\mathbf{P}$ given the data $\mathbf{d}$, $\mathcal{L}(\mathbf{P}|\mathbf{d})$ is the likelihood of the data given the model parameters and $p(\mathbf{P})$ is the prior probability distribution for the model parameters.

For our multi-wavelength analysis, the likelihood can be written as (Sereno et al. 2013)

$$\mathcal{L} \propto \exp\left[-\left(\chi^2_{\rm GL} + \chi^2_{\rm SB} + \chi^2_{\rm T} + \chi^2_{\rm SZ}\right)/2\right]. \quad (52)$$

Our method relies on a minimum number of assumptions, e.g. we do not require equilibrium. Under this general setting, the matter part (GL) and the ICM part (X+SZ) communicate only through the orientation, which is shared by gas and matter, and the shape parameters, whose relations are defined through priors.

Under the strong assumption of spherical symmetry, the mass and the concentration, which only appear in the $\chi^2_{\rm GL}$, are determined by lensing alone. Under the triaxial assumption, the inference of mass and concentration is affected by the X+SZ part too through the shape and orientation parameters. In fact, the estimate of the concentration is strongly correlated with shape and orientation (Oguri et al. 2005; Sereno, Jetzer & Lubini 2010a).

The present analysis presents some major developments with respect to the methodology used in Sereno et al. (2013). The priors cover a larger range of scenarios than those used in Sereno et al. (2013) and are updated to latest results from numerical simulations. Priors for shape and orientation have been introduced in Sections 2 and 3.

Furthermore, we now fit the 2D map of the X-ray SB rather than the averaged 1D profile only.

The last major development with respect to Sereno et al. (2013) and Umetsu et al. (2015) is in our treatment of the likelihood. Sereno et al. (2013) used a step procedure, where the data were first fitted in terms of projected parameters ($\kappa_{\rm s}$, $\epsilon_{\rm mat}$, $e_{\rm ICM,\,\|}$,...) and then the inferred probability of the projected parameters (approximated with either a smooth kernel distribution or a multi-variate Gaussian distribution) was used as likelihood to infer the intrinsic parameters (M_{200}, $q_{\rm mat,1}$, $\cos\vartheta$,...). Now, we fit the data in terms of the intrinsic parameters in a single step, see equation (52).

7 RESULTS

The Bayesian regression scheme was applied to infer the properties of total matter and gas. We used different priors to perform the regression under five schemes.

(i) Flat prior for the minor-to-major axial ratio of the matter distribution (*q*-flat) and gas rounder than total matter ($q_{\rm ICM} \geq q$). This is our reference setting.
(ii) *q*-flat and shared triaxiality ($\mathcal{T}_{\rm mat} = \mathcal{T}_{\rm ICM}$).
(iii) *q*-flat and gas following the potential ($q_{\rm ICM} = q_\Phi$).
(iv) *N*-body prior for the matter shape (*q*-*N* body) and $q_{\rm ICM} \geq q$.
(v) Spherical symmetry.

Apart from the spherical case, we always take the gas to be rounder than the matter distribution, either through $q_{\rm ICM} \geq q$, $\mathcal{T}_{\rm mat} = \mathcal{T}_{\rm ICM}$ or $q_{\rm ICM} = q_\Phi$, and we assume a priori randomly oriented clusters. When not stated otherwise, priors for other parameters are set to their default distributions listed in Table 1.

Results are presented in Table 3. The marginalized 2D and 1D inferred probability distributions are illustrated in Fig. 10 for the reference case.

With no regard to the priors, estimates of mass and concentration are consistent. Only the results based on the spherical assumption deviate significantly. Stronger priors, i.e. *q*-*N* body or $q_{\rm ICM} = q_\Phi$,

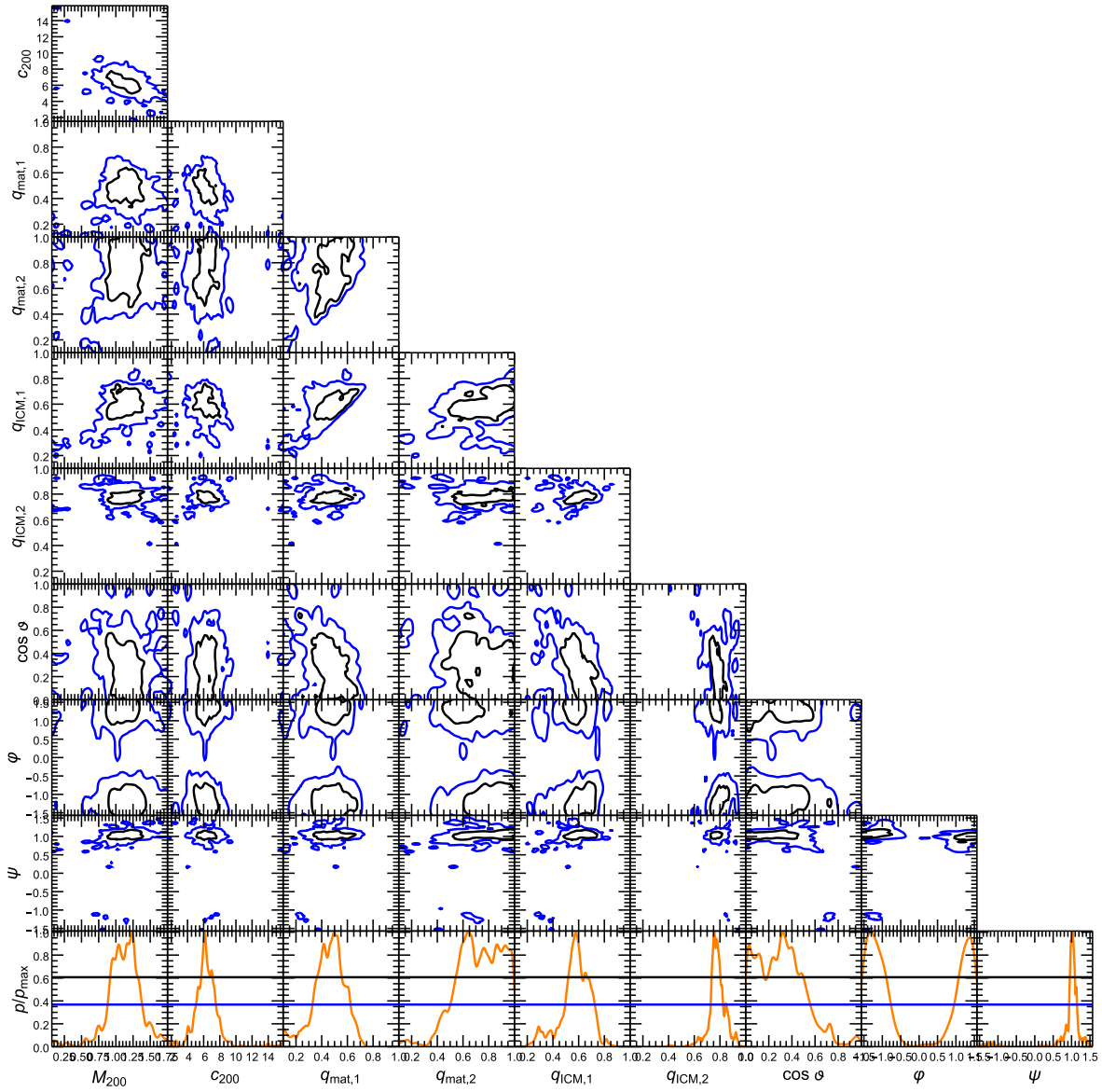

Figure 10. Probability distributions of the parameters of mass distribution, shape and orientation as derived from the regression assuming a flat prior for the matter axial ratios and $q_{\rm ICM} \geq q$. The black and blue contours include the 1σ, 2σ confidence regions in two dimensions, here defined as the regions within which the probability is larger than $\exp[-2.3/2]$ or $\exp[-6.17/2]$ of the maximum, respectively. The bottom row plots the marginalized 1D distributions, renormalized to the maximum probability. The blue and black levels denote the confidence limits in one dimension, i.e. $\exp[-1/2]$ and $\exp[-4/2]$ of the maximum.

enforce smaller values of the axial ratios, but they still are compatible with other less informative schemes, i.e. $q_{\rm ICM} \geq q$ or $\mathcal{T}_{\rm mat} = \mathcal{T}_{\rm ICM}$.

Large inclination angles ($\cos\vartheta \lesssim 0.3$) are always favoured.

The parameters of the gas distributions are remarkably stable through the different setting of priors, which mostly affect shape. The estimate of the central gas density is strongly anti-correlated with the elongation parameter $e_{\rm ICM,\parallel}$ (Sereno et al. 2012). The shape and orientation parameters vary but still conjure to keep $e_{\rm ICM,\parallel}$ stable, and the gas density is stable too.

The estimate of the intrinsic ellipsoidal core radius $\zeta_{\rm c}$ is prior dependent too. Assuming an ellipsoidal geometry, if the projected radius is very well constrained and the shape parameters are stable, the intrinsic radius can be deprojected through a geometrical factor, see Appendix A. Any variation in the geometrical factor, which can be induced by biased priors, should be compensated by variations in the core radius to keep the projected one nearly fixed. However, our estimates of $\zeta_{\rm c}$ are remarkably stable, which is a consequence of the well-determined shape.

The parametric model offers an excellent fit to the data, see Figs 3, 4, 6 and 8. Whereas the SB and the temperature constraints are uncorrelated, the SZ aperture photometry is not. If a point is above the fit, the following leans to do the same.

In Figs 3, 4, 6 and 8, we plot the predicted profiles as computed for the typical parameter values, i.e. the biweighted estimators of the marginalized posterior distributions, see Table 3. These are not the 'best-fitting' parameters found with a maximum likelihood analysis for the respective plots, seen in isolation, still they provide an excellent fitting to the plotted profiles.

