The Astrophysical Journal, 806:207 (27pp), 2015 June 20 doi:10.1088/0004-637X/806/2/207

THREE-DIMENSIONAL MULTI-PROBE ANALYSIS OF THE GALAXY CLUSTER A1689*

Keiichi Umetsu[1], Mauro Sereno[2,3], Elinor Medezinski[4,5], Mario Nonino[6], Tony Mroczkowski[7], Jose M. Diego[8], Stefano Ettori[3,9], Nobuhiro Okabe[10,11], Tom Broadhurst[12], and Doron Lemze[5]
[1] Institute of Astronomy and Astrophysics, Academia Sinica, P. O. Box 23-141, Taipei 10617, Taiwan; keiichi@asiaa.sinica.edu.tw
[2] Dipartimento di Fisica e Astronomia, Alma Mater Studiorum—Università di Bologna Viale Berti Pichat 6/2, I-40127 Bologna, Italy
[3] INAF-Osservatorio Astronomico di Bologna, via Ranzani 1, I-40127 Bologna, Italy
[4] The Racah Institute of Physics, The Hebrew University of Jerusalem, Jerusalem 91904, Israel
[5] Department of Physics and Astronomy, The Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218, USA
[6] INAF-Osservatorio Astronomico di Trieste, via G.B. Tiepolo 11, I-34143 Trieste, Italy
[7] National Research Council Fellow at the U.S. Naval Research Laboratory, 4555 Overlook Ave SW, Washington, DC 20375, USA
[8] IFCA, Instituto de Física de Cantabria (UC-CSIC), Av. de Los Castros s/n, 39005 E-Santander, Spain
[9] INFN, Sezione di Bologna, Viale Berti Pichat 6/2, I-40127 Bologna, Italy
[10] Department of Physical Science, Hiroshima University, 1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526, Japan
[11] Kavli Institute for the Physics and Mathematics of the Universe (WPI), Todai Institutes for Advanced Study, University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8583, Japan
[12] Ikerbasque, Basque Foundation for Science, Alameda Urquijo, 36-5 Plaza Bizkaia, E-48011 Bilbao, Spain
Received 2015 March 7; accepted 2015 May 14; published 2015 June 18

ABSTRACT

We perform a three-dimensional multi-probe analysis of the rich galaxy cluster A1689, one of the most powerful known lenses on the sky, by combining improved weak-lensing data from new wide-field $BVR_Ci'z'$ Subaru/Suprime-Cam observations with strong-lensing, X-ray, and Sunyaev–Zel'dovich effect (SZE) data sets. We reconstruct the projected matter distribution from a joint weak-lensing analysis of two-dimensional shear and azimuthally integrated magnification constraints, the combination of which allows us to break the mass-sheet degeneracy. The resulting mass distribution reveals elongation with an axis ratio of ∼0.7 in projection, aligned well with the distributions of cluster galaxies and intracluster gas. When assuming a spherical halo, our full weak-lensing analysis yields a projected halo concentration of $c_{200c}^{2D} = 8.9 \pm 1.1$ ($c_{vir}^{2D} \sim 11$), consistent with and improved from earlier weak-lensing work. We find excellent consistency between independent weak and strong lensing in the region of overlap. In a parametric triaxial framework, we constrain the intrinsic structure and geometry of the matter and gas distributions, by combining weak/strong lensing and X-ray/SZE data with minimal geometric assumptions. We show that the data favor a triaxial geometry with minor–major axis ratio 0.39 ± 0.15 and major axis closely aligned with the line of sight ($22^\circ \pm 10^\circ$). We obtain a halo mass $M_{200c} = (1.2 \pm 0.2) \times 10^{15}\, M_\odot\, h^{-1}$ and a halo concentration $c_{200c} = 8.4 \pm 1.3$, which overlaps with the $\gtrsim 1\sigma$ tail of the predicted distribution. The shape of the gas is rounder than the underlying matter but quite elongated with minor–major axis ratio 0.60 ± 0.14. The gas mass fraction within 0.9 Mpc is 10^{+3}_{-2}%, a typical value for high-mass clusters. The thermal gas pressure contributes to ∼60% of the equilibrium pressure, indicating a significant level of non-thermal pressure support. When compared to *Planck*'s hydrostatic mass estimate, our lensing measurements yield a spherical mass ratio of $M_{Planck}/M_{GL} = 0.70 \pm 0.15$ and 0.58 ± 0.10 with and without corrections for lensing projection effects, respectively.

Key words: cosmology: observations – dark matter – galaxies: clusters: individual (A1689) – gravitational lensing: strong – gravitational lensing: weak

1. INTRODUCTION

The evolution of the abundance of galaxy clusters with cosmic epoch is sensitive to the amplitude and growth rate of primordial density fluctuations as well as to the cosmic volume-redshift relation because massive clusters lie in the high-mass exponential tail of the halo mass function (Haiman et al. 2001; Watson et al. 2014). Therefore, large cluster samples defined from cosmological surveys can provide an independent means of examining any viable cosmological model, including the current concordance Λ cold dark matter (ΛCDM) model defined in the framework of general relativity, complementing cosmic microwave background (CMB), large-scale galaxy clustering, and supernova observations.

Clusters provide various probes of the role and nature of "dark matter" (DM) that dominates the material universe (Clowe et al. 2006), or modified gravity theories as an alternative to DM (Rapetti et al. 2010), physics governing the final state of self-gravitating collisionless systems in an expanding universe (Navarro et al. 1996, 1997; Taylor & Navarro 2001; Hjorth & Williams 2010), and screening mechanisms in long-range modified models of gravity whereby general relativity is restored (Narikawa et al. 2013).

Substantial progress has been made in recent years in constructing statistical samples of clusters thanks to dedicated surveys (e.g., Planck Collaboration et al. 2014, 2015c; Bleem et al. 2015). Cluster samples are often defined by X-ray or Sunyaev–Zel'dovich effect (SZE) observables, so that the masses are indirectly inferred from scaling relations, which are often based on the assumption of hydrostatic equilibrium (HSE) and then statistically calibrated using weak lensing or internal dynamics using a subset of massive clusters at lower redshifts (Rines et al. 2013; Gruen et al. 2014). Since the level of mass bias from indirect observations assuming HSE is likely

* Based in part on data collected at the Subaru Telescope, which is operated by the National Astronomical Society of Japan.

mass dependent (Sereno et al. 2014a) and sensitive to calibration systematics of the instruments (Donahue et al. 2014; Israel et al. 2015), a systematic effort is needed to enable a self-consistent calibration of mass–observable relations using robust, direct cluster mass measurements (Ford et al. 2014; Jimeno et al. 2014; Merten et al. 2014; Umetsu et al. 2014; von der Linden et al. 2014; Hoekstra et al. 2015; Simet et al. 2015) and well-defined selection functions (e.g., Benitez et al. 2014).

The great attraction of gravitational lensing in the cluster regime is its ability to map the mass distribution on an individual cluster basis, independent of and free from assumptions about the physical and dynamical state of the cluster system (Miyazaki et al. 2007; Okabe & Umetsu 2008; Hamana et al. 2009). Clusters act as efficient gravitational lenses, producing various observable effects, including deflection, distortion, and magnification of the images of background sources (Bartelmann & Schneider 2001). In the weak regime, the lensing signals are approximately linearly related to the gravitational potential, so that one can determine the distribution of lensing matter at large scales in a model-independent manner (e.g., Umetsu et al. 1999, 2011b). In the strong regime, several sets of multiply lensed images with known redshifts can be used to constrain the mass distribution in the cluster cores (e.g., Jauzac et al. 2014; Zitrin et al. 2014).

A practical difficulty of obtaining precise mass measurements from cluster lensing, however, is significant scatter present in the projected lensing signals due to inherent variations (at a fixed halo mass) in halo concentration, asphericity, orientation, and the presence of correlated large-scale structure (Rasia et al. 2012). The projection effects due to such intrinsic profile variations alone can produce a $\lesssim$20% uncertainty in lensing mass estimates for $\sim 10^{15}\,M_\odot$ clusters (Becker & Kravtsov 2011; Gruen et al. 2015).

A possible way to overcome this problem is to simultaneously determine the mass, concentration, shape, and orientation of a given cluster by combining lensing data with independent probes or information about its line of sight elongation (Sereno 2007; Corless et al. 2009; Limousin et al. 2013). Gravitational lensing probes the structure and morphology of the matter distribution in projection. X-ray observations constrain the characteristic size and orientation of the intracluster medium (ICM) in the sky plane. The elongation of the ICM along the line of sight can be constrained from the combination of X-ray and thermal SZE observations (De Filippis et al. 2005; Sereno et al. 2012). Recently, Sereno et al. (2013) developed a parametric triaxial framework to combine and couple independent morphological constraints from lensing and X-ray/SZE data, using minimal geometric assumptions about the matter and gas distributions but without assuming HSE.

The first critical step in a three-dimensional (3D) cluster analysis is an unbiased, direct recovery of the projected cluster mass distribution from weak lensing. A fundamental limitation of measuring shear only is the *mass-sheet* degeneracy (Schneider & Seitz 1995). This degeneracy can be broken by using the complementary combination of shear and magnification (Schneider et al. 2000; Umetsu et al. 2011b; Umetsu 2013). Umetsu et al. (2011b) have shown that the magnification effect can significantly enhance the accuracy and precision of lensing-derived cluster mass profiles when added to weak-lensing shear measurements.

Our aim in this paper is to develop and apply a comprehensive set of techniques and methods for 3D analysis of galaxy clusters based on the multi-probe framework of Sereno et al. (2013). To this end, we first generalize the one-dimensional (1D) weak-lensing inversion method of Umetsu et al. (2011b) to a two-dimensional (2D) description of the mass distribution without assuming particular functional forms, i.e., in a *free-form* fashion. In this approach, we combine the spatial shear pattern with azimuthally averaged magnification information, imposing integrated constraints on the mass distribution.

Taking advantage of new $BVR_Ci'z'$ imaging obtained with Suprime-Cam on the 8.3 m Subaru Telescope, we perform a new weak-lensing analysis of the rich cluster A1689 at $z = 0.183$ and then apply our methods to weak-lensing, strong-lensing, X-ray, and SZE data sets we have obtained for the cluster. The cluster is among the best studied clusters (Tyson & Fischer 1995; Taylor et al. 1998; Andersson & Madejski 2004; Broadhurst et al. 2005b; Halkola et al. 2006; Limousin et al. 2007; Umetsu & Broadhurst 2008; Peng et al. 2009; Coe et al. 2010; Kawaharada et al. 2010; Sereno et al. 2012, 2013; Nieuwenhuizen & Morandi 2013) and one of the most powerful known lenses on the sky, characterized by a large Einstein radius of $\theta_{\rm Ein} = 47\farcs0 \pm 1\farcs2$ for a fiducial source at $z_{\rm s} = 2$ (see Table 1; Coe et al. 2010); this indicates a high degree of mass concentration in projection (Broadhurst & Barkana 2008). To date, 61 candidate systems of 165 multiply lensed images have been identified (Broadhurst et al. 2005b; Coe et al. 2010; Diego et al. 2015) from Advanced Camera for Surveys (ACS) observations with the *Hubble Space Telescope* (*HST*). Despite significant efforts, the degree of concentration inferred from different lensing analyses is somewhat controversial (see Coe et al. 2010; Sereno et al. 2013), and it is still unclear if and to what degree this cluster is over-concentrated.