We fitted the 2D-WL map, so the result for the averaged profile in circular annuli showed in Fig. 8 is given for illustration purposes. Similarly, in Fig. 3 we plotted the SB averaged in

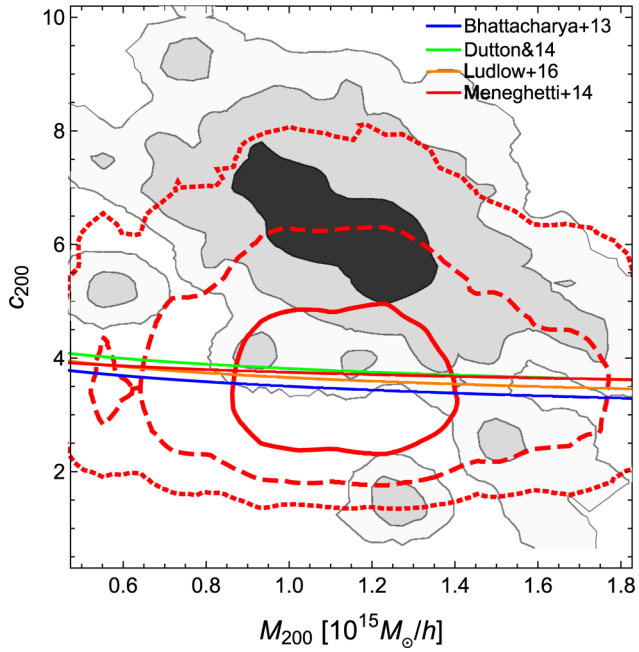

Figure 11. Marginalized probability distribution of mass and concentration. The grey shadowed regions include the 1σ, 2σ, 3σ confidence region in two dimensions, here defined as the regions within which the probability density is larger than $\exp[-2.3/2]$, $\exp[-6.17/2]$ and $\exp[-11.8/2]$ of the maximum, respectively. The regression assumed a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution (q-flat) and $q_{\rm ICM} \geq q$. The blue, green, orange and red lines plot the mass–concentration relations of Bhattacharya et al. (2013), Dutton & Macciò (2014), Ludlow et al. (2016) and Meneghetti et al. (2014), respectively. The red contours trace the predicted concentration from Meneghetti et al. (2014) given the observed mass distribution and the predicted scatter of the theoretical mass–concentration relation. If needed, published relations were rescaled to our reference cosmology.

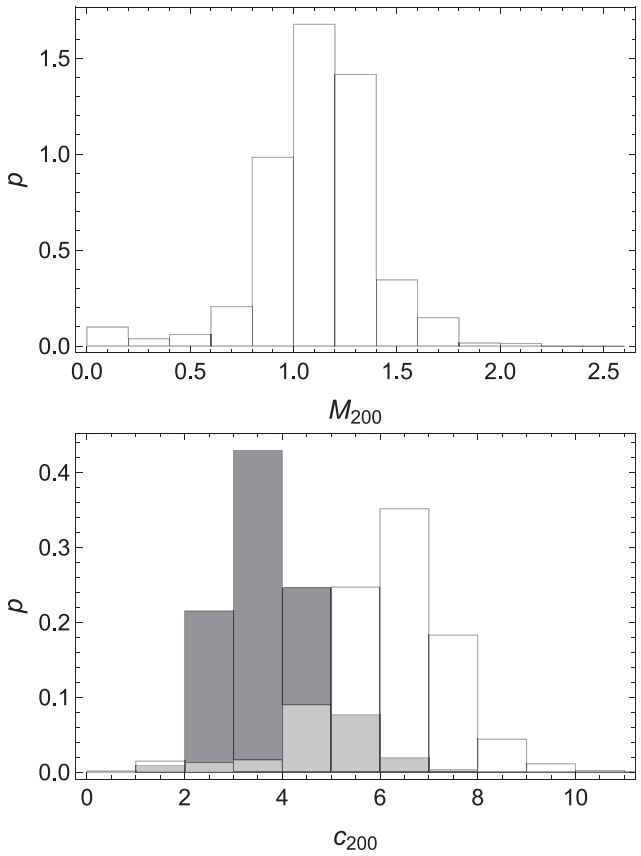

Figure 12. Marginalized PDFs (probability density functions), plotted as white histograms, of mass (top panel) and concentration (bottom panel). The regression assumed a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution (q-flat) and $q_{\rm ICM} \geq q$. The grey histogram in the bottom panel shows the theoretical prediction based on the inferred mass distribution and the scattered mass–concentration relation from Meneghetti et al. (2014).

circular annuli in the inner regions too, even though we fitted the 2D map.

7.1 Mass and concentration

MACS1206 is massive, $M_{200} = (1.1 \pm 0.2) \times 10^{15}\,{\rm M}_\odot\,h^{-1}$. The measured concentration, $c_{200} = 6.3 \pm 1.2$, as estimated in the reference regression, is slightly higher than but still consistent with predictions. Recent theoretical estimates based on N-body simulations of dark matter haloes (Bhattacharya et al. 2013; Dutton & Macciò 2014; Ludlow et al. 2016) graze the 68.3 per cent confidence region, see Fig. 11. The most sensible comparison is to Meneghetti et al. (2014), which studied a large set of nearly 1400 cluster-sized haloes simulated at high spatial and mass resolution from the MUSIC-2 N-body/hydrodynamical runs. The sample was originally constructed by selecting all haloes in the simulation box that were more massive than $10^{15}\,{\rm M}_\odot\,h^{-1}$ at redshift $z = 0$. The evolved haloes are distributed over the redshift range $0.25 \leq z \leq 0.67$ and are suitable to make predictions about several properties of the clusters included in the CLASH sample

Meneghetti et al. (2014) defined clusters as regular if they showed unperturbed X-ray SB distributions. These haloes have small centroid shift, ellipticity and power ratios, and they have large SB concentrations. Meneghetti et al. (2014) used the term regular with reference to the X-ray appearance, so that regular clusters may be dynamically unrelaxed.

From the comparison to the sample of simulated clusters above the completeness mass limit, MACS1206 is the only CLASH cluster less regular than the mean of the simulations. The comparison shows that the regularity of the CLASH clusters is not extreme, in the sense that the simulated sample has an extended tail of very regular clusters.

Meneghetti et al. (2014) measured the theoretical mass–concentration relation under different selection criteria and considered either projected or 3D concentrations and masses. The most sensible comparison for our analysis of MACS1206 is with the NFW fitting in 3D of the extended sample of simulated clusters, where no selection was applied except that based on the relaxation state.

The distribution of concentrations expected using the mass–concentration relation given the measured mass distribution is compared to the observed concentration in Fig. 11 and in the bottom panel of Fig. 12 for the marginalized distribution. Agreement is substantial.

Triaxial analyses favour agreement of measured concentrations of massive lensing clusters with theoretical predictions (Oguri et al. 2005; Sereno & Zitrin 2012). In fact, they do not suffer by the orientation bias affecting lensing selected clusters preferentially elongated along the line of sight. This configuration makes the concentration estimated under the spherical hypothesis biased

Table 4. Published mass and concentration measurements of MACS1206 from GL. The used data sets are in column 3, where γWL and μWL denote shear and magnification weak lensing data, respectively. Values in square brackets were assumed as fixed. The masses are in units of $10^{15}\,M_\odot\,h^{-1}$.

Author	Geometry	Data set	GL observatory	M_{200}	c_{200}
Foëx et al. (2012)	Spherical	γWL	CFHT	$1.06^{+0.20}_{-0.13}$	[4]
Merten et al. (2015)	Spherical	γWL, SL	Subaru, *HST*	0.86 ± 0.11	4.3 ± 1.5
Umetsu et al. (2016)	Spherical	γWL, μWL, SL	Subaru, *HST*	1.28 ± 0.29	3.7 ± 1.1
This work	Spherical	γWL, μWL, SL	Subaru, *HST*	1.06 ± 0.16	5.0 ± 0.7
This work	Triaxial	γWL, μWL, SL, X, SZ	Subaru, *HST*	1.14 ± 0.23	6.3 ± 1.2

Table 5. Ellipsoidal and spherically enclosed mass estimates, as derived from the regression assuming a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution and $q_{\rm ICM} \geq q$. The ellipsoidal mass M_Δ is computed within the ellipsoid of semi-major axis ζ_Δ. The overdensity radii are given in units of Mpc h^{-1}. The enclosed masses are in units of $10^{15}\,M_\odot\,h^{-1}$. Typical values and dispersions are computed as biweighted estimators.

Overdensity	Ellipsoidal		Spherically enclosed	
Δ	ζ_Δ	M_Δ	r_Δ	$M_{\rm sph}(< r_\Delta)$
2500c	0.65 ± 0.09	0.42 ± 0.07	0.44 ± 0.02	0.39 ± 0.06
500c	1.40 ± 0.20	0.85 ± 0.16	0.95 ± 0.05	0.79 ± 0.14
200c	2.10 ± 0.31	1.14 ± 0.23	1.43 ± 0.09	1.08 ± 0.20
$\Delta_{\rm vir}$	2.49 ± 0.37	1.27 ± 0.27	1.69 ± 0.11	1.21 ± 0.24
200m	2.68 ± 0.40	1.33 ± 0.29	1.83 ± 0.12	1.27 ± 0.26

high. The opposite holds for clusters elongated orthogonally to the line of sight.

MACS1206 is X-ray selected, and it is elongated in the plane of the sky, see Section 7.3. Mass and concentration derived under the spherical assumption are then biased low, see Table 3. Our unbiased result is in good agreement with theoretical predictions.

Our estimates under the spherical hypothesis are consistent with previous lensing-based results from the literature, see Table 4. Differences with the triaxial estimate show that the geometrical bias is significant when compared to the statistical errors.

We list the values of ellipsoidal and spherical overdensity mass in Table 5. Ellipsoidal or spherically enclosed mass estimates are similar at a given overdensity.

7.2 Matter shape

MACS1206 shows a triaxial shape, see Figs 13 and 14. The axial ratios given the cluster mass and redshift are compatible with the theoretical predictions, even though the mean intermediate-to-major axial ratio slightly exceeds expectations.

In the more conservative reference analysis, when we only require the gas to be rounder than the matter, we cannot preferentially distinguish the prolate from the oblate configuration, see Fig. 13. However, with the slightly more informative prior assuming the same triaxial parameter for both matter and gas ($\mathcal{T}_{\rm mat} = \mathcal{T}_{\rm ICM}$), prolate configurations are preferred.

7.3 Halo orientation

We observe MACS1206 face-on, i.e. the main axis is near the plane of the sky. Large inclination angles, nearly orthogonal to the line of sight, are preferred, see Fig. 15. This is also confirmed by the analysis of the elongation parameter, which shows that the width in the plane of the sky is larger than the size along the line of sight,

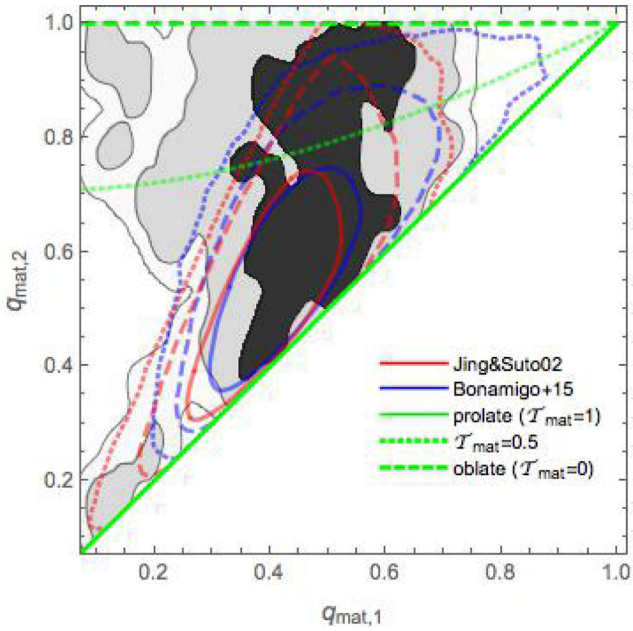

Figure 13. Probability distribution of the matter axial ratios. The regression assumed a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution (q-flat) and $q_{\rm ICM} \geq q$. The grey shadowed regions include the 1σ, 2σ, 3σ confidence region in two dimensions, here defined as the region within which the probability density is larger than $\exp[-2.3/2]$, $\exp[-6.17/2]$ and $\exp[-11.8/2]$ of the maximum, respectively. The blue and red contours plot the theoretical predictions from Bonamigo et al. (2015) and Jing & Suto (2002), respectively, smoothed for the inferred mass distribution. Contours are drawn at 1σ (full), 2σ (dotted) and 3σ (dashed). The green full, dashed and long dashed lines denote the loci of points corresponding to prolate ($T_{\rm mat} = 1$), triaxial ($T_{\rm mat} = 0.5$) and oblate ($T_{\rm mat} = 0$) haloes.

$e_\parallel < 1$, see Fig. 16. This effect is more pronounced for the matter distribution $e_{\rm mat,\parallel} \lesssim e_{\rm ICM,\parallel}$, but this is expected given the priors.

The second Euler angle φ is poorly constrained with a bimodal distribution, whereas the precision on ψ is reflective of the accuracy on the measured orientation angles of matter (from lensing) and gas (from X-ray SB).

Based on the highly elliptical mass distribution in projection as inferred at large cluster radii from the Subaru WL analysis, Umetsu et al. (2012) argued that that major axis of MACS1206 is not far from the sky plane. We could unambiguously prove this based on our multi-probe 3D analysis. This conclusion could not be obtained based on SZ data alone. Romero et al. (2016) noted that the Bolocam contours of MACS1206 do not exhibit much ellipticity and, by comparison with X-ray data analysed under the hypothesis of circular symmetry in the plane of the sky, found a major-to-minor axial ratio of 1.24 ± 0.29, where the major axis is along the line of sight.

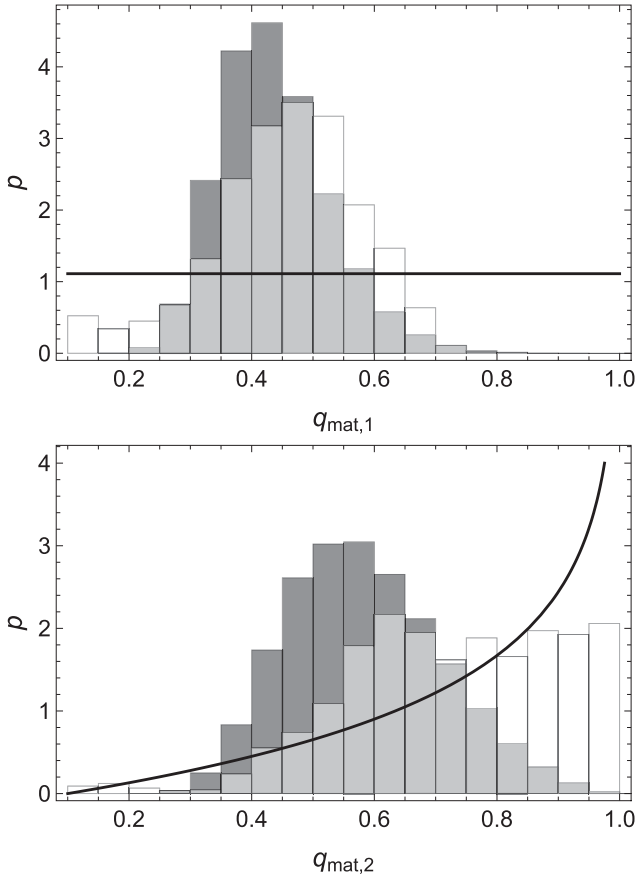

Figure 14. Marginalized PDFs (plotted as white histograms) of the minor-to-major (top panel) and intermediate-to-major (bottom) axial ratios of the matter distribution. The regression assumed a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution (q-flat) and $q_{\rm ICM} \geq q$. Grey histograms show the theoretical predictions based on the inferred mass distribution and the predicted probability distribution for axial ratios from Bonamigo et al. (2015). The black lines show the marginalized a priori distributions under the q-flat prior.

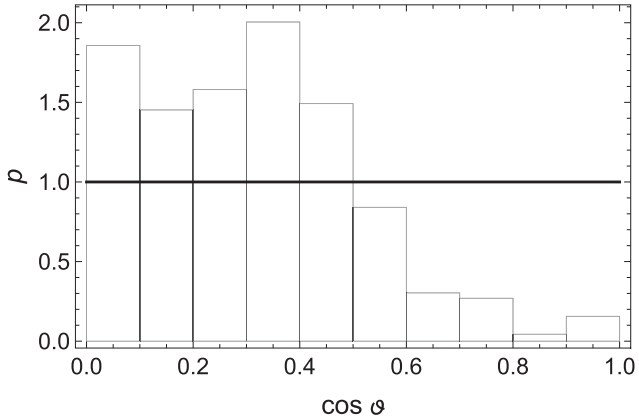

Figure 15. Marginalized PDFs (plotted as white histogram) of the cosine of the inclination angle, $\cos\vartheta$. The regression assumed a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution (q-flat) and $q_{\rm ICM} \geq q$. The black line denotes random orientation.

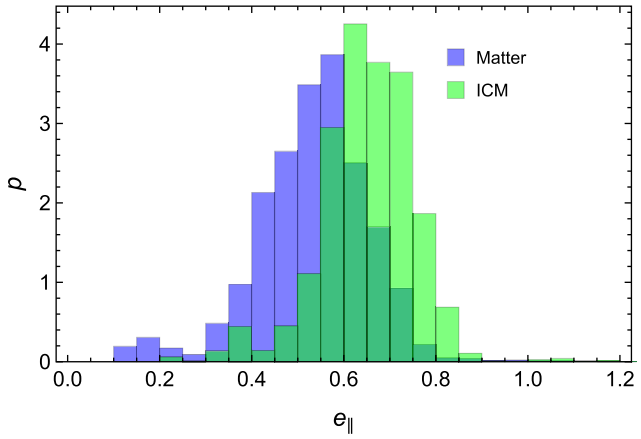

Figure 16. Marginalized PDFs (plotted as white histogram) of the line-of-sight elongation parameter $e_\parallel$ of the total matter (blue) and of the gas distribution (green). The regression assumed a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution and $q_{\rm ICM} \geq q$.

7.4 Gas distribution

The inferred gas distribution is quite common to massive clusters. The slope β is ∼0.6, and there is some evidence for an inner spike ($\eta \sim 0.6$). There is no clear evidence for a truncation of the density profile, which would occur at $\zeta_{\rm t} \gtrsim 1.5\,{\rm Mpc}\,h^{-1}$, beyond the observational range. This makes the constraints on the outer slope $\gamma_{\rm ICM}$ of poor significance.

The temperature distribution shows a cool core with a quite large cool core radius, $\zeta_{\rm cc} \sim \zeta_{\rm c}$. However, the temperature profile is sampled in just one point in the inner $100\,{\rm kpc}\,h^{-1}$, the presumptive size of the cool core, making this estimate a likely artefact of poor sampling and extrapolation.

7.5 Gas shape

A consequence of the pressure equilibrium is the *X-ray shape theorem* (Buote & Canizares 1994, 1996, 1998), i.e. the gas in strict HE follows the isopotential surfaces of the underlying matter distribution. If the gas pressure can be written as a function of gas density and temperature and the gas is adequately described by a single phase, the hydrostatic equation demands that the potential, the gas density, the gas pressure, the gas temperature and the X-ray volume emissivity all share the same constant surfaces in three dimensions. This geometric test for dark matter is robust and independent of the temperature profile, which can be poorly constrained.

The first application of the shape test compared the 2D ellipticities of the X-ray SB with the gravitational potential after projection in the hypothesis of spheroidal symmetry for the emitting system (Buote & Canizares 1994, 1996, 1998). Here, we can directly compare the axial ratios in three dimensions without restricting assumptions on the cluster shape.

Our method can determine the intrinsic structure, shape and orientation of the gas without a priori assuming equilibrium. The direct comparison of gas shape, as determined by the regression, to the potential shape, as computed from the inferred mass distribution, is then a test of equilibrium.

The regular morphology of MACS1206 makes it an excellent candidate to the application of the shape theorem. The projected gravitational potential of MACS1206 as inferred from galaxy kinematics under the hypothesis of spherical symmetry agrees well

Table 6. Comparison of the axial ratios and derived geometrical quantities of the distributions of gas (column 1) and gravitational potential (columns 2 and 3). The distributions share the same orientation, see Table 3. The regression assumed a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution (q-flat) and $q_{\rm ICM} \geq q$. The shape parameters of the potential are computed considering the effective axial ratios at $r_{200}/3$ (column 2) and at $2\,r_{200}/3$ (column 3). Projected orientation angles are measured in degrees north-over-east.

	ICM	$\Phi(r_{200}/3)$	$\Phi(2\,r_{200}/3)$
q_1	0.59 ± 0.11	0.77 ± 0.06	0.80 ± 0.05
q_2	0.78 ± 0.06	0.88 ± 0.08	0.90 ± 0.07
ϵ	0.22 ± 0.02	0.12 ± 0.07	0.10 ± 0.06
θ_ϵ	-58.52 ± 3.53	-56.50 ± 8.07	-56.51 ± 8.06
$e_\parallel$	0.66 ± 0.09	0.83 ± 0.05	0.85 ± 0.05
$e_1/e_{\rm mat,1}$	0.93 ± 0.06	0.72 ± 0.04	0.67 ± 0.04
$\mathcal{T}$	0.63 ± 0.16	0.54 ± 0.31	0.54 ± 0.31

with the reconstruction of the gravitational potential from weak and strong GL or X-ray measurements (Stock et al. 2015).