The paper is organized as follows. After summarizing the basic theory of cluster weak lensing, we present in Section 2 the formalism that we use for our weak-lensing analysis. In Section 3, we describe our Subaru observations and data processing. In Section 4, we present our Subaru weak-lensing analysis. Section 5 presents our *HST* strong-lensing analysis. In Section 6 we outline the triaxial modeling and describe the statistical framework for the 3D cluster analysis. In Section 7 we present the multi-probe analysis of lensing and X-ray/SZE data. In Section 8 we discuss the results and their implications for the intrinsic properties of A1689. Finally, a summary of our work is given in Section 9.

Throughout this paper, we use the AB magnitude system and adopt a concordance ΛCDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $h \equiv 0.7 h_{70} = 0.7$ where $H_0 = h \times 100\,{\rm km\,s^{-1}\,Mpc^{-1}}$. In this cosmology, $1'$ corresponds to $129\,{\rm kpc}\,h^{-1} \simeq 185\,{\rm kpc}\,h_{70}^{-1}$ for this cluster. The reference sky position is the center of the brightest cluster galaxy (BCG): R.A.(J2000.0) = 13 : 11 : 29.52, decl.(J2000.0) = −01 : 20 : 27.59 (Table 1). We use the standard notation r_Δ to denote the spherical overdensity radius within which the mean interior density is Δ times the critical density $\rho_{\rm c}$ of the universe at the cluster redshift. For its ellipsoidal counterpart R_Δ, see Section 6.1. All quoted errors are 68.3% (1σ) confidence limits (CL) unless otherwise stated.

Table 1
Properties of the Galaxy Cluster A1689

Parameter	Value
ID	A1689
Optical center position (J2000.0)	
R.A.	13:11:29.52
Decl.	–01:20:27.59
X-ray center position (J2000.0)	
R.A.	13:11.29.50
Decl.	–01:20:29.92
SZE center position (J2000.0)	
R.A.	13:11.29.57
Decl.	–01:20:29.87
Redshift	0.183
X-ray temperature (keV)	10.4
Einstein radius ($''$)	47.0 ± 1.2 at $z_s = 2$

Notes. The optical cluster center is defined as the center of the BCG from Reference [2]. Units of right ascension are hours, minutes, and seconds, and units of declination are degrees, arcminutes, and arcseconds. The X-ray properties were taken from Reference [3]. The X-ray center is defined as the X-ray emission centroid derived from *XMM-Newton* observations. See also Reference [1]. The SZE center is determined from the joint analysis of interferometric BIMA/OVRO/SZA observations described in Section 7.2. The BCG is located within $2''.3$ ($\simeq 5$ kpc h^{-1}) of the X-ray center. The X-ray and SZE centroid positions agree to within $1''$. The Einstein radius is constrained by detailed strong lens modeling by Reference [4].
References. [1] Andersson & Madejski (2004); [2] Limousin et al. (2007), [3] Kawaharada et al. (2010), [4] Coe et al. (2010).

2. WEAK-LENSING METHODOLOGY

2.1. Weak Lensing Basics

In the cluster regime, the lensing convergence, $\kappa = \Sigma/\Sigma_c$, is the projected mass density $\Sigma(\boldsymbol{\theta})$ in units of the critical surface density for lensing, $\Sigma_c = (c^2 D_s)/(4\pi G D_l D_{ls}) \equiv c^2/(4\pi G D_l \beta)$ with D_l, D_s, and D_{ls} the lens, source, and lens-source angular diameter distances, respectively; $\beta(z) = D_{ls}(z)/D_s(z)$ represents the geometric lensing strength for a source at redshift z, where $\beta(z) = 0$ for $z \leqslant z_l$.

The gravitational shear $\gamma = \gamma_1 + i\gamma_2$ can be directly observed from ellipticities of background galaxies in the weak regime, $\kappa \ll 1$. The shear and convergence are related by

$$\gamma(\boldsymbol{\theta}) = \int d^2\theta'\, D(\boldsymbol{\theta} - \boldsymbol{\theta}')\kappa(\boldsymbol{\theta}') \tag{1}$$

with $D(\boldsymbol{\theta}) = (\theta_2^2 - \theta_1^2 - 2i\theta_1\theta_2)/(\pi\,|\boldsymbol{\theta}|^4)$ (Kaiser & Squires 1993). The observable quantity for quadrupole weak lensing in general is not γ but the complex reduced shear,

$$g(\boldsymbol{\theta}) = \frac{\gamma(\boldsymbol{\theta})}{1 - \kappa(\boldsymbol{\theta})}. \tag{2}$$

The g field is invariant under $\kappa(\boldsymbol{\theta}) \to \lambda\kappa(\boldsymbol{\theta}) + 1 - \lambda$ and $\gamma(\boldsymbol{\theta}) \to \lambda\gamma(\boldsymbol{\theta})$ with an arbitrary constant $\lambda \neq 0$, known as the mass-sheet degeneracy (Schneider & Seitz 1995). This degeneracy can be broken, for example, by measuring the magnification $\mu(\boldsymbol{\theta})$ in the subcritical regime,

$$\mu(\boldsymbol{\theta}) = \frac{1}{[1 - \kappa(\boldsymbol{\theta})]^2 - |\gamma(\boldsymbol{\theta})|^2} \equiv \frac{1}{\Delta(\boldsymbol{\theta})}, \tag{3}$$

which transforms as $\mu(\boldsymbol{\theta}) \to \lambda^2\mu(\boldsymbol{\theta})$.

Let us consider a population of source galaxies described by their redshift distribution function, $\overline{N}(z)$. In general, we apply different size, magnitude, and color cuts in background selection for measuring shear and magnification, which results in different $\overline{N}(z)$. In contrast to the former effect, the latter does not require source galaxies to be spatially resolved, but it requires a stringent flux limit against incompleteness effects. The mean lensing depth for a given population ($X = g$, μ) is

$$\langle\beta\rangle_X = \left[\int_0^\infty dz\, w(z)\overline{N}_X(z)\beta(z)\right]\left[\int_0^\infty dz\, w(z)\overline{N}_X(z)\right]^{-1}, \tag{4}$$

where $w(z)$ is a weight factor (see Section 3.3).

We introduce the relative lensing strength of a given source population relative to a fiducial source in the far background as $\langle W\rangle_X = \langle\beta\rangle_X/\beta_\infty$ (Bartelmann & Schneider 2001) with $\beta_\infty \equiv \beta(z \to \infty; z_l)$. The associated critical density is $\Sigma_{c,\infty}(z_l) = c^2/(4\pi G D_l \beta_\infty)$. Hereafter, we use the far-background fields $\kappa_\infty(\boldsymbol{\theta})$ and $\gamma_\infty(\boldsymbol{\theta})$ to describe the projected cluster mass distribution.

2.2. Discretized Mass Distribution

We discretize the convergence field $\kappa_\infty(\boldsymbol{\theta}) = \Sigma_{c,\infty}^{-1}\Sigma(\boldsymbol{\theta})$ into a regular grid of pixels and approximate $\kappa_\infty(\boldsymbol{\theta})$ by a linear combination of basis functions $B(\boldsymbol{\theta} - \boldsymbol{\theta}')$ as

$$\kappa_\infty(\boldsymbol{\theta}) = \Sigma_{c,\infty}^{-1}\sum_{m=1}^{N_{pix}} B\big(\boldsymbol{\theta} - \boldsymbol{\theta}_m\big)\,\Sigma_m, \tag{5}$$

where our model (signal) $\boldsymbol{s} = \{\Sigma_m\}_{m=1}^{N_{pix}}$ is a vector of parameters containing mass coefficients. To avoid the loss of information due to oversmoothing, we take the basis function to be the Dirac delta function $B(\boldsymbol{\theta} - \boldsymbol{\theta}_m) = (\Delta\theta)^2\delta_D^2(\boldsymbol{\theta} - \boldsymbol{\theta}_m)$ with $\Delta\theta$ a constant spacing, so that $\boldsymbol{s}$ represents the cell-averaged projected mass density. The $\gamma_\infty(\boldsymbol{\theta})$ field can be expressed as

$$\gamma_\infty(\boldsymbol{\theta}) = \Sigma_{c,\infty}^{-1}\sum_{m=1}^{N_{pix}} \mathcal{D}\big(\boldsymbol{\theta} - \boldsymbol{\theta}_m\big)\Sigma_m \tag{6}$$

with $\mathcal{D} \equiv D \otimes B$ an effective kernel (Equation (1)). Hence, both κ_∞ and γ_∞ can be written as linear combinations of $\boldsymbol{s}$.

Because of the choice of the basis function, an unbiased extraction of mass coefficients $\{\Sigma_m\}_{m=1}^{N_{pix}}$ (or certain linear combinations of Σ_m) can be done by performing a spatial integral of Equation (5) over a certain area. In practical applications, such operations include smoothing (Figure 1), azimuthal averaging for a mass profile reconstruction (Section 5.3), and profile fitting with smooth functions (Section 7).

2.3. Weak-lensing Observables

2.3.1. Reduced Shear

The quadrupole image distortion due to lensing is described by the reduced shear, $g = g_1 + ig_2$. We calculate the weighted average $g_m \equiv g(\boldsymbol{\theta}_m)$ of individual shear estimates on a regular cartesian grid ($m = 1, 2, \ldots, N_{pix}$) as

$$g_m = \left[\sum_k S\big(\boldsymbol{\theta}_{(k)}, \boldsymbol{\theta}_m\big)w_{(k)}g_{(k)}\right]\left[\sum_k S\big(\boldsymbol{\theta}_{(k)}, \boldsymbol{\theta}_m\big)w_{(k)}\right]^{-1} \tag{7}$$