We computed the effective shape of the potential as detailed in Appendix B. The inferred gas shape is broadly compatible with the potential, see Table 6 and Figs 17 and 18, even though rounder potential shapes are compatible with the data. In particular, the ratio of the measured eccentricities slightly exceeds the expected value of ∼0.7, see the bottom panel of Fig. 17, and the projected isocontours are more elongated, see Table 6. The compatibility of gas and potential shapes is best evidenced by the marginalized distributions of the axial ratios, see Fig. 18.

The shape theorem is expected to be violated to some degree. The degree of HE was investigated by Biffi et al. (2016) on a sample of 29 massive clusters extracted from cosmological hydrodynamical simulations including several physical processes, i.e. stellar and AGN feedback. The radial balance between the gravitational and hydrodynamical forces can be assessed via comparison of the gas accelerations generated. They found an average deviation from equilibrium of 10–20 per cent out to the virial radius. However, the result is strongly dependent on the properties of the selected clusters.

7.6 Triaxiality

A working hypothesis suitable for clusters in near equilibrium is that the triaxial parameters of the gas and of the matter distribution are similar. This is the case for MACS1206. First, the X-ray and the lensing maps share the same orientation in the plane of the sky, see Fig. 19.

Secondly, our direct measurements of the triaxial parameters are compatible, see Fig. 20. However, when no strong assumption is made a priori on the matter shape and on the relation between the axial ratios, as in our reference case where we just assumed a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution, the triaxiality parameter of the matter halo cannot be significantly constrained and its final distribution follows the initial prior.

In contrast, the triaxiality parameter of the gas distribution is strongly constrained by the data.

7.7 Gas fraction

We computed the gas mass fraction in spherically enclosed regions, $f_{\rm gas} = M_{\rm sph,gas}(<r)/M_{\rm sph,tot}(<r)$, using the posterior samples of the ellipsoidal cluster model. The estimate does not rely on the assumption of equilibrium. Despite being computed in spherical regions, this estimate of the cumulative gas mass fraction is free from the assumptions of spherical symmetry.

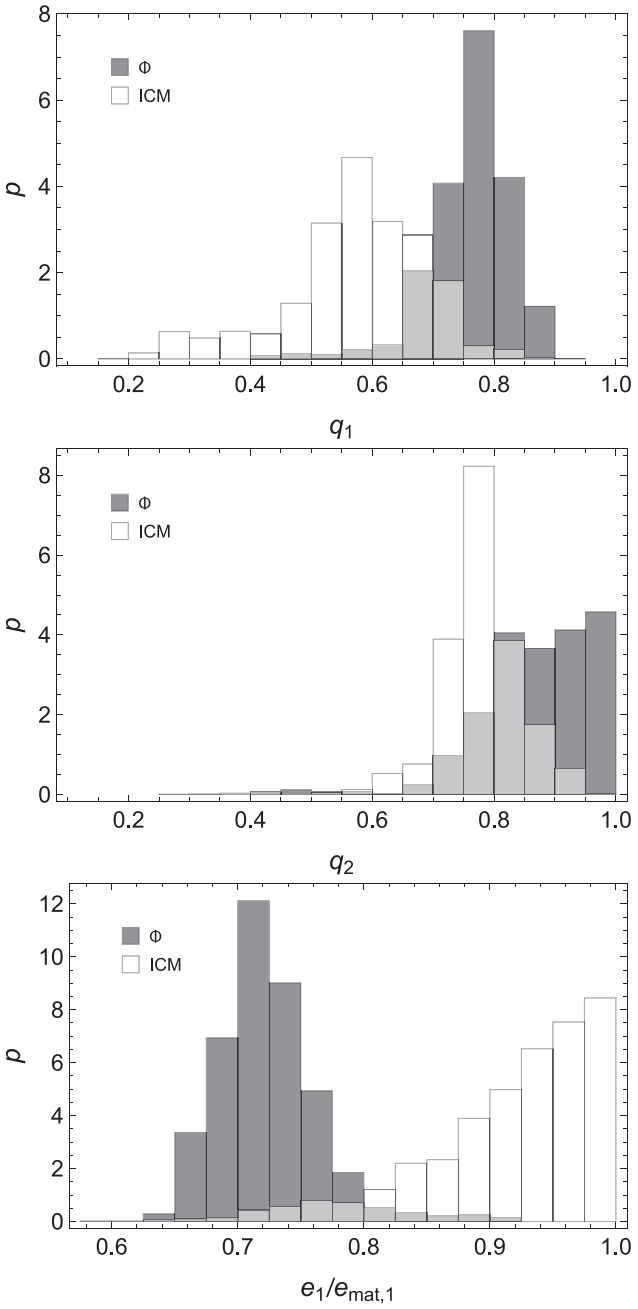

Figure 17. Marginalized PDFs of the axial ratios and the eccentricity of the gas distribution (plotted as white histograms) and of the gravitational potential (grey histograms). From top to bottom: minor-to-major axial ratio, intermediate-to-major axial ratio and eccentricity. The regression assumed a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution (q-flat) and $q_{\rm ICM} \geq q$.

The gas fraction profile is shown in Fig. 21. At the overdensity spherical radius $r_{\rm sph,\,2500}$, we found $f_{\rm gas} = 0.128 \pm 0.014$; within $r_{\rm sph,\,500}$, we found $f_{\rm gas,\,500} = 0.177 \pm 0.023$. Based on scaling relations of X-ray luminosity, temperature and gas mass, and employing total masses from weak GL measurements, Mantz et al. (2016) found a constraint on the gas mass fraction of $f_{\rm gas,\,500} = 0.125 \pm 0.005$ in a sample of 40 clusters identified as being dynamically

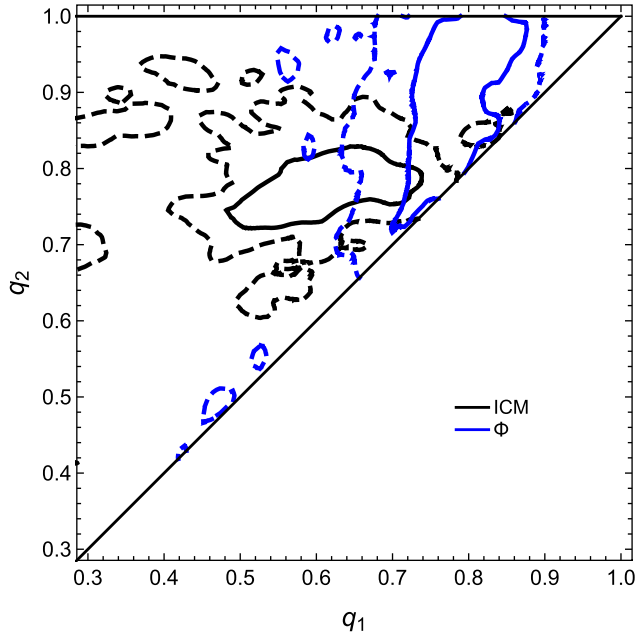

Figure 18. Probability distribution of gas (black) and potential (blue) axial ratios. The regression assumed a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution (q-flat) and $q_{\rm ICM} \geq q$. Contours are drawn at 1σ (full), and 2σ (dashed) confidence region in two dimensions, here defined as the region within which the probability density is larger than $\exp[-2.3/2]$ or $\exp[-6.17/2]$ of the maximum, respectively.

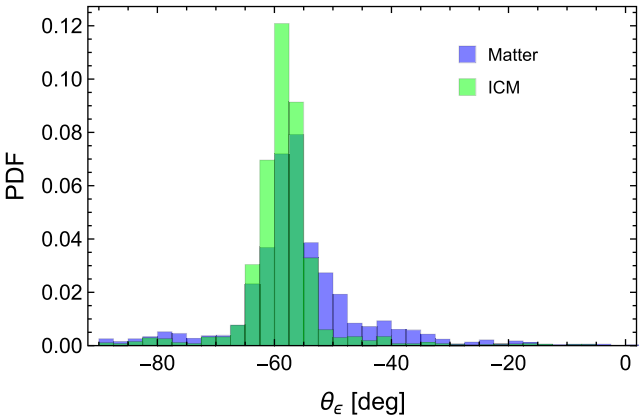

Figure 19. Marginalized PDFs (plotted as histograms) of the orientation angle in the plane of the sky θ_ϵ (measured in degrees north-over-east NE) of the total matter (blue) and of the gas distribution (green). The regression assumed a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution and $q_{\rm ICM} \geq q$.

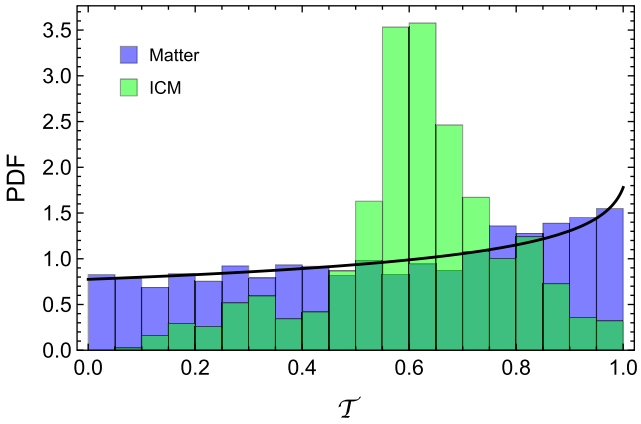

Figure 20. Marginalized PDFs (plotted as histograms) of the triaxial parameter $\mathcal{T}$ of the total matter (blue) and of the gas distribution (green). The regression assumed a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution and $q_{\rm ICM} \geq q$. The black line shows the prior.