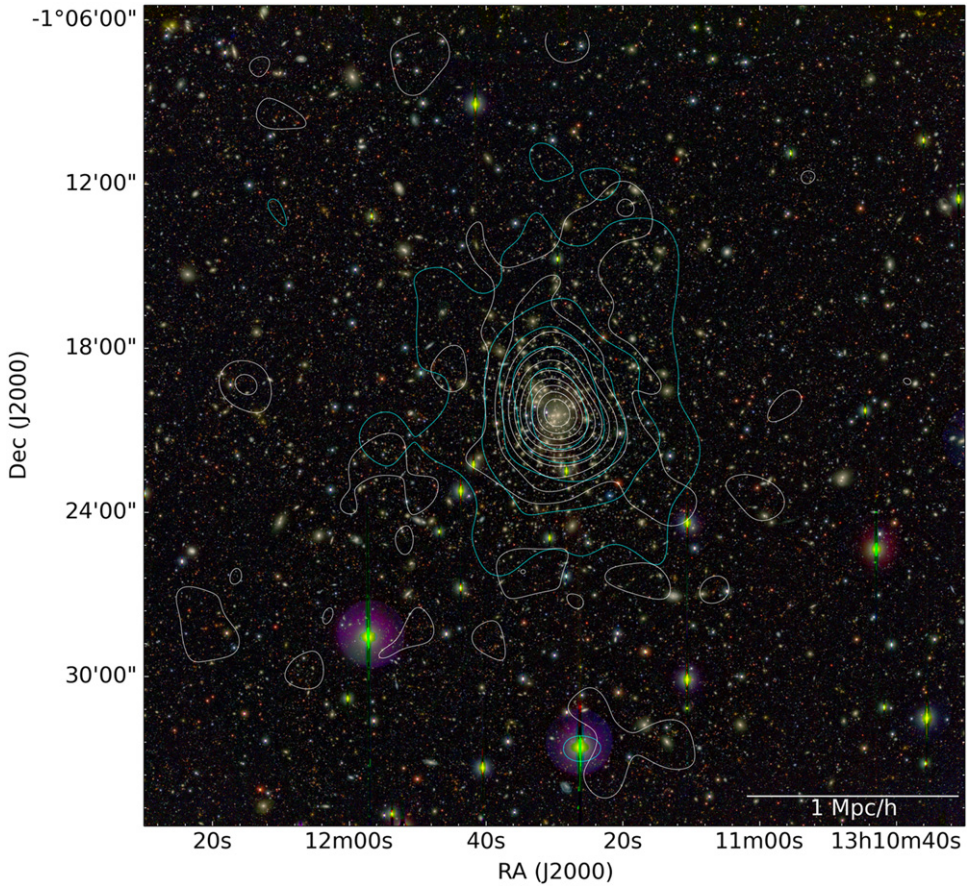

Figure 1. Subaru $BVR_Ci'z'$ composite color image centered on the galaxy cluster A1689 ($z = 0.183$), overlaid with mass contours from our joint shear-and-magnification weak-lensing analysis of Subaru data. The image is $30' \times 30'$ in size. The mass map is smoothed with a Gaussian of FWHM $= 1.'5$. The horizontal bar represents 1 Mpc h^{-1} at the cluster redshift. The lowest contour level and the contour interval are $\Delta\kappa = 0.06$. The cyan contours show the smoothed projected distribution of cluster red-sequence galaxies. North is up and east is to the left.

where $S(\boldsymbol{\theta}_{(k)}, \boldsymbol{\theta}_m)$ is a spatial window function, $g_{(k)}$ is an estimate of $g(\boldsymbol{\theta})$ for the kth object at $\boldsymbol{\theta}_{(k)}$, and $w_{(k)}$ is its statistical weight given by $w_{(k)} = 1/(\sigma^2_{g(k)} + \alpha_g^2)$ with $\sigma^2_{g(k)}$ the error variance of $g_{(k)}$ and α_g^2 the softening constant variance. We choose $\alpha_g = 0.4$, a typical value of the mean rms $\sqrt{\overline{\sigma_g^2}}$ found in Subaru observations (e.g., Umetsu et al. 2009).

The source-averaged theoretical expectation for the estimator (7) is approximated by (see Appendix A.1)

$$\hat{g}(\boldsymbol{\theta}_m) = \frac{\langle W\rangle_g \gamma_\infty(\boldsymbol{\theta}_m)}{1 - f_{W,g}\langle W\rangle_g \kappa_\infty(\boldsymbol{\theta}_m)}, \tag{8}$$

where $\langle W\rangle_g$ is the source-averaged relative lensing strength (Section 2.1), and $f_{W,g} = \langle W^2\rangle_g / \langle W\rangle_g^2$ is a dimensionless quantity of the order unity. The variance $\sigma^2_{g,m} \equiv \sigma^2_g(\boldsymbol{\theta}_m)$ for $g_m = g_{1,m} + ig_{2,m}$ is expressed as

$$\sigma^2_{g,m} = \left[\sum_k S^2(\boldsymbol{\theta}_{(k)}, \boldsymbol{\theta}_m) w^2_{(k)} \sigma^2_{g(k)}\right] \left[\sum_k S^2(\boldsymbol{\theta}_{(k)}, \boldsymbol{\theta}_m) w_{(k)}\right]^{-2}. \tag{9}$$

In this work, we adopt the top-hat window of radius $\theta_{\rm f}$ (Merten et al. 2009), $S(\boldsymbol{\theta}, \boldsymbol{\theta}') = H(\theta_{\rm f} - |\boldsymbol{\theta} - \boldsymbol{\theta}'|)$, with $H(x)$ the Heaviside function defined such that $H(x) = 1$ if $x \geqslant 0$ and $H(x) = 0$ otherwise. The covariance matrix for g_m is

$$\mathrm{Cov}\left(g_{\alpha,m}, g_{\beta,n}\right) \equiv \delta_{\alpha\beta}\left(C_g\right)_{mn} = \frac{\delta_{\alpha\beta}}{2}\sigma_{g,m}\sigma_{g,n}\xi_H\left(|\boldsymbol{\theta}_m - \boldsymbol{\theta}_n|\right), \tag{10}$$

where $\xi_H(x;\, \theta_{\rm f})$ is the autocorrelation of a pillbox of radius $\theta_{\rm f}$ (White et al. 1999; Park et al. 2003), given by

$$\xi_H(x) = \frac{2}{\pi}\left[\cos^{-1}\left(\frac{x}{2\theta_{\rm f}}\right) - \left(\frac{x}{2\theta_{\rm f}}\right)\sqrt{1 - \left(\frac{x}{2\theta_{\rm f}}\right)^2}\right] \tag{11}$$

for $|x| \leqslant 2\theta_{\rm f}$ and $\xi_H(x) = 0$ for $|x| > 2\theta_{\rm f}$.

2.3.2. Magnification Bias

Deep multi-band photometry allows us to explore the faint end of the luminosity function of red quiescent galaxies at $z \sim 1$ (Ilbert et al. 2010), for which the effect of magnification bias is dominated by the geometric area distortion, and thus is not sensitive to the exact form of the source luminosity function. In this work, we perform magnification measurements using a flux-limited sample of red background galaxies.

If the magnitude shift $\delta m = 2.5 \log_{10} \mu$ due to magnification is small compared to that on which the logarithmic slope of the luminosity function varies, their number counts can be locally approximated by a power law at the limiting flux (Broadhurst

et al. 1995). The expectation value for the source counts $N_\mu(\boldsymbol{\theta}_m)$ on a grid of equal-area cells ($m = 1, 2, \ldots$) is modified by lensing magnification as (see Appendix A.2)

$$E\left[N_\mu(\boldsymbol{\theta}_m)\right] = \overline{N}_\mu \Delta^{1-\alpha}(\boldsymbol{\theta}_m),$$
$$\Delta(\boldsymbol{\theta}) = \left[1 - \langle W\rangle_\mu \kappa_\infty(\boldsymbol{\theta})\right]^2 - \langle W\rangle_\mu^2 |\gamma_\infty(\boldsymbol{\theta})|^2, \quad (12)$$

where $\overline{N}_\mu$ is the unlensed mean source counts per cell, α is the unlensed count slope evaluated at the flux limit F, $\alpha = -d\log\overline{N}_\mu(>F)/d\log F$,[13] and $\langle W\rangle_\mu$ is the source-averaged relative lensing strength (Section 2.1).

The net magnification effect on the source counts vanishes when $\alpha = 1$. In the regime where $\alpha \ll 1$, the bias is dominated by the expansion of the sky area, producing a net count depletion. For a population with $\alpha > 1$, the bias is positive and a net density enhancement results (e.g., Hildebrandt et al. 2011; Ford et al. 2012, 2014). The faint blue population lying at $z \sim 2$ (e.g., Lilly et al. 2007; Medezinski et al. 2010, 2013) tends to have a steep intrinsic slope close to the lensing-invariant one, $\alpha = 1$.

The covariance matrix of $N_\mu(\boldsymbol{\theta})$ includes both sample covariance and Poisson variance (Hu & Kravtsov 2003):

$$\mathrm{Cov}\left[N_\mu(\boldsymbol{\theta}_m), N_\mu(\boldsymbol{\theta}_n)\right] \equiv \left(C_N\right)_{mn} = \left(\overline{N}_\mu\right)^2 \omega_{mn} + \delta_{mn} N_\mu(\boldsymbol{\theta}_m), \quad (13)$$

where ω_{mn} is the cell-averaged angular correlation function

$$\omega_{mn} = \frac{1}{\Omega_{\rm cell}^2}\int d^2\theta\, d^2\theta'\, S_m(\boldsymbol{\theta}) S_n(\boldsymbol{\theta}')\omega(\boldsymbol{\theta} - \boldsymbol{\theta}') \quad (14)$$

with $\omega(\boldsymbol{\theta})$ the angular two-point correlation function of the source galaxies, $S_m(\boldsymbol{\theta})$ the boxcar window function of the mth cell, and $\Omega_{\rm cell} = \int d^2\theta\, S_m(\boldsymbol{\theta})$. For deep lensing observations, the angular correlation length of background galaxies can be small (e.g., Connolly et al. 1998) compared to the typical resolution $\sim 1'$ of reconstructed mass maps. Therefore, the correlation between different cells can be generally ignored, whereas the unresolved correlation on small angular scales accounts for increase of the variance of $N_\mu(\boldsymbol{\theta})$ (Van Waerbeke et al. 2000). We thus approximate C_N by

$$\left(C_N\right)_{mn} \approx \left[\left\langle \delta N_\mu^2(\boldsymbol{\theta}_m)\right\rangle + N_\mu(\boldsymbol{\theta}_m)\right]\delta_{mn}, \quad (15)$$

with $\langle \delta N_\mu^2(\boldsymbol{\theta}_m)\rangle$ the variance of the mth counts.

To enhance the signal-to-noise ratio (S/N), we azimuthally average $N_\mu(\boldsymbol{\theta})$ in contiguous, concentric annuli and calculate the surface number density $\{n_{\mu,i}\}_{i=1}^{N_{\rm bin}}$ of background galaxies as a function of clustercentric radius:

$$n_{\mu,i} = \frac{\eta_i}{\Omega_{\rm cell}}\sum_m \mathcal{P}_{im} N_\mu(\boldsymbol{\theta}_m) \quad (16)$$

with $\mathcal{P}_{im} = (\sum_m A_{mi})^{-1} A_{mi}$ the radial projection matrix normalized 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 ($0 \leqslant A_{mi} \leqslant 1$), and $\eta_i (\geqslant 1)$ is the mask correction factor for the ith annular bin, $\eta_i = \left[\sum_m (1 - f_m) A_{mi}\right]^{-1}\sum_m A_{mi}$, with f_m the fraction of the mask area in the mth cell, due to bad pixels, saturated objects, foreground and cluster member galaxies (see Section 3.2 of Umetsu et al. 2014).