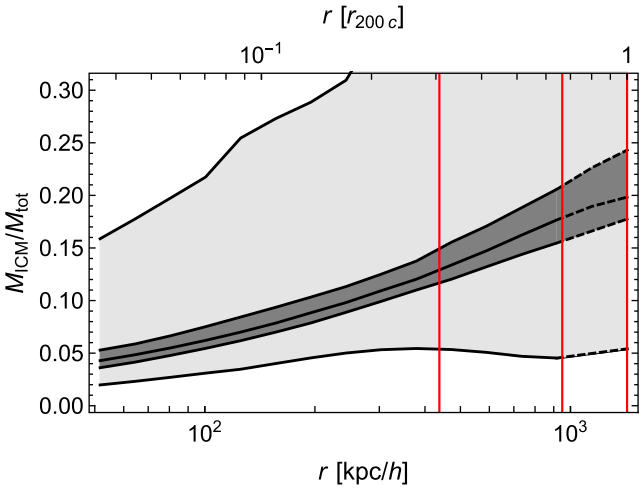

Figure 21. Ratio of spherically enclosed gas mass to total mass as a function of the spherical radius r. The regression assumed a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution and $q_{\rm ICM} \geq q$. The middle line tracks the median. The grey shaded regions represent the 68.3 and 95.4 per cent regions around the median, as computed from symmetric quantiles. Dashed lines indicate extrapolations beyond the region covered by X-ray observations. Vertical red lines mark $r_{\rm sph,\,2500}$, $r_{\rm sph,\,500}$ and $r_{\rm sph,\,200}$, as measured in spherical regions.

relaxed and hot, consistent with previous measurements using hydrostatic mass estimates for relaxed clusters (Mantz et al. 2016). Accounting for the intrinsic scatter affecting gas fraction estimates, our result is consistent with this general trend.

Umetsu et al. (2012) combined lensing results with *Chandra* gas mass measurements and found a cumulative gas mass fraction of $f_{\rm gas} = 0.137^{+0.045}_{-0.030}$ at 1 Mpc $= 0.7$ Mpc h^{-1} under the hypothesis of spherical symmetry. Our extrapolated result is $f_{\rm gas}(<1\,{\rm Mpc}) = 0.157 \pm 0.019$, in good agreement.

When compared to the cosmic baryon fraction $f_{\rm B} \simeq 0.156$ (Planck Collaboration XIII 2016a), we find no significant gas depletion, $Y_{\rm B} = f_{\rm gas,\,500}/f_{\rm B} \sim 1$. Some depletion is predicted by hydrodynamic simulations including radiative cooling, star formation and AGN feedback (Battaglia et al. 2013; Planelles et al. 2013). Planelles et al. (2013) found that the gas fraction of massive clusters is nearly independent of the physical processes, and it is characterized by a negligible redshift evolution, $Y_{\rm B} = 0.85 \pm 0.03$ at r_{500}. At smaller radii, $Y_{\rm B}$ slightly decreases, in agreement with what found in MACS1206, by an amount that depends on the physics included in the simulations.

MACS1206 may be baryon rich. However, angular variance of the anisotropically distributed gas that originates from the recent formation epoch of clusters and from the strong internal baryon-to-dark-matter density bias, along with density clumpiness, can bias high the measured gas fraction too (Battaglia et al. 2013).

7.8 Non-thermal pressure

Non-thermal pressure $P_{\rm nth}$ can significantly contribute to the overall balance (Rasia et al. 2014). Neglecting the contribution from bulk and/or turbulent motions can systematically bias low the X-ray mass determination (Meneghetti et al. 2010; Rasia et al. 2012). High-resolution cosmological simulations showed that the contribution can be significant in the core of relaxed clusters too (Lau, Kravtsov & Nagai 2009; Molnar et al. 2010).

Based on 10 simulations of massive relaxed clusters, Molnar et al. (2010) found that pressure support from subsonic random gas motions can contribute up to 40 per cent in the inner regions and up to 20 per cent within one tenth of the virial radius. The non-thermal contribution is expected to increase with radius in the very outer regions (Shaw et al. 2010; Martizzi & Agrusa 2016). Lau et al. (2009) found a non-thermal pressure contribution of the order of 5–15 per cent at about one tenth of the viral radius, also increasing with radius in the outer regions.

Chiu & Molnar (2012) tested the assumption of strict HE in MS2137.3−2353 and found a significant contribution from non-thermal pressure in the core region, independently of the assumed shape of the cluster.

Since our inference did not rely on the assumption of HE, the equilibrium hypothesis can be used a posteriori to assess the level of non-thermal pressure. The generalized equilibrium condition reads

$$\nabla P_{\rm tot} = -\rho_{\rm ICM} \nabla \Phi_{\rm mat}, \tag{53}$$

where $P_{\rm tot} (= P_{\rm th} + P_{\rm nth})$ is the total pressure, $\rho_{\rm ICM}$ the gas density and $\Phi_{\rm mat}$ the gravitational potential. When the equilibrium is hydrostatic, the pressure is only thermal, $P_{\rm th} = k_{\rm B} T n_{\rm ICM}$ for an ideal gas.

Due to the shape theorem for gas in equilibrium (Buote & Canizares 1994), we considered the posteriori probabilities for the cluster parameters inferred under the prior of flat distribution for the matter shape and $q_{\rm ICM} = q_\Phi$. The gravitational potential of the ellipsoidal NFW halo was computed using the formulae in Lee & Suto (2003).

The result is presented in Fig. 22, where the ratio of thermal to equilibrium gas pressure, $P_{\rm th}/P_{\rm tot}$, is plotted as function of the ellipsoidal radius of the ICM distribution, $\zeta_{\rm ICM}$.

We find that the contribution of thermal pressure is dominant. The apparent drop in the inner regions can be an artefact due to the poor modelling of the temperature profile in the inner ∼100 kpc, see Fig. 4.

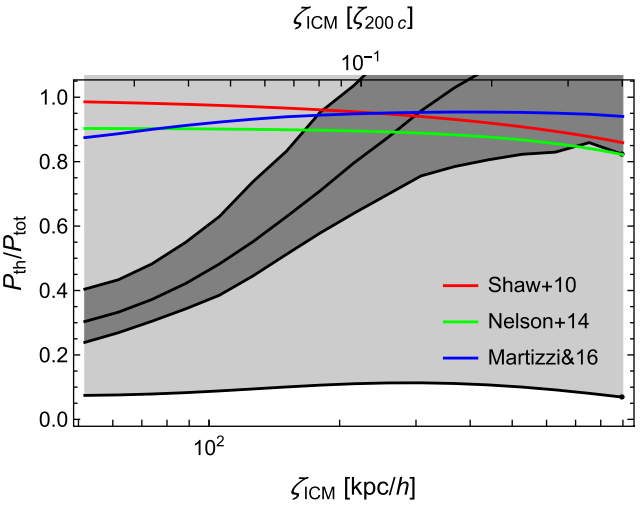

Figure 22. Ratio of the thermal gas pressure to the total equilibrium pressure as a function of the ellipsoidal radius $\zeta_{\rm ICM}$ measured along the major axis of the gas distribution. The regression assumed a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution and $q_{\rm ICM} = q_\Phi$. The middle line tracks the median. The grey shaded regions show the 68.3 and the 99.7 per cent regions around the median. The red, green and blue lines plot the theoretical predictions from Shaw et al. (2010), Nelson et al. (2014) and Martizzi & Agrusa (2016), respectively.

8 SYSTEMATICS

The level of systematics errors plaguing an analysis can be checked by comparing results derived with distinct methodologies or exploiting different data sets. We considered independent treatments of either the lensing, X-ray or SZ part of our analysis. We only changed one module of the total χ^2 for each test, keeping all the other probes fixed. Results are summarized in Table 7. Whereas we considered different data sets or likelihoods, we used the same Bayesian inference and the same treatment of the posterior. As priors, we assumed a flat distribution for the minor-to-major matter axial ratio and $q_{\rm ICM} \geq q$.

8.1 Lensing

As an alternative to the lensing treatment, we considered the convergence maps obtained with the SaWLens (Strong and Weak Lensing) method (Merten et al. 2015), which can consistently combine WL and SL with no a priori assumptions about the underlying mass distribution. The convergence map can then be fitted to get unbiased parameters for the lensing model of choice.

SaWLens performs a reconstruction of the lensing potential on adaptively refined grids. For the inner regions of MACS1206, Merten et al. (2015) used the position and redshift of 33 critical line estimators derived from the 13 multiple-image systems (four of which spectroscopically confirmed) and measured the shapes of ≲600 background galaxies in seven broad-band Advanced Camera for Surveys filters. In the outer regions, Merten et al. (2015) exploited the Subaru data for the shear analysis.

For our analysis, we considered two different grid sizes. Even though each grid exploits all data sets, it is most sensitive to some of them. The low-resolution grid, covering 25 arcmin × 25 arcmin with a pixel resolution of 50 arcsec, is well suited for WL on the wide field, and is dominated by the data from the Subaru ground-based telescope.

The fine grained grid, covering the inner 150 arcsec × 150 arcsec with a pixel resolution of ∼8 arcsec, traces SL features near the innermost core of the cluster and exploits the WL constraints from the *HST* on a much smaller field of view but with considerably higher spatial resolution.

When combining the two maps, we excluded the inner square of side 2 arcmin of the low-resolution grid to prevent overlap with the SL region already sampled by the fine-resolution grid.

Statistical uncertainty within SaWLens is based on bootstrap resamplings and noise realizations. However, the number of resampled maps is not big enough to compute the uncertainty covariance matrix of the 2D-WL map. In these cases, regularization schemes already employed in lensing analyses (Umetsu et al. 2012, 2015) are not so effective. In fact, the covariance matrix is singular not due to intrinsic features of the data sample but due to the number of bootstrap realizations being smaller than or comparable to the number of pixels. We then assumed the covariance matrix to be diagonal.

Table 7. Results under different methods or data sets. All regressions assumed a flat prior for the axial ratio $q_{\rm mat,1}$ of the matter distribution and $q_{\rm ICM} \geq q$. Differently from the reference analysis (reported in column 2), we considered the SaWLens convergence maps (column 3), a 1D analysis of the X-ray SB (column 4) or a treatment of the SZ effect based on Bolocam data only (column 5). Units are as in Table 1. Typical values and dispersions are computed as biweighted estimators.

	Reference	SaWLens	1D-X	No *Planck*
M_{200}	1.137 ± 0.229	1.115 ± 0.113	1.035 ± 0.196	1.071 ± 0.187
c_{200}	6.277 ± 1.188	4.400 ± 0.556	6.247 ± 1.106	5.586 ± 0.879
$q_{\rm mat,1}$	0.466 ± 0.119	0.540 ± 0.145	0.434 ± 0.114	0.557 ± 0.107
$q_{\rm mat,2}$	0.735 ± 0.176	0.794 ± 0.094	0.688 ± 0.158	0.749 ± 0.140
$\cos\vartheta$	0.297 ± 0.204	0.302 ± 0.239	0.266 ± 0.229	0.333 ± 0.209
φ	−0.438 ± 1.609	−0.549 ± 1.530	0.200 ± 1.306	0.096 ± 0.987
ψ	1.027 ± 0.123	0.979 ± 0.149	0.936 ± 0.203	0.975 ± 0.156
$q_{\rm ICM,\,1}$	0.587 ± 0.109	0.614 ± 0.122	0.528 ± 0.124	0.684 ± 0.090
$q_{\rm ICM,\,2}$	0.779 ± 0.057	0.778 ± 0.069	0.787 ± 0.096	0.814 ± 0.067
n_0	0.010 ± 0.001	0.010 ± 0.001	0.011 ± 0.001	0.011 ± 0.001
$\zeta_{\rm c}$	169.780 ± 12.02	165.950 ± 5.880	174.490 ± 6.728	145.060 ± 5.839
$\zeta_{\rm t}/\zeta_{\rm c}$	8.250 ± 1.635	8.057 ± 1.290	7.272 ± 1.293	8.041 ± 1.092
β	0.600 ± 0.025	0.592 ± 0.027	0.613 ± 0.024	0.568 ± 0.020
η	0.627 ± 0.047	0.643 ± 0.054	0.592 ± 0.042	0.574 ± 0.046
$\gamma_{\rm ICM}$	1.810 ± 0.632	2.336 ± 0.434	2.093 ± 0.477	2.151 ± 0.527
T_0	24.245 ± 2.646	21.469 ± 2.961	21.426 ± 2.885	18.384 ± 1.790
$\zeta_{cT}/\zeta_{\rm c}$	9.069 ± 1.034	8.518 ± 0.882	8.666 ± 1.285	17.646 ± 2.173
c_T	2.576 ± 0.313	2.790 ± 0.158	2.891 ± 0.097	2.201 ± 0.482
$T_{\rm cc}$	4.221 ± 1.328	3.654 ± 1.107	3.894 ± 0.861	4.316 ± 1.258
$\zeta_{\rm cc}/\zeta_{\rm c}$	1.140 ± 0.211	0.986 ± 0.132	0.950 ± 0.126	1.006 ± 0.133

This conservative assumption may underestimate the χ^2, and consequently overestimate the confidence regions of the parameters but can still recover the main features of mass distribution.

The pixels of the low-resolution map are large enough to contain a significant number of background galaxies (∼10 galaxies pixel^{-1}). Even though the reconstruction of the convergence map from the shear field is not local, the main contribution comes from the galaxies in the pixel itself, which reduces the correlations among pixels.

Since we modelled the whole cluster as one three-dimensional NFW halo and we do not assume any subhaloes, our conservative approach can be still effective. However, due to the limited knowledge of the covariance matrix, we used the SaWLens map only for model comparison.

Even though SaWLens fits all data at once, the different grids and the different weights of the lensing data sets make the low- and the fine-resolution maps nearly independent. For the inner regions, we considered the azimuthally averaged convergence in six equally spaced angular annuli between 5 arcsec and 1 arcmin. To measure the uncertainties, we re-computed the convergence for the noise realizations and we computed the uncertainty covariance matrix from the sample distribution.

As far as data sets are considered, the SaWLens analysis exploited the shear signal in the outer regions, whereas our reference analysis considered magnification and number counts too.

The results based on the SaWLens maps are in good agreement with our reference model even though the SaWLens-based concentration is smaller and uncertainties on mass and concentration are smaller. As a consequence of the smaller concentration, the shape is a bit rounder too. Even though the shift is compatible with the statistical uncertainties of our reference model, we notice that the NFW profile is not an excellent modelling of the SaWLens convergence profile, which is steep in the inner regions and flat in the outer regions, see Fig. 8. In fact, the smaller formal uncertainties are more a signal of problems in modelling than of increased accuracy, which is not reasonable since the data set is the same apart from the magnification data. However, to fully understand if this deviation is statistical significant, we would need an accurate computation of the SaWLens uncertainty covariance matrix for the WL regime, which at this time is still missing.

8.2 X-ray SB

To check the analysis of the X-ray SB, we considered an alternative procedure. The measurements of ellipticity and slope are nearly uncorrelated (De Filippis et al. 2005). The analysis of the 2D SB can then be approximated as a 1D problem. The ellipticity and orientation angle are evaluated in a first step, and the SB is measured in elliptical annuli (following the morphology determined before) in a second step (Sereno et al. 2012). The χ^2 function can be written as

$$\chi^2_{\rm SB} = \sum_{i=1}^{N_S} \left(\frac{S_{{\rm X},i} - \hat{S}_{{\rm X},i}}{\delta_{S,i}} \right)^2 + \left(\frac{\epsilon_{\rm ICM} - \hat{\epsilon}_{\rm ICM}}{\delta\epsilon_{\rm ICM}} \right)^2 + \left(\frac{\theta_{{\rm ICM},\epsilon} - \hat{\theta}_{{\rm ICM},\epsilon}}{\delta\theta_{{\rm ICM},\epsilon}} \right)^2 , \quad (54)$$

with $\hat{S}_{\rm X}$, $\hat{\epsilon}_{\rm ICM}$ and $\hat{\theta}_{{\rm ICM},\epsilon}$ are the model predictions for the corresponding X-ray observables with measurements uncertainties $\delta_{S,\,i}$, $\delta\epsilon_{\rm ICM}$ and $\delta\theta_{{\rm ICM},\,\epsilon}$, respectively. $\epsilon_{\rm ICM}$ and $\theta_{{\rm ICM},\,\epsilon}$ are the ellipticity and the orientation angle, respectively, of the X-ray isophotes in the plane of the sky.

The ellipticity and the orientation angle can be determined as the parameters of the ellipse enclosing a fraction of the total cluster light. We considered light thresholds of 50, 60, 70 and 80 per cent, and we determined the parameters as the biweight estimators. We found $\epsilon_{\rm ICM} = 0.23 \pm 0.07$ and orientation angle $\theta_{{\rm ICM},\epsilon} = -54.2 \pm 2.7$ deg (measured north over east) in remarkable agreement with the 2D analysis, see Table 6. These parameters

were then used to define the elliptical bins in which the SB profile is resolved.

The agreement with the full 2D analysis is substantial, showing that modelling the gas shape as an ellipsoid with fixed axial ratios and fixed orientation is a very good approximation at the present level of accuracy and precision.

8.3 SZe

The integrated Compton parameter was alternatively computed relying on the Bolocam data only. Results are in very good agreement with the reference analysis exploiting *Planck* data too. The *Planck* data are important to better resolve the elongation of the gas distribution, which favours more triaxial structure. *Planck* data also improve the modelling of the gas profile at large radii, favouring a slightly larger core radius and a temperature profile decrement at large radii.

9 CONCLUSIONS

MACS1206 is a remarkably regular, massive, face-on cluster. Its concentration is in line with theoretical predictions. The measured triaxial shape is common to clusters simulated in the ΛCDM concordance cosmology. The gas has settled in the potential well and its distribution traces the isopotential surfaces. This is evidence under the shape theorem that the cluster is in pressure equilibrium. The level of baryonic depletion is small. The thermal pressure can balance the cluster in HE.

Multi-probe analyses are needed to achieve one of the main goals of precise and accurate cosmology: unbiased cluster mass measurements at the per cent level (Benson et al. 2013). The compatibility of independent observables, from X-ray to lensing to radio observations of the SZe, under the same coherent picture guarantees that the measurements of mass and other intrinsic properties are unbiased and systematic-free. In our modelling, we do not rely on the assumptions of spherical symmetry or HE, which could bias results.

At the same time, the joint exploitation of different data sets improves the statistical accuracy and enables us to expand the scope of the analysis. We can measure the cluster shape and the non-thermal pressure too.

The multi-wavelength analysis we presented in this paper is a development of the Bayesian method introduced by Sereno et al. (2013). We used the same modelling of the mass and gas distribution and the same inference scheme but we extended and refined the priors on mass and gas; we fitted the 2D map of the X-ray SB; we used the exact likelihoods for GL and X-ray plus SZe rather than approximating them as smooth kernel or multi-variate distributions of the projected parameters. Whereas in the ideal case the two approaches are equivalent, the new approach is more flexible and relies less on geometrical assumptions.

In forthcoming papers, we plan to analyse the full CLASH sample.

ACKNOWLEDGEMENTS

The authors thank M. Bonamigo and M. Limousin for stimulating discussions and C. Grillo for useful comments. SE, MM and MS acknowledge the financial contribution from contracts ASI-INAF I/009/10/0, PRIN-INAF 2012 'A unique dataset to address the most compelling open questions about X-Ray Galaxy Clusters', and PRIN-INAF 2014 1.05.01.94.02 'Glittering Kaleidoscopes in the sky: the multifaceted nature and role of galaxy clusters'. SE acknowledges the financial contribution from contracts NARO15 ASI-INAF I/037/12/0 and ASI 2015-046-R.0. MM acknowledges support from the Italian Ministry of Foreign Affairs and International Cooperation, Directorate General for Country Promotion, and from ASI via contract ASI/INAF/I/023/12/0. JS was supported by NSF/AST-1617022. KU acknowledges support from the Ministry of Science and Technology of Taiwan through grants MOST 103-2112-M-001-030-MY3 and MOST 103-2112-M-001-003-MY3. This research has made use of NASA's Astrophysics Data System (ADS) and of the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