The theoretical expectation for the estimator (16) is

$$\hat{n}_{\mu,i} = \bar{n}_\mu \sum_m \mathcal{P}_{im}\, \Delta^{1-\alpha}(\boldsymbol{\theta}_m) \quad (17)$$

with $\bar{n}_\mu = \overline{N}_\mu/\Omega_{\rm cell}$. The bin-to-bin covariance matrix for the estimator (16) is obtained as

$$\mathrm{Cov}(n_{\mu,i}, n_{\mu,j}) \equiv \left(C_\mu\right)_{ij} = \frac{\eta_i\eta_j}{\Omega_{\rm cell}^2}\sum_{m,n}\mathcal{P}_{im}\mathcal{P}_{jn}\left(C_N\right)_{mn}. \quad (18)$$

Note that since C_N is diagonal, C_μ is also diagonal:

$$\left(C_\mu\right)_{ij} \equiv \sigma_{\mu,i}^2\delta_{ij}. \quad (19)$$

2.4. Mass Reconstruction

Given a model $\boldsymbol{m}$ and observed (fixed) data $\boldsymbol{d}$, the posterior probability $P(\boldsymbol{m}|\boldsymbol{d})$ is proportional to the product of the likelihood $\mathcal{L}(\boldsymbol{m}) \equiv P(\boldsymbol{d}|\boldsymbol{m})$ and the prior probability $P(\boldsymbol{m})$. In our 2D inversion problem, $\boldsymbol{m}$ is a vector containing the signal parameters $\boldsymbol{s}$ (Section 2.2) and calibration parameters $\boldsymbol{c}$ (Section 2.4.3), $\boldsymbol{m} \equiv (\boldsymbol{s}, \boldsymbol{c})$.

The total likelihood function $\mathcal{L}$ for combined weak-lensing data $\boldsymbol{d}$ is given as a product of the two separate likelihoods, $\mathcal{L} = \mathcal{L}_g\mathcal{L}_\mu$, where $\mathcal{L}_g$ and $\mathcal{L}_\mu$ are the likelihood functions for shear and magnification, respectively. We assume that the errors on the data follow a Gaussian distribution, so that $\mathcal{L} \propto \exp(-\chi^2/2)$, with χ^2 the standard misfit statistic.

2.4.1. Shear Log-likelihood Function

The log-likelihood function $l_g \equiv -\ln\mathcal{L}_g$ for 2D shear data can be written in the general form (ignoring constant terms) as (Oguri et al. 2010; Umetsu et al. 2012)

$$l_g = \frac{1}{2}\sum_{m,n=1}^{N_{\rm pix}}\sum_{\alpha=1}^{2}\left[g_{\alpha,m} - \hat{g}_{\alpha,m}(\boldsymbol{m})\right]\left(\mathcal{W}_g\right)_{mn}\left[g_{\alpha,n} - \hat{g}_{\alpha,n}(\boldsymbol{m})\right] \quad (20)$$

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

$$\left(\mathcal{W}_g\right)_{mn} = M_m M_n\left(C_g^{-1}\right)_{mn}, \quad (21)$$

with $(C_g^{-1})_{mn}$ the inverse covariance matrix for the 2D shear data and M_m a mask weight, defined such that $M_m = 0$ if the mth cell is masked out and $M_m = 1$ otherwise.

2.4.2. Magnification Log-likelihood Function

Similarly, the log-likelihood function for magnification-bias data $l_\mu \equiv -\ln\mathcal{L}_\mu$ can be written as

$$l_\mu = \frac{1}{2}\sum_{i=1}^{N_{\rm bin}}\left[n_{\mu,i} - \hat{n}_{\mu,i}(\boldsymbol{m})\right]\left(\mathcal{W}_\mu\right)_{ij}\left[n_{\mu,j} - \hat{n}_{\mu,j}(\boldsymbol{m})\right], \quad (22)$$

where $\hat{n}_{\mu,i}(\boldsymbol{m})$ is the theoretical prediction for the observed counts $n_{\mu,i}$ (see Equations (16) and (17)), and $(\mathcal{W}_\mu)_{ij}$ is the

[13] In the weak-lensing literature, $s \equiv d\log_{10} N(<m)/dm = 0.4\alpha$ in terms of the limiting magnitude m is often used instead of α (e.g., Umetsu et al. 2011b, 2014; Medezinski et al. 2013).

magnification weight matrix,

$$\left(\mathcal{W}_\mu\right)_{ij} = \left(C_\mu^{-1}\right)_{ij} = \frac{\delta_{ij}}{\sigma^2_{\mu,i}} \tag{23}$$

(Equations (18) and (19)). We use Monte Carlo integration to calculate the radial projection matrix $\mathcal{P}_{im}$ (Equation (16)) of size $N_{\rm bin} \times N_{\rm pix}$, which is needed to predict $\{\hat{n}_{\mu,i}(\boldsymbol{m})\}_{i=1}^{N_{\rm bin}}$ for a given $\boldsymbol{m} = (\boldsymbol{s}, \boldsymbol{c})$.

The l_μ function imposes a set of azimuthally integrated constraints on the underlying projected mass distribution. Since magnification is locally related to κ, this will essentially provide the otherwise unconstrained normalization of $\Sigma(\boldsymbol{\theta})$ over a set of concentric rings where count measurements are available. We note that no assumption is made of azimuthal symmetry or isotropy of the 2D mass distribution $\Sigma(\boldsymbol{\theta})$.

2.4.3. Calibration Parameters

We account for the calibration uncertainty in the observational nuisance parameters,

$$\boldsymbol{c} = \left(\langle W\rangle_g, f_{W,g}, \langle W\rangle_\mu, \bar{n}_\mu, \alpha\right). \tag{24}$$

To do this, we include in our analysis Gaussian priors on $\boldsymbol{c}$ given by means of quadratic penalty terms with mean values and errors directly estimated from data.

2.4.4. Best-fit Solution and Covariance Matrix

The log posterior $F(\boldsymbol{m}) = -\ln P(\boldsymbol{m}|\boldsymbol{d})$ is expressed as a linear sum of the log-likelihood and prior terms. The maximum-likelihood (ML) solution, $\hat{\boldsymbol{m}}$, is obtained by minimizing $F(\boldsymbol{m})$ with respect to $\boldsymbol{m}$. In our implementation, we use the conjugate-gradient method (Press et al. 1992) to find the solution. Here we employ an analytic expression for the gradient function $\boldsymbol{\nabla} F(\boldsymbol{m})$ obtained in the nonlinear, subcritical regime. To be able to quantify the errors on the reconstruction, we evaluate the Fisher matrix at $\boldsymbol{m} = \hat{\boldsymbol{m}}$ as

$$\mathcal{F}_{pp'} = \left\langle \frac{\partial^2 F(\boldsymbol{m})}{\partial m_p \partial m_{p'}} \right\rangle \Bigg|_{\boldsymbol{m}=\hat{\boldsymbol{m}}} \tag{25}$$

where the angular brackets represent an ensemble average, and the indices (p, p') run over all model parameters $\boldsymbol{m} = (\boldsymbol{s}, \boldsymbol{c})$. We estimate the error covariance matrix as

$$\mathrm{Cov}(m_p, m_{p'}) \equiv C_{pp'} = \left(\mathcal{F}^{-1}\right)_{pp'}. \tag{26}$$

3. SUBARU OBSERVATIONS

Here we present a description of our data analysis of A1689 based on deep Subaru $BVR_{\rm C}i'z'$ images. In this study, we analyze the data using the same methods and procedures as in Umetsu et al. (2014), who performed a weak-lensing analysis of 20 high-mass clusters selected from the CLASH survey (Postman et al. 2012). For details of our reduction and analysis pipelines, we refer to Section 4 of Umetsu et al. (2014).

3.1. Data and Photometry

We analyze deep $BVR_{\rm C}i'z'$ images of A1689 observed with the wide-field camera Suprime-Cam ($34' \times 27'$; Miyazaki et al. 2002) at the prime focus of the 8.3 m Subaru Telescope. We combine both existing archival data taken from SMOKA[14] with observations acquired by the team on the nights of 2010 March 17–18 (S10A-019). The observation details of A1689 are summarized in Table 2. Figure 1 shows a $BVR_{\rm C}i'z'$ composite color image of the cluster field, produced using the publicly available Trilogy software (Coe et al. 2012). The image is overlaid by mass contours determined from our weak-lensing analysis (see Section 4.2).

Table 2
Subaru/Suprime-Cam Data

Filter	Exposure time[a] (ks)	Seeing[b] (arcsec)	$m_{\rm lim}$[c] (AB mag)
B	2.40	0.91	27.1
V	4.08	0.84	27.0
$R_{\rm C}$	6.42	0.70 (0.60)	27.0
i'	4.08	0.84	26.4
z'	8.02	0.81	26.2

Note. The $R_{\rm C}$ band is used as the filter to measure object shapes for the weak-lensing analysis, where we separately stack data from different epochs. The $R_{\rm C}$-band seeing in parentheses is the average of values derived from separate stacks.
[a] Total exposure time.
[b] Seeing FWHM in the full stack of images.
[c] Limiting magnitude for a 3σ detection within a $2''$ aperture.

Our imaging reduction pipeline derives from Nonino et al. (2009) and has been optimized separately for accurate photometry and shape measurements. For multi-band photometry, standard reduction steps include bias subtraction, superflat-field correction, and point-spread function (PSF, hereafter) matching between exposures in the same band. An accurate astrometric solution is derived with the SCAMP software (Bertin 2006), using the the Sloan Digital Sky Survey (Adelman-McCarthy et al. 2008) as an external reference catalog.[15] The Swarp software (Bertin et al. 2002) is used to stack individual exposures on a common World Coordinate System (WCS) grid with pixel scale of $0.''2$.

The photometric zero-points for the co-added images were derived using *HST*/ACS magnitudes of cluster elliptical-type galaxies. These zero points were further refined by fitting spectral energy distribution (SED) templates with the BPZ code (Bayesian photometric redshift estimation; Benítez 2000; Benítez et al. 2004) to 1445 galaxies having spectroscopic redshifts.[16] This leads to a final photometric accuracy of $\sim$0.01 mag in all passbands. The magnitudes were corrected for Galactic extinction according to Schlegel et al. (1998). The multi-band photometry was measured using SExtractor (Bertin & Arnouts 1996) in dual-image mode on PSF-matched images created by ColorPro (Coe et al. 2006).

3.2. Shape Measurement

We use our shear analysis pipeline based on the IMCAT package (Kaiser et al. 1995, KSB) incorporating improvements developed by Umetsu et al. (2010). On the basis of simulated

[14] http://smoka.nao.ac.jp
[15] This research has made use of the VizieR catalog access tool, CDS, Strasbourg, France.
[16] The data used here are part of an extensive multi-object spectroscopy survey carried out with the VIMOS spectrograph on the VLT (Czoske 2004). For details, see Lemze et al. (2009).