REFERENCES

Alam S. M. K., Ryden B. S., 2002, ApJ, 570, 610
Arnaud K. A., 1996, in Jacoby G. H., Barnes J., eds, ASP Conf. Ser. Vol. 101, Astronomical Data Analysis Software and Systems V. Astron. Soc. Pac., San Francisco, p. 17
Baldi A., Ettori S., Molendi S., Gastaldello F., 2012, A&A, 545, A41
Basilakos S., Plionis M., Maddox S. J., 2000, MNRAS, 316, 779
Battaglia N., Bond J. R., Pfrommer C., Sievers J. L., 2013, ApJ, 777, 123
Beers T. C., Flynn K., Gebhardt K., 1990, AJ, 100, 32
Benson B. A., Church S. E., Ade P. A. R., Bock J. J., Ganga K. M., Henson C. N., Thompson K. L., 2004, ApJ, 617, 829
Benson B. A. et al., 2013, ApJ, 763, 147
Bhattacharya S., Habib S., Heitmann K., Vikhlinin A., 2013, ApJ, 766, 32
Biffi V. et al., 2016, ApJ, 827, 112
Binggeli B., 1980, A&A, 82, 289
Binney J., 1985, MNRAS, 212, 767
Binney J., de Vaucouleurs G., 1981, MNRAS, 194, 679
Birkinshaw M., 1999, Phys. Rep., 310, 97
Biviano A. et al., 2013, A&A, 558, A1
Bonamigo M., Despali G., Limousin M., Angulo R., Giocoli C., Soucail G., 2015, MNRAS, 449, 3171
Buote D. A., Canizares C. R., 1994, ApJ, 427, 86
Buote D. A., Canizares C. R., 1996, ApJ, 457, 177
Buote D. A., Canizares C. R., 1998, MNRAS, 298, 811
Buote D. A., Humphrey P. J., 2012, MNRAS, 420, 1693
Cash W., 1979, ApJ, 228, 939
Chiu I.-N. T., Molnar S. M., 2012, ApJ, 756, 1
Cooray A. R., 2000, MNRAS, 313, 783
Corless V. L., King L. J., 2007, MNRAS, 380, 149
Corless V. L., King L. J., Clowe D., 2009, MNRAS, 393, 1235
Czakon N. G. et al., 2015, ApJ, 806, 18
De Filippis E., Sereno M., Bautz M. W., Longo G., 2005, ApJ, 625, 108
de Theije P. A. M., Katgert P., van Kampen E., 1995, MNRAS, 273, 30
Doré O., Bouchet F. R., Mellier Y., Teyssier R., 2001, A&A, 375, 14
Dutton A. A., Macciò A. V., 2014, MNRAS, 441, 3359
Ebeling H., Edge A. C., Henry J. P., 2001, ApJ, 553, 668
Ebeling H., Ma C. J., Kneib J.-P., Jullo E., Courtney N. J. D., Barrett E., Edge A. C., Le Borgne J.-F., 2009, MNRAS, 395, 1213
Eckert D., Ettori S., Pointecouteau E., Molendi S., Paltani S., Tchernin C., 2016, preprint (arXiv:1611.05051)
Eichner T. et al., 2013, ApJ, 774, 124
Ettori S., Morandi A., Tozzi P., Balestra I., Borgani S., Rosati P., Lovisari L., Terenziani F., 2009, A&A, 501, 61
Fasano G., Vio R., 1991, MNRAS, 249, 629
Foëx G., Soucail G., Pointecouteau E., Arnaud M., Limousin M., Pratt G. W., 2012, A&A, 546, A106
Fox D. C., Pen U.-L., 2002, ApJ, 574, 38
Fruscione A. et al., 2006, in Silva D. R., Doxsey R. E., eds, Proc. SPIE, Vol. 6270, Observatory Operations: Strategies, Processes, and Systems. SPIE, Bellingham, p. 62701V
Gavazzi R., 2005, A&A, 443, 793

Hubble E. P., 1926, ApJ, 64, 321
Jing Y. P., Suto Y., 2002, ApJ, 574, 538
Johnston D. E. et al., 2007, preprint (arXiv:0709.1159)
Kalberla P. M. W., Burton W. B., Hartmann D., Arnal E. M., Bajaja E., Morras R., Pöppel W. G. L., 2005, A&A, 440, 775
Kawahara H., 2010, ApJ, 719, 1926
Kazantzidis S., Kravtsov A. V., Zentner A. R., Allgood B., Nagai D., Moore B., 2004, ApJ, 611, L73
Komatsu E. et al., 2011, ApJS, 192, 18
Lau E. T., Kravtsov A. V., Nagai D., 2009, ApJ, 705, 1129
Lau E. T., Nagai D., Kravtsov A. V., Zentner A. R., 2011, ApJ, 734, 93
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Lee J., Suto Y., 2003, ApJ, 585, 151
Lee J., Suto Y., 2004, ApJ, 601, 599
Limousin M., Morandi A., Sereno M., Meneghetti M., Ettori S., Bartelmann M., Verdugo T., 2013, Space Sci. Rev., 177, 155
Ludlow A. D., Bose S., Angulo R. E., Wang L., Hellwing W. A., Navarro J. F., Cole S., Frenk C. S., 2016, MNRAS, 460, 1214
McCarthy I. G., Holder G. P., Babul A., Balogh M. L., 2003, ApJ, 591, 526
Mahdavi A., Chang W., 2011, ApJ, 735, L4
Mantz A. B., Allen S. W., Morris R. G., Schmidt R. W., 2016, MNRAS, 456, 4020
Martizzi D., Agrusa H., 2016, MNRAS, preprint (arXiv:1608.04388)
Mazzotta P., Rasia E., Moscardini L., Tormen G., 2004, MNRAS, 354, 10
Meneghetti M., Rasia E., Merten J., Bellagamba F., Ettori S., Mazzotta P., Dolag K., Marri S., 2010, A&A, 514, A93
Meneghetti M. et al., 2014, ApJ, 797, 34
Merten J. et al., 2015, ApJ, 806, 4
Mohr J. J., Evrard A. E., Fabricant D. G., Geller M. J., 1995, ApJ, 447, 8
Molnar S. M., Chiu I.-N., Umetsu K., Chen P., Hearn N., Broadhurst T., Bryan G., Shang C., 2010, ApJ, 724, L1
Morandi A. et al., 2012, MNRAS, 425, 2069
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., Lau E. T., Nagai D., Rudd D. H., Yu L., 2014, ApJ, 782, 107
Noerdlinger P. D., 1979, ApJ, 234, 802
Oguri M., Lee J., Suto Y., 2003, ApJ, 599, 7
Oguri M., Takada M., Umetsu K., Broadhurst T., 2005, ApJ, 632, 841
Okabe N., Takada M., Umetsu K., Futamase T., Smith G. P., 2010, PASJ, 62, 811
Orwell G., 1949, Nineteen Eighty-Four. Secker & Warburg, London
Paz D. J., Lambas D. G., Padilla N., Merchán M., 2006, MNRAS, 366, 1503
Pierre M. et al., 2016, A&A, 592, A1
Planck Collaboration XIII, 2016a, A&A, 594, A13
Planck Collaboration XXVII, 2016b, A&A, 594, A27
Planelles S., Borgani S., Dolag K., Ettori S., Fabjan D., Murante G., Tornatore L., 2013, MNRAS, 431, 1487
Plionis M., Basilakos S., Tovmassian H. M., 2004, MNRAS, 352, 1323
Postman M. et al., 2012, ApJS, 199, 25
Puchwein E., Bartelmann M., 2006, A&A, 455, 791
Rasia E. et al., 2012, New J. Phys., 14, 055018
Rasia E. et al., 2014, ApJ, 791, 96
Reblinsky K., 2000, A&A, 364, 377
Reese E. D., Kawahara H., Kitayama T., Ota N., Sasaki S., Suto Y., 2010, ApJ, 721, 653
Romanowsky A. J., Kochanek C. S., 1998, ApJ, 493, 641
Romero C. et al., 2016, ApJ, preprint (arXiv:1608.03980)
Ryden B. S., 1996, ApJ, 461, 146
Sayers J., Golwala S. R., Ameglio S., Pierpaoli E., 2011, ApJ, 728, 39
Sayers J. et al., 2016, ApJ, 832, 26
Seitz C., Schneider P., 1997, A&A, 318, 687
Sereno M., 2007, MNRAS, 380, 1207
Sereno M., Ettori S., 2016, preprint (arXiv:1603.06581)
Sereno M., Umetsu K., 2011, MNRAS, 416, 3187
Sereno M., Zitrin A., 2012, MNRAS, 419, 3280
Sereno M., De Filippis E., Longo G., Bautz M. W., 2006, ApJ, 645, 170
Sereno M., Jetzer P., Lubini M., 2010a, MNRAS, 403, 2077
Sereno M., Lubini M., Jetzer P., 2010b, A&A, 518, A55
Sereno M., Ettori S., Baldi A., 2012, MNRAS, 419, 2646
Sereno M., Ettori S., Umetsu K., Baldi A., 2013, MNRAS, 428, 2241
Shaw L. D., Nagai D., Bhattacharya S., Lau E. T., 2010, ApJ, 725, 1452
Smith R. E. et al., 2003, MNRAS, 341, 1311
Stark A. A., 1977, ApJ, 213, 368
Stock D. et al., 2015, A&A, 584, A63
Sunyaev R. A., Zeldovich Y. B., 1970, Ap&SS, 7, 3
Suto D., Kitayama T., Nishimichi T., Sasaki S., Suto Y., 2016, PASJ, 68, 97
Tchernin C. et al., 2016, A&A, 595, A42
Thakur P., Chakraborty D. K., 2001, MNRAS, 328, 330
Umetsu K. et al., 2009, ApJ, 694, 1643
Umetsu K., Broadhurst T., Zitrin A., Medezinski E., Hsu L.-Y., 2011, ApJ, 729, 127
Umetsu K. et al., 2012, ApJ, 755, 56
Umetsu K. et al., 2014, ApJ, 795, 163
Umetsu K. et al., 2015, ApJ, 806, 207
Umetsu K., Zitrin A., Gruen D., Merten J., Donahue M., Postman M., 2016, ApJ, 821, 116
Vega-Ferrero J., Yepes G., Gottlöber S., 2016, preprint (arXiv:1603.02256)
Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., 2006, ApJ, 640, 691
Voit G. M., 2005, Rev. Mod. Phys., 77, 207
West M. J., 1994, MNRAS, 268, 79
Zaroubi S., Squires G., Hoffman Y., Silk J., 1998, ApJ, 500, L87
Zitrin A. et al., 2009, MNRAS, 396, 1985
Zitrin A., Bartelmann M., Umetsu K., Oguri M., Broadhurst T., 2012a, MNRAS, 426, 2944
Zitrin A. et al., 2012b, ApJ, 749, 97
Zitrin A. et al., 2013, ApJ, 762, L30
Zitrin A. et al., 2015, ApJ, 801, 44

APPENDIX A: PROJECTION

The problem of finding the volume density distribution of systems whose surface density contours are similar ellipses is a recurrent astronomical problem. Here, we follow the formalism presented by Stark (1977), who first discussed the projection of an ideal triaxial galaxy on to the plane of the sky, and further developed by Binggeli (1980) and Binney (1985). The formalism was later introduced in the context of GL (Oguri, Lee & Suto 2003; Corless & King 2007; Sereno, Lubini & Jetzer 2010b; Sereno & Umetsu 2011) and multi-wavelengths analyses of galaxy clusters (De Filippis et al. 2005; Sereno et al. 2006, 2012, 2013; Sereno 2007; Buote & Humphrey 2012). Here, we review the main results in terms of the notation used in the present paper.