Subaru/Suprime-Cam images (Massey et al. 2007; Oguri et al. 2012), Umetsu et al. (2010) showed that the lensing signal can be recovered with $|m| \sim 5\%$ of the multiplicative shear calibration bias (as defined by Heymans et al. 2006; Massey et al. 2007), and $c \sim 10^{-3}$ of the residual shear offset, which is about one order of magnitude smaller than the typical shear signal in cluster outskirts. Accordingly, we include for each galaxy a shear calibration factor of 1/0.95 ($g \rightarrow g/0.95$) to account for residual calibration.

In this work, we perform weak-lensing shape analysis using the same procedures adopted in the CLASH weak-lensing analysis of Umetsu et al. (2014). Here, we only highlight key aspects of our analysis pipeline.

1. *Object detection*. Objects are detected using the IMCAT peak finder, *hfindpeaks*, using a set of Gaussian kernels of varying sizes. This algorithm produces object parameters such as the peak position, the *best-matched* Gaussian scale length, r_g, and an estimate of the significance of the peak detection, ν.
2. *Crowding effects*. Objects having any detectable neighbors within $3r_g$ are identified. All such close pairs of objects are rejected to avoid possible shape measurement errors due to crowding. The detection threshold is set to $\nu = 7$ for close-pair identification. After this close-pair rejection, objects with low detection significance $\nu < 10$ are excluded from our analysis.
3. *Shear calibration*. We calibrate KSB's isotropic correction factor P_g as a function of object size (r_g) and magnitude, using galaxies detected with high significance $\nu > 30$ (Umetsu et al. 2010). This is to minimize the inherent shear calibration bias in the presence of noise. We correct for the isotropic smearing effect caused by seeing as well as by the window function used in the shape estimate as $g_\alpha = e_\alpha / P_g$ with e_α the anisotropy-corrected object ellipticity.

To measure the shapes of background galaxies, we use the R_C-band data, which have the best image quality in our data sets (Table 2). Two separate co-added R_C-band images are created, one from 2009 (observed by Matsuda et al.) and another from 2010 (observed by Umetsu et al.). We separately stack data obtained at different epochs. We do not smear individual exposures before stacking, so as not to degrade the weak-lensing signal. After PSF anisotropy correction, the mean residual stellar ellipticity is consistent with zero, and the rms residual stellar ellipticity in each stack is $\sigma(\delta e^*_\alpha) \sim 2.5 \times 10^{-3}$ per component. A shape catalog is created for each epoch separately. These subcatalogs are then combined by properly weighting and stacking the calibrated shear estimates for galaxies in the overlapping region (see Section 4.3 of Umetsu et al. 2014).

3.3. Background Galaxy Selection

A careful background selection is critical for a cluster weak-lensing analysis, so that unlensed objects do not dilute the true lensing signal of the background (Medezinski et al. 2007; Umetsu & Broadhurst 2008; Okabe et al. 2013; Hwang et al. 2014). In particular, dilution due to contamination by cluster members can lead to a substantial underestimation of the true signal at small cluster radii, $r \lesssim r_{2500c}$ (Medezinski et al. 2010; Okabe et al. 2010). The relative importance of the dilution effect indicates that, the impact of background purity and depth is more important than that of shot noise ($\propto \bar{n}_g^{-1/2}$).

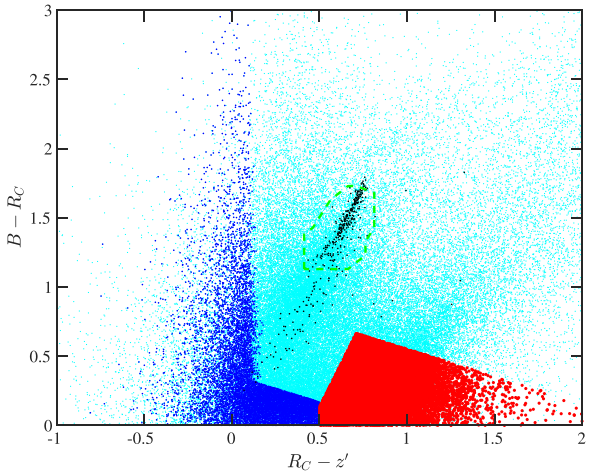

Figure 2. "Blue" and "red" background galaxy samples selected for the weak-lensing analysis (lower-left blue and lower-right red regions, respectively) on the basis of Subaru BR_Cz' color–color–magnitude selection. All galaxies (cyan) are shown in the diagram. At small clustercentric radius ($<4'$), an overdensity of cluster galaxies is identified as our "green" sample (green), comprising mostly the red sequence of cluster ellipticals and some blue trail of later-type cluster members. The background samples are well isolated from the green region and satisfy other criteria as discussed in Section 3.3. The black dots represent a dynamically selected spectroscopic sample of 377 cluster galaxies found within a projected distance of $12'$ ($\sim r_{200c}$) from the cluster center. Our background selection successfully excludes all except 2 spectroscopically confirmed cluster members (see Section 3.3).

We use the color–color (CC) selection method of Medezinski et al. (2010) to define uncontaminated samples of background galaxies from which to measure the shear and magnification effects. Here we refer the reader to Medezinski et al. (2010) for further details. Our multi-color approach and its variants have been successfully applied to a large number of clusters (Medezinski et al. 2010, 2011, 2013; Umetsu et al. 2010, 2011b, 2012, 2014; Coe et al. 2012; Oguri et al. 2012; Covone et al. 2014; Sereno et al. 2014b).

We use the Subaru BR_Cz' photometry, which spans the full optical wavelength range, to perform CC selection of background samples. In Figure 2, we show the $B - R_C$ versus $R_C - z'$ distribution of all galaxies to our limiting magnitudes (cyan). We select two distinct populations that encompass the red and blue branches of background galaxies in CC space, each with typical redshift distributions peaked around $z \sim 1$ and ~2, respectively (see Figures 5 and 6 of Lilly et al. 2007; Medezinski et al. 2011). The color boundaries of our "blue" and "red" background samples are shown in Figure 2.

As a cross-check we calculate the tangential (g_+) and cross ($g_\times$) reduced-shear components in clustercentric radial bins, which we show in Figure 3. In the absence of higher-order effects, weak lensing produces only curl-free tangential distortions, g_+. The presence of $\times$ modes can thus be used to check for systematic errors. Using the weak-lensing-matched blue and red samples, we find a consistent, rising distortion signal all the way to the cluster center. For all cases, the $\times$-component is consistent with a null signal detection well within 2σ at all radii.

For the number counts to measure magnification, we define flux-limited photometry samples of background galaxies. Here we limit the data to $z' = 25.6$ mag in the reddest band

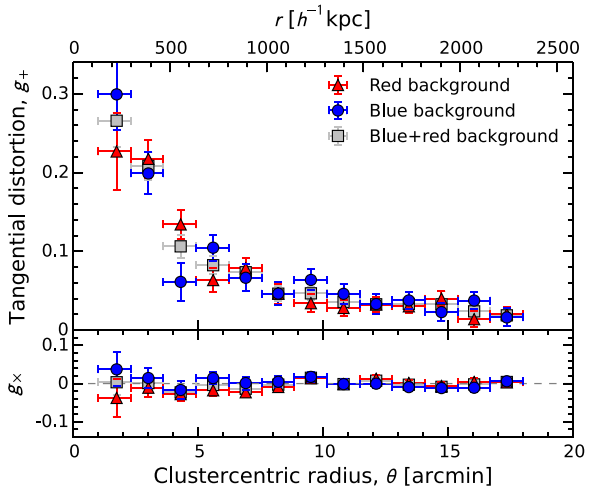

Figure 3. Azimuthally averaged radial profiles of the tangential lens distortion g_+ (upper panel) and the 45° rotated (×) component $g_\times$ (lower panel) for our red (triangles), blue (circles), and blue+red (squares) galaxy samples derived from Subaru multi-color photometry (Table 3).

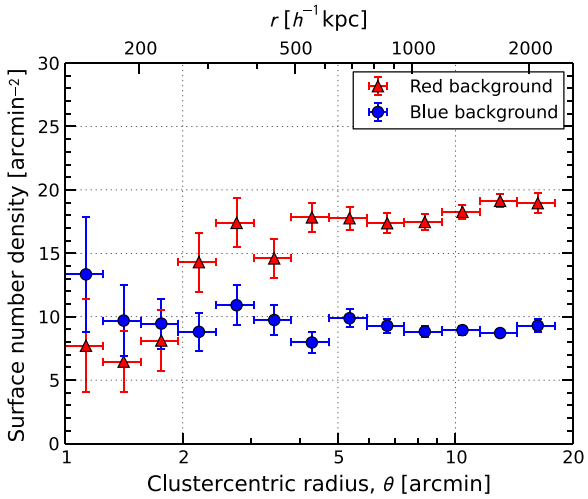

Figure 4. Coverage- and mask-corrected surface number density profiles of Subaru BR_Cz'-selected galaxy samples (Table 4). The results are shown for our red (triangles) and blue (circles) background samples. The error bars include contributions from Poisson counting uncertainties and contamination due to intrinsic clustering of each source population. For the red sample, a systematic radial depletion of the source counts is seen toward the cluster center owing to magnification of the sky area, while the faint blue counts are nearly constant with the distance from the cluster center. See also Figure 5.

(Table 4), corresponding to the 5σ limiting magnitude within $2''$ diameter aperture. We plot in Figure 4 the coverage- and mask-corrected surface number density as a function of clustercentric radius, for the blue and red samples. No clustering is observed toward the center, demonstrating that there is no detectable contamination by cluster members in the background samples. The red sample reveals a systematic decrease in their counts toward the cluster center, caused by magnfication of the sky area (Section 3.3). The faint blue counts, on the other hand, are nearly constant with cluster radius, as expected by their steep count slope (Table 4). A more quantitative magnification analysis is given in Section 4.1.

For validation purposes, we compare in Figure 2 our background samples with a dynamically selected spectroscopic sample of 377 cluster galaxies (black) found within a projected distance of $12'$ ($\sim r_{200c}$) from the cluster center. We find that our background selection procedure successfully excludes all except 2 spectroscopically confirmed cluster members (see also Coe et al. 2012; Umetsu et al. 2012), corresponding to a negligible contamination fraction of ~0.5%. We note that, in the blue background region, there are 4 cluster members, of which two are excluded by the magnitude cuts used to reject bright foreground/cluster galaxies.

We estimate the mean depths ($\langle\beta\rangle$, $\langle\beta^2\rangle$) of the background samples (Tables 3 and 4), which are necessary when converting the observed lensing signal into physical mass units. For this, we follow the prescription outlined in Section 4.4 of Umetsu et al. (2014). We utilize BPZ to measure photo-zs using our PSF-corrected Subaru $BVR_Ci'z'$ photometry. Following Umetsu et al. (2012), we employ BPZ's ODDS parameter as the weight factor $w(z)$ in Equation (4). The resulting depth estimates are summarized in Tables 3 and 4.

4. SUBARU WEAK-LENSING ANALYSIS

We use our z'-band limited sample of red background galaxies (Table 4) for magnifciation measurements and a full composite sample of blue+red galaxies (Table 3) for shear measurements. In Section 4.1, we perform a 1D weak-lensing analysis of A1689 to derive azimuthally averaged lensing profiles from our new Subaru data (Section 3), and examine the consistency of complementary shear and magnification measurements. In Section 4.2, we apply the 2D inversion method developed in Section 2 and reconstruct the projected 2D mass distribution from joint shear+magnification measurements.

Table 3
Background Galaxy Samples For Weak-lensing Shape Measurements

Sample	N_g	$\bar{n}_g$ [a] (arcmin^{-2})	$\bar{z}_{\rm eff}$ [b]	$\langle D_{\rm ls}/D_{\rm s}\rangle$	f_W
Red	12674	12.0	1.10	0.79 ± 0.04	1.00
Blue	9238	8.7	1.62	0.84 ± 0.04	1.01
Blue+red	21912	20.7	1.22	0.80 ± 0.04	1.01

Notes.
[a] Mean surface number density of source background galaxies.
[b] Effective source redshift corresponding to the mean lensing depth $\langle\beta\rangle = \langle D_{\rm ls}/D_{\rm s}\rangle$, defined as $\beta(\bar{z}_{\rm eff}) = \langle\beta\rangle$.

4.1. Weak-lensing Profiles of A1689

A1689 exhibits a small offset $d_{\rm off} \simeq 5\,{\rm kpc}\,h^{-1}$ ($\simeq 2\farcs3$) between the BCG and X-ray centroids (Table 1), ensuring a well-defined center. The X-ray and SZE centroids agree to within $1''$ (Table 1). Here we will adopt the BCG position as the cluster center for a radial profile analysis.

We derive azimuthally averaged radial profiles of tangential reduced shear (g_+) and magnification bias (n_μ) from Subaru data. We calculate the lensing profiles in $N_{\rm bin} = 13$ discrete radial bins, spanning the range $[\theta_{\rm min}, \theta_{\rm max}] = [1', 18']$ with a constant logarithmic spacing, $\Delta \ln\theta = \ln(\theta_{\rm max}/\theta_{\rm min})/N_{\rm bin} \simeq 0.22$. The inner radial limit $r_{\rm min} \equiv D_l\theta\ {\rm min} \simeq 129\ {\rm kpc}\,h^{-1}$ is sufficiently greater than the Einstein radius $\theta_{\rm Ein} = 47\farcs0 \pm 1\farcs2$ ($z_s = 2$; Table 1), and it also satisfies $r_{\rm min} > 2d_{\rm off} \simeq 10\ {\rm kpc}\,h^{-1}$, so that the miscentering effects on mass profile reconstructions are negligible (Johnston et al. 2007; Umetsu et al. 2011a; Du & Fan 2014). The outer boundary $\theta_{\rm max} = 18'$, or

Table 4
Background Galaxy Samples for Magnification-bias Measurements

Sample	$z'_{\rm cut}$ [a] (AB mag)	N_μ	$\overline{n}_\mu$ [b] (arcmin^{-2})	α [c]	$\langle z\rangle$ [d]	$\overline{z}_{\rm eff}$ [e]	$\langle D_{\rm ls}/D_{\rm s}\rangle$
Red	25.6	26136	19.0 ± 0.5	0.39 ± 0.08	1.13	1.05	0.73 ± 0.04
Blue	25.6	12143	8.8 ± 0.3	0.82 ± 0.12	1.81	1.39	0.82 ± 0.04

Notes.
[a] Fainter magnitude cut of the background sample. Apparent magnitude cuts are applied in the reddest CC-selection band available (z') to avoid incompleteness near the detection limit.
[b] Coverage- and mask-corrected normalization of unlensed background source counts.
[c] Logarithmic slope of the unlensed source counts $\alpha = 2.5\left[d\,\log_{10}\overline{N}_\mu(<z')/dz'\right]_{z'=z'_{\rm cut}}$.
[d] Mean photometric redshift of the sample obtained with the BPZ code, defined similarly to Equation (4).
[e] Effective source redshift corresponding to the mean lensing depth $\langle\beta\rangle = \langle D_{\rm ls}/D_{\rm s}\rangle$, defined as $\beta(\overline{z}_{\rm eff}) = \langle\beta\rangle$.

$r_{\rm max} \equiv D_{\rm l}\theta_{\rm max} \simeq 2.3$ Mpc h^{-1}, is large enough to encompass the entire virial region with $r_{\rm vir} \simeq 2$ Mpc h^{-1} (Umetsu & Broadhurst 2008), but sufficiently small compared to the size of the Suprime-Cam field of view so as to ensure accurate PSF anisotropy correction. The number of bins $N_{\rm bin} = 13$ is chosen such that the detection S/N is of the order of unity per bin, which is optimal for an inversion problem.

In this work, we follow the prescription outlined in Section 3.2.2 of Umetsu et al. (2014) to perform magnification measurements using the Subaru $BR_{\rm C}z'$-selected red galaxy sample (Table 4), which exhibits a clear depletion signal (Figure 4). We have properly accounted and corrected for masking of background galaxies due to cluster galaxies, foreground objects, and saturated pixels (see also Section 2.3.2). Unlike the nonlocal distortion signal, the magnification signal falls off sharply with increasing cluster radius. We thus estimate the count normalization and slope $(\overline{n}_\mu, \alpha)$ from the source counts in cluster outskirts (Umetsu et al. 2011b, 2012, 2014; Medezinski et al. 2013), specifically at $12' \ (\sim r_{200c}) < \theta < \theta_{\rm max}$.

Figure 5 shows the radial profiles of (g_+, n_μ). A clear depletion of red galaxies is seen toward the center owing to geometric magnification of the sky area. The statistical significance of the detection of the tangential distortion is 22σ. The detection significance of the magnification signal is 9σ, which is ∼40% of that of distortion.

Here we construct the radial mass profile of A1689 from a joint likelihood analysis of shear and magnification measurements (Figure 5), using the method of Umetsu et al. (2011b). We have 26 constraints $\{g_{+,i}, n_{\mu,i}\}_{i=1}^{N_{\rm bin}}$ in 13 log-spaced clustercentric radial bins. The model is described by $N_{\rm bin} + 1 = 14$ parameters, $\{\Sigma_{\rm min}, \Sigma_i\}_{i=1}^{N_{\rm bin}}$, where $\Sigma_{\rm min} \equiv \Sigma(<\theta_{\rm min})$ is the average surface mass density interior to $\theta_{\rm min}$, and Σ_i is the surface mass density averaged in the ith radial bin. To perform a reconstruction, we express the lensing observables (g_+, μ^{-1}) in terms of Σ using the relations given in Appendix B. Additionally, we account for the calibration uncertainty in the observational parameters $\boldsymbol{c} = (\langle W\rangle_g, f_{W,g}, \langle W\rangle_\mu, \overline{n}_\mu, \alpha)$ as given in Tables 3 and 4. Following Umetsu et al. (2014), we fix $f_{W,g}$ to the observed value (Table 3).

The results are shown in Figures 5 and 6. The ML solution has a reduced χ^2 of 11.5 for 12 degrees of freedom (dof), indicating good consistency between the shear and magnification measurements having different potential systematics. This is demonstrated in Figure 5, which compares the observed lensing profiles with the respective joint reconstructions

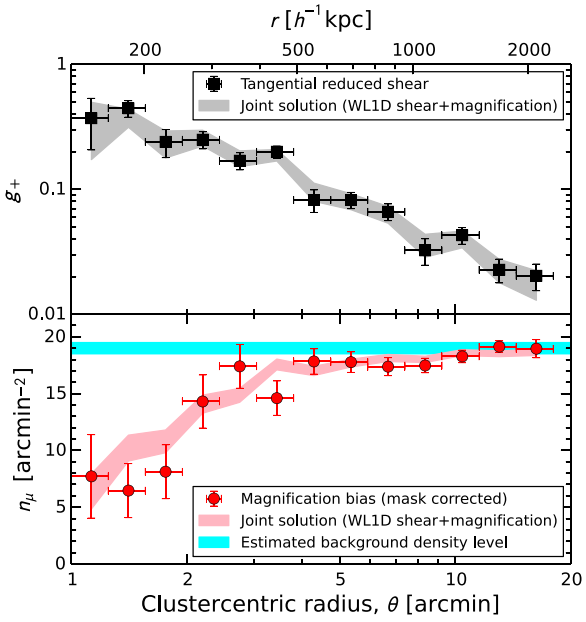

Figure 5. Azimuthally averaged cluster weak-lensing profiles obtained from Subaru multi-color observations of A1689. The upper panel shows the tangential reduced shear profile g_+ (black squares) based on the full background sample. The lower panel shows the magnification-bias profile n_μ (red circles) of a z'-band limited sample of red background galaxies. For each observed profile, the shaded area represents the joint reconstruction (68% CL) from the combined shear+magnification measurements. The horizontal bar (cyan shaded region) shows the constraints on the unlensed count normalization estimated from the source counts in cluster outskirts.

(68% CL). The resulting mass profile $\Sigma(\theta)$ is shown in the upper panel of Figure 6. The error bars represent the 1σ errors from the diagonal part of the total covariance matrix C (Umetsu et al. 2014). The corresponding cumulative mass profile is shown in the lower panel of Figure 6.

4.2. Weak-lensing Mapmaking of A1689

We apply our 2D inversion method (Section 2) to our new Subaru observations (Sections 3) for obtaining an unbiased recovery of the projected matter distribution $\Sigma(\boldsymbol{\theta})$ in A1689. In this approach, we combine the observed spatial shear pattern $(g_1(\boldsymbol{\theta}), g_2(\boldsymbol{\theta}))$ with the azimuthally averaged magnification measurements $\{n_{\mu,i}\}_{i=1}^{N_{\rm bin}}$ (Section 4.1), which impose a set of azimuthally integrated constraints on the underlying $\Sigma(\boldsymbol{\theta})$ field, thus effectively breaking the mass-sheet degeneracy. The

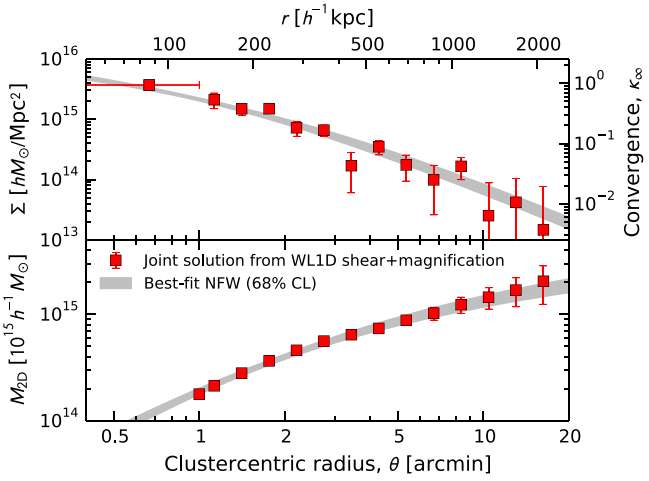

Figure 6. Surface mass density profile $\Sigma(\theta)$ (upper panel, red squares) derived from a Subaru 1D weak-lensing analysis of the combination of shear and magnification measurements shown in Figure 5. The lower panel shows the corresponding cumulative mass profile $M_{\rm 2D}(<\theta)$ (red squares). The gray area in each panel represents the best-fit projected Navarro–Frenk–White profile (68% CL) for the mass profile solution $\Sigma(\theta)$.

algorithm takes into account the nonlinear subcritical regime of the lensing properties.