The observed system is an ellipsoid whose principal axes define the intrinsic coordinate system. The semi-major axis l_s is oriented along the third axis of the intrinsic system and the minor-to-major and intermediate-to-major axial ratios are $q_1 \leq q_2 \leq 1$. The orientation of the ellipsoid in the observer system is defined by the Euler angles ϑ, φ and ψ.

The system appears in the plane of the sky as an ellipse with projected axial ratio

$$q_\perp = \sqrt{\frac{j+l-\sqrt{(j-l)^2+4k^2}}{j+l+\sqrt{(j-l)^2+4k^2}}}, \tag{A1}$$

where j, k and l are defined as

$$j = \left(\frac{\sin\vartheta}{q_1 q_2}\right)^2 + \cos^2\vartheta\left[\left(\frac{\cos\varphi}{q_1}\right)^2 + \left(\frac{\sin\varphi}{q_2}\right)^2\right], \tag{A2}$$

$$k = \left(\frac{1}{q_1^2} - \frac{1}{q_2^2}\right)\sin\varphi\cos\varphi\cos\vartheta, \tag{A3}$$

$$l = \left(\frac{\sin\varphi}{q_1}\right)^2 + \left(\frac{\cos\varphi}{q_2}\right)^2. \quad \text{(A4)}$$

We also use the ellipticity

$$\epsilon_\perp = 1 - q_\perp. \quad \text{(A5)}$$

The (tangent of the) orientation angle in the plane of the sky of the projected ellipse (measured north over east as usual in astronomy) is

$$\tan\theta_\epsilon = \tan\left[\Delta\theta_{\psi=0} + \frac{\pi}{2}H(\Delta_{\pi/2}) - \psi\right], \quad \text{(A6)}$$

where H is the Heaviside function and $\Delta\theta_{\psi=0}$ is the angle between the projection of the major axis of the ellipsoid and the principal axes of the ellipse,

$$\Delta\theta_{\psi=0} = \frac{1}{2}\arctan\left[\frac{2k}{j-l}\right]. \quad \text{(A7)}$$

The apparent principal axis that lies furthest from the projection of the 3D major axis on to the plane of the sky is the apparent major axis if (Binney 1985)

$$\Delta_{\pi/2} = (j-l)\cos(2\Delta\theta_{\psi=0}) + 2k\sin(2\Delta\theta_{\psi=0}) \leq 0. \quad \text{(A8)}$$

The semi-major axis $l_\perp$ of the projected ellipse in the plane of the sky, i.e. perpendicularly to the line of sight, is

$$l_\perp = \frac{l_s}{e_\parallel\sqrt{f}}, \quad \text{(A9)}$$

where

$$f = \sin^2\vartheta\left[\left(\frac{\sin\varphi}{q_1}\right)^2 + \left(\frac{\cos\varphi}{q_2}\right)^2\right] + \cos^2\vartheta \quad \text{(A10)}$$

and

$$e_\parallel = \frac{q_\perp}{q_1 q_2 f^{3/4}}. \quad \text{(A11)}$$

The half-size $l_\parallel$ along the line of sight of the ellipsoid projected perpendicularly to the line of sight, i.e. as seen from above, is

$$l_\parallel = \frac{l_s}{\sqrt{f}}. \quad \text{(A12)}$$

The parameter $e_\parallel$ quantifies the extent of the cluster along the line of sight. It can be expressed as the ratio of the size of the cluster along the line of sight to the size in the plane of the sky,

$$e_\parallel = \frac{l_\parallel}{l_\perp}. \quad \text{(A13)}$$

The larger $e_\parallel$, the more the orientation bias towards the observer. If $e_\parallel > 1$, the cluster is more elongated along the line of sight than wide in the plane of the sky.

A general form for the volume density is $\rho_{3D} = \rho_s f_\rho(x)$, where ρ_s sets the density scale and x is a dimensionless variable. The functional f_ρ describes the density profile and is characterized by a number of parameters, e.g. the concentration, the outer slope, the truncation radius.

The projected surface distribution is obtained as

$$\Sigma_{2D} = \int_\parallel \rho_{3D}\,\mathrm{d}l, \quad \text{(A14)}$$

where the subscript $_\parallel$ denotes integration along the line of sight.

If the volume density is constant on similar ellipsoids, the functional f_ρ can be expressed in terms of $x = \zeta/l_s$, where ζ is the ellipsoidal radius. The projected isocontours are elliptical,

$$\Sigma_{2D} = \Sigma_s f_\Sigma(\xi/l_\perp), \quad \text{(A15)}$$

where ξ is the observed elliptical radius, the density scale is given by

$$\Sigma_s = \rho_s l_\parallel, \quad \text{(A16)}$$

and the functional form is obtained as

$$f_\Sigma(x) = 2\int_x^\infty \frac{f_\rho(x')x'}{\sqrt{x'^2 - x^2}}\mathrm{d}x', \quad \text{(A17)}$$

analogously to the spherical case.

APPENDIX B: POTENTIAL SHAPE

Even though the isopotential surfaces of ellipsoidal dark haloes are not exact ellipsoids, the isopotential surfaces are still well approximated as ellipsoids whose shape slightly varies with the radius (Lee & Suto 2003). Here, we follow Kawahara (2010) and derive the axial ratios of the potential by numerical integration. We considered the gravitational potential Φ of a triaxial NFW halo, but results are nearly independent of the peculiar matter density profile.

The isopotential surface (in the intrinsic system aligned with the axial ratios) defined by $\Phi(x, y, z) = \text{const.} = \Phi(0, 0, c_\Phi)$ is approximated by the triaxial ellipsoid with minor axis a_Φ, intermediate axis b_Φ and major axis c_Φ, such that $\Phi(a_\Phi, 0, 0) = \Phi(0, b_\Phi, 0) = \Phi(0, 0, c_\Phi)$; the axial ratios are then $q_{\Phi,1} = a_\Phi/c_\Phi$ and $q_{\Phi,2} = b_\Phi/c_\Phi$.

The shape variation is small between ζ_{500} and ζ_{vir}, the scales most interested by WL observations. In Figs B1, B2 and B3, we plot the intrinsic minor-to-major axial ratio of the potential, the projected ellipticity and the orientation angle of the projected potential, respectively, as a function of the radius. Radial variations are small even for very elongated clusters ($q_{mat,1} \lesssim 0.3$), and far below the usual observational accuracy between ζ_{500} and ζ_{vir}. The potential and the matter distribution have nearly the same triaxiality parameter and $e_{\Phi,i}/e_{mat,i} \sim 0.7$.

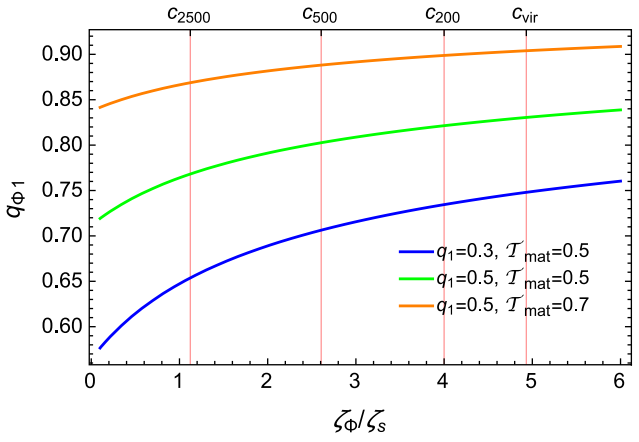

Figure B1. Minor-to-major axial ratio of the isopotential surfaces of an NFW halo with $c_{200} = 4$ at $z = 0.3$ as a function of the ellipsoidal radius.

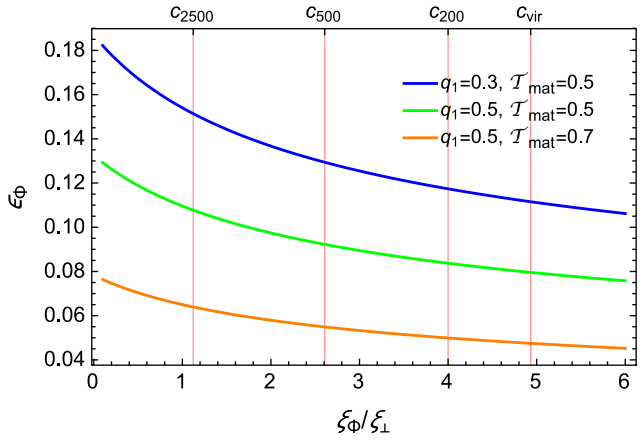

Figure B2. Projected ellipticity of the isopotential surfaces of an NFW halo with $c_{200} = 4$ as a function of the elliptical radius. The orientation of the halo is fixed by $\cos\vartheta = 0.5$ and $\phi = \pi/3$.

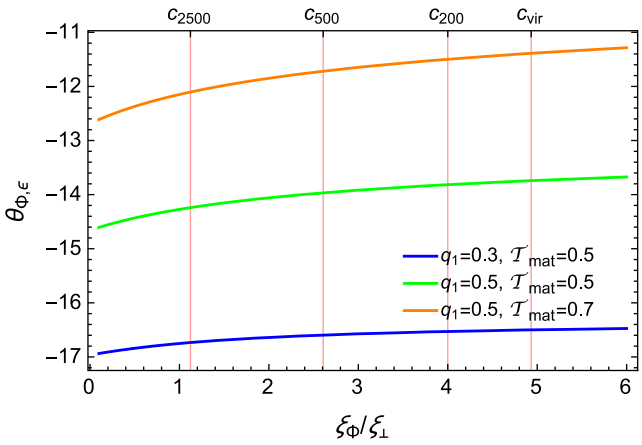

Figure B3. Orientation angle of the isopotential surfaces of an NFW halo with $c_{200} = 4$ as a function of the elliptical radius. The orientation of the halo is fixed by $\cos\vartheta = 0.5$, $\phi = \pi/3$ and $\psi = 0$.

This paper has been typeset from a TEX/LATEX file prepared by the author.