For mapmaking, we pixelize the lensing fields into a 56×56 grid with $\Delta\theta = 0.'5$ spacing, covering the central $28' \times 28'$ field. The model $\boldsymbol{m} = (\boldsymbol{s}, \boldsymbol{c})$ is specified by $N_{\rm pix} = 56^2$ parameters, $\boldsymbol{s} = \{\Sigma(\boldsymbol{\theta}_m)\}_{m=1}^{N_{\rm pix}}$, and a set of calibration parameters $\boldsymbol{c}$ to marginalize over. We utilize the FFTW implementation of fast Fourier transforms to calculate $\gamma_\infty(\boldsymbol{\theta})$ from $\kappa_\infty(\boldsymbol{\theta})$ using Equation (6). To minimize spurious aliasing effects from the periodic boundary condition, the maps are zero padded to twice the original length in each spatial dimension (e.g., Seljak 1998; Umetsu & Broadhurst 2008).

We use a top-hat window of $\theta_{\rm f} = 0.'4$ (Section 2.3.1) to average over a local ensemble of galaxy ellipticities ($N = \pi \bar{n}_g \theta_{\rm f}^2 \sim 10$; Merten et al. 2014) at each grid point, accounting for the intrinsic ellipticity distribution of background sources. The filter size corresponds to an effective resolution of $2D_1\theta_{\rm f} \simeq 100$ kpc h^{-1} at the cluster redshift. To avoid potential systematic errors, we exclude from our analysis (Section 2.4.1) those pixels lying within central $\theta_{\rm cut} = 1'$ where $\Sigma(\boldsymbol{\theta})$ can be close to or greater than the critical value $\Sigma_{\rm c}$, as well as those containing no background galaxies with usable shape measurements. For distortion measurements $(g_1(\boldsymbol{\theta}), g_2(\boldsymbol{\theta}))$ from the full background sample (Table 3), this leaves us with a total of 3093 usable measurement pixels (blue points in Figure 7), corresponding to 6186 constraints. For magnification measurements, we have 13 azimuthally averaged constraints $\{n_{\mu,i}\}_{i=1}^{N_{\rm bin}}$ in log-spaced clustercentric annuli (Figure 7). The total number of constraints is thus $N_{\rm data} = 6199$, yielding $N_{\rm data} - N_{\rm pix} = 3063$ dof.

In Figure 8, we show the resulting $\Sigma(\boldsymbol{\theta})$ field reconstructed from a joint analysis of the 2D shear and azimuthally averaged magnification data. The χ^2 value for the ML solution is $\chi^2(\hat{\boldsymbol{m}}) = 4046$ for 3063 dof. Here, for visualization purposes, the $\Sigma(\boldsymbol{\theta})$ field is smoothed with a Gaussian of FWHM $= 1'$. The main mass peak coincides well with the cluster center.

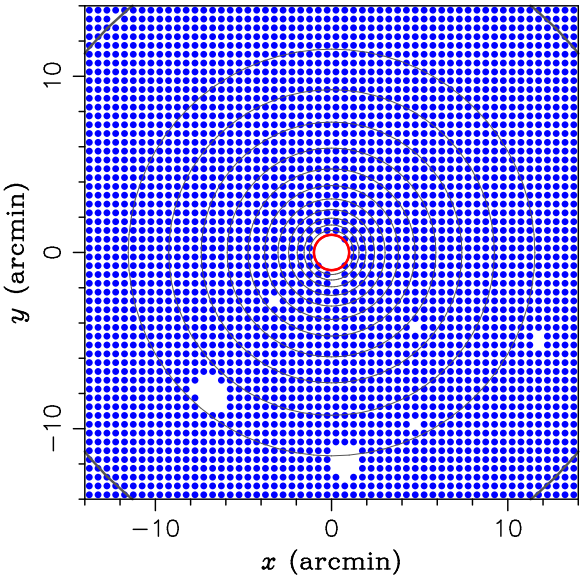

Figure 7. Spatial distribution of weak-lensing constraints averaged onto a grid of 56×56 pixels, covering a field of $28' \times 28'$ centered on the BCG. Each point represents a single pixel with two-component reduced shear constraints (g_1, g_2) averaged within a top-hat region with radius $\theta_{\rm f} = 0.'4$. We exclude from our analysis those pixels lying within the inner $\theta_{\rm cut} = 1'$ region (red circle) and those having no background galaxies with usable shape measurements (see Figure 1). There are 3093 pixels with reduced-shear constraints, yielding 6186 constraints from 2D shear measurements. Azimuthally averaged magnification constraints are obtained in 13 logarithmically spaced, clustercentric annuli spanning the range $[\theta_{\rm min}, \theta_{\rm max}] = [1', 18']$.

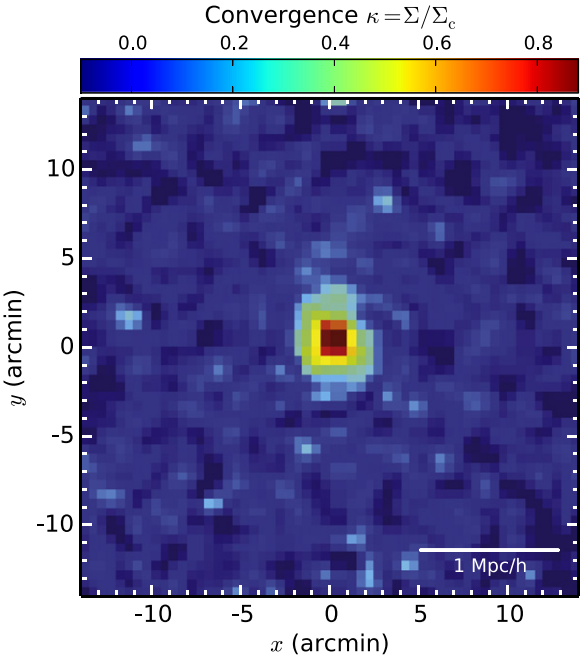

Figure 8. Projected mass distribution $\Sigma(\boldsymbol{\theta})$ of A1689 reconstructed from a Subaru weak-lensing analysis of 2D gravitational shear and azimuthally averaged magnification data. The mass maps is $28' \times 28'$ in size (3.6 Mpc h^{-1} on a side) and centered on the BCG. The color bar indicates the lensing convergence $\kappa = \langle \Sigma_{\rm c}^{-1} \rangle \Sigma$, scaled to the mean depth of weak-lensing observations, $1/\langle \Sigma_{\rm c}^{-1} \rangle = 4.66 \times 10^{15}\, h\, M_\odot$ Mpc^{-2}. For visualization purposes, the mass map is smoothed with a $1'$ FWHM Gaussian. North is to the top, east to the left. The horizontal bar represents 1 Mpc h^{-1} at the cluster redshift.

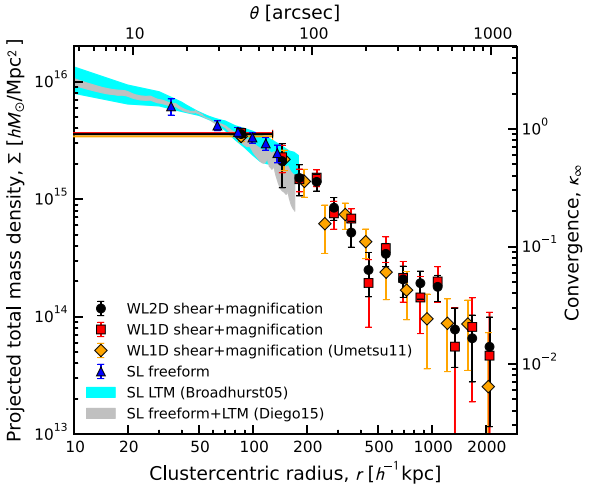

Figure 9. Comparison of projected mass density profiles $\Sigma(r)$ derived from our Subaru 1D weak-lensing analysis (squares; Section 4.1), Subaru 2D weak-lensing analysis (circles; Section 4.2), and free-form strong-lensing analysis of *HST* data (triangles; Section 5). The cyan shaded area represents the mass profile with 1σ uncertainty from a strong-lensing analysis of Broadhurst et al. (2005b) based on the light-traces-mass (LTM) assumption. The gray shaded area shows the strong-lensing results (68% CL) from Diego et al. (2015) using a hybrid scheme combining both free-form grid and LTM substructure components. The diamonds with error bars show the results from our earlier 1D weak-lensing analysis (Umetsu et al. 2011b) based on Subaru *Vi′* data. Good agreement between the strong and weak lensing results is seen in the region of overlap. There is also good agreement between the different lensing methods and data sets.

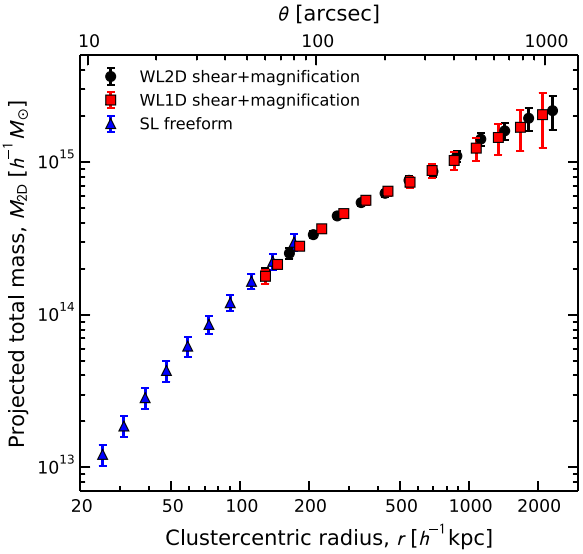

Figure 10. Comparison of projected cumulative mass profiles $M_{2D}(<r)$ of A1689 derived from our Subaru 1D weak-lensing analsyis (squares; Section 4.1), Subaru 2D weak-lensing analysis (circles; Section 4.2), and *HST* strong-lensing analysis (triangles; Section 5).

The projected mass distribution is elongated in the north–south direction (Figure 1; see also Section 7.1) and very similar to the distribution of cluster members (Figure 1).

In Figure 9, we compare the projected mass profiles $\Sigma(\theta)$ obtained from our 1D and 2D analyses of the shear+magnification data. Here we have used the method described in Appendix C to construct an optimally weighted radial projection of the $\boldsymbol{\Sigma}$ map. Our 1D- and 2D-based Σ profiles are consistent within 1σ at all cluster radii, and both are in good agreement with the 1D results of Umetsu et al. (2011b) from the joint shear+magnification analysis of the Subaru *Vi′* data. Similarly, our 1D and 2D weak-lensing results are in excellent agreement with each other in terms of the cumulative mass $M_{2D}(<\theta)$ as shown in Figure 10.

5. *HST* STRONG-LENSING ANALYSIS

5.1. Image Systems

A1689 has been a subject of detailed strong-lensing studies by numerous authors (e.g., Broadhurst et al. 2005b; Halkola et al. 2006; Limousin et al. 2007; Coe et al. 2010; Diego et al. 2015). Thus far, a total of 61 multiple-image candidate systems of 165 images were identified from extremely deep optical and near-infrared data from *HST* and Subaru (Diego et al. 2015).

To study global structural properties of the cluster, we focus our strong-lensing analysis on the principal modes of the cluster mass distribution, responsible for the massive, smooth halo component (see Section 7.1.2). To this end, we conservatively select a subset of systems based on the following criteria: (i) we use only spectroscopically confirmed systems. (ii) We consider only systems whose members were consistently identified in different studies. (iii) We limit our analysis to those lying within 80″ from the BCG, so that multiple images spread fairly evenly over the analysis region. (iv) We discard systems of very close pairs. They are primarily sensitive to substructures rather than the principal modes of the mass distribution, which we are interested in.

These criteria leave us with 12 systems (ID 1, 2, 4, 5, 6, 7, 11, 15, 18, 22, 24, 29, according to the original notation in Broadhurst et al. 2005b), for a total of 44 multiple images spanning the range 1″.4–72″.3 in cluster radius.

5.2. PixeLens Free-form Mass Reconstruction

Free-form models describe the lens on a grid of pixels or a set of basis functions, allowing for a wide range of solutions (Coles 2008). We have performed a free-form strong-lensing analysis of the central region using the PixeLens software (Saha & Williams 2004), which produces pixelated maps of the surface mass density. Each map is constrained to exactly reproduce the positions and parities of all given multiple images. PixeLens generates a statistical ensemble of models through which uncertainties and degeneracies in solutions can be explored (Coles 2008).

Our PixeLens analysis procedure largely follows Sereno & Zitrin (2012) and Sereno et al. (2013). To determine robust sampling strategies optimized to recover the smooth cluster signal, we tested the PixeLens algorithm using simulated sets of multiple images in analytic lenses. The results suggest that the best strategy is to limit each analysis to three image systems, for a total of a dozen of images, and to reconstruct maps with ∼10 pixels in the radial direction, avoiding oversampling (Lubini & Coles 2012). We thus divide the strongly lensed images in four groups of three systems each and analyze each group separately. We end up with four triples consisting of systems 1, 5, and 11 (11 images), systems 2, 6, and 22 (11 images), systems 4, 15, and 29 (12 images), and systems 7, 18, and 24 (10 images). Image systems with similar configurations are divided into different groups.

For each group, we compute 500 κ maps within 80″ from the BCG on a circular grid of 349 pixels (10 pixels along the radial

direction) with a pixel size of $8''$ ($\simeq 17.2\,\mathrm{kpc}\,h^{-1}$). These optimal settings allow us to avoid the known problem of too flat density profiles recovered with PixeLens modeling (see Grillo et al. 2010; Umetsu et al. 2012), which otherwise could bias cluster mass estimates. As discussed by Grillo et al. (2010, see their appendix), this bias can arise from a combination of the mass-sheet degeneracy (Schneider & Seitz 1995) and the assumed prior on the positive definiteness of every pixel of the surface mass density map.

In the following, we restrict our analysis to the region where the cluster mass distribution is accurately recovered by PixeLens. We exclude the central $20\,\mathrm{kpc}\,h^{-1}$ region to minimize the effects of miscentering and baryonic physics (Umetsu et al. 2012, 2014). For each group of reconstruction, we determine the outer cutoff radius beyond which the logarithmic density slope is steeper than −2, the asymptotic minimum slope for the projected Navarro–Frenk–White density profile (NFW, Navarro et al. 1997). The maximum radius is $63.'7$ (188 mass pixels) in three cases and $54.'9$ (140 mass pixels) for the group with the triple 4–15–29.

5.3. Comparison of Weak and Strong Lensing Results

We show in Figure 9 the radial mass distribution of A1689 from our *HST* strong-lensing analysis. The results are shown along with the previous strong-lensing results by Broadhurst et al. (2005b) and Diego et al. (2015), as well as with independent weak-lensing results from shear and magnification information (Sections 4.1 and 4.2). The strong-lensing model of Broadhurst et al. (2005b) is based on the light-traces-mass (LTM) assumption, so that the *HST* photometry of cluster red-sequence galaxies was used as an initial guess for their lens solution. Diego et al. (2015) used a hybrid (free-form + LTM) approach combining Gaussian pixel grid and cluster member components for describing large- and small-scale contributions to the deflection field, respectively. They constrained the range of solutions with sufficient accuracy to allow the detection of new counter images for further improving the lensing solution of A1689. This comparison shows clear consistency among a wide variety of lensing methods with different assumptions and potential systematics, demonstrating the robustness of our results (see also Figure 10). Excellent agreement is also found between our strong-lensing mass profile and that of Limousin et al. (2007).

6. TRIAXIAL MODELING OF THE CLUSTER MATTER DISTRIBUTION

Since we can only observe clusters in projection, determining the intrinsic 3D shape and orientation of an aspherical cluster is an intrinsically underconstrained problem (Sereno 2007). In this section, we describe the modeling of the 3D cluster matter distribution as an ellipsoidal halo following Sereno et al. (2013). In this approach, we exploit the combination of X-ray and SZE observations to constrain the elongation of the ICM along the line of sight. We use minimal geometric assumptions about the matter and gas distributions to couple the constraints from lensing and SZE/X-ray data. The parameter space is explored in a Bayesian inference framework. This multi-probe method allows us to improve constraints on the intrinsic shape and orientation of the cluster mass distribution without assuming HSE.

6.1. Matter Distribution

We model the cluster mass distribution with a triaxial NFW density profile as motivated by cosmological *N*-body simulations (Jing & Suto 2002; Kasun & Evrard 2005). The radial dependence of the spherical NFW density profile is given by (Navarro et al. 1996, 1997)

$$\rho(r) = \frac{\rho_{\rm s}}{(r/r_{\rm s})\left(1 + r/r_{\rm s}\right)^2} \tag{27}$$

with $\rho_{\rm s}$ the characteristic density and $r_{\rm s}$ the inner characteristic radius at which the logarithmic slope of the density profile is −2. We generalize the spherical NFW model to obtain a triaxial density profile by replacing r and $r_{\rm s}$ with the respective ellipsoidal radii R and $R_{\rm s}$, defined such that

$$R^2 = c^2\left(\frac{X^2}{a^2} + \frac{Y^2}{b^2} + \frac{Z^2}{c^2}\right) = \frac{X^2}{q_a^2} + \frac{Y^2}{q_b^2} + Z^2, \tag{28}$$

where $q_a = a/c$ and $q_b = b/c$ ($a \leqslant b \leqslant c$) are the minor–major and intermediate–major axis ratios, respectively.[17] The corresponding eccentricities are $e_a = \sqrt{1 - q_a^2}$ and $e_b = \sqrt{1 - q_b^2}$. The degree of triaxiality is defined as $\mathcal{T} = e_b^2/e_a^2$ (Sereno et al. 2013).

We define an ellipsoidal overdensity radius R_Δ (e.g., Corless et al. 2009; Sereno & Umetsu 2011; Buote & Humphrey 2012b) such that the mean interior density contained within an ellipsoidal volume of semimajor axis R_Δ is $\Delta \times \rho_{\rm c}$. The total mass enclosed within R_Δ is $M_\Delta = (4\pi/3)\Delta q_a q_b \rho_{\rm c} R_\Delta^3$. We use $\Delta = 200$ to define the halo mass, $M_{200{\rm c}}$. The triaxial concentration parameter is defined by $c_{200{\rm c}} = R_{200{\rm c}}/R_{\rm s}$. The characteristic density is then expressed as $\rho_{\rm s} = M_\Delta\big/(4\pi q_a q_b R_\Delta^3) \times c_\Delta^3\big/[\ln(1 + c_\Delta) - c_\Delta/(1 + c_\Delta)]$ (Buote & Humphrey 2012b).

A triaxial halo is projected on to the sky plane as elliptical isodensity contours (Stark 1977), which can be expressed as a function of the intrinsic halo axis ratios $(a/c, b/c)$ and orientation angles (ϑ, ϕ, ψ) with respect to the observer's line of sight. Here we adopt the z–x–z convention of Euler angles to be consistent with Stark (1977; see, e.g., Sereno et al. 2012). The angle ϑ describes the inclination of the major (Z) axis with respect to the line of sight.

For a given projection, the elliptical projected mass distribution can be described as a function of the elliptical radius ζ defined in terms of the observer's coordinates (X', Y') in the plane of the sky:

$$\zeta^2 = \frac{1}{f}\left(jX'^2 + 2\,kX'Y' + lY'^2\right) \equiv \frac{X''^2}{q_{\perp X}^2} + \frac{Y''^2}{q_{\perp Y}^2} \tag{29}$$

where $q_{\perp X}$ and $q_{\perp Y}$ ($q_{\perp X} \geqslant q_{\perp Y}$) are

$$q_{\perp X}^2 = \frac{2f}{j + l - \sqrt{(j - l)^2 + 4k^2}},$$
$$q_{\perp Y}^2 = \frac{2f}{j + l + \sqrt{(j - l)^2 + 4k^2}}, \tag{30}$$

[17] The intrinsic axis ratios (q_a, q_b) here correspond to $(\eta_{{\rm DM},a}, \eta_{{\rm DM},b})$ of Limousin et al. (2013) in their notation.

with

$$j = \cos^2\vartheta\left(\frac{c^2}{a^2}\cos^2\phi + \frac{c^2}{b^2}\sin^2\phi\right) + \frac{c^2}{a^2}\frac{c^2}{b^2}\sin^2\vartheta,$$
$$k = \sin\phi\cos\phi\cos\vartheta\left(\frac{c^2}{a^2} - \frac{c^2}{b^2}\right),$$
$$l = \frac{c^2}{a^2}\sin^2\phi + \frac{c^2}{b^2}\cos^2\phi,$$
$$f = \sin^2\vartheta\left(\frac{c^2}{a^2}\sin^2\phi + \frac{c^2}{b^2}\cos^2\phi\right) + \cos^2\vartheta. \quad (31)$$

Here we have chosen the new coordinate system (X'', Y'') such that the X'' axis is aligned with the major axis of the projected ellipse.

The minor–major axis ratio $q_\perp \equiv q_{\perp Y}/q_{\perp X}$ of the elliptical density contours is given by[18]

$$q_\perp(a/c, b/c, \vartheta, \phi) = \left[\frac{j + l - \sqrt{(j-l)^2 + 4k^2}}{j + l + \sqrt{(j-l)^2 + 4k^2}}\right]^{1/2}. \quad (32)$$

The principal axes of the isodensities are rotated by an angle ψ with respect to the projection on to the sky of the intrinsic major axis Z, where $2\psi = \arctan[2k/(j-l)]$ (Sereno 2007). As observable parameters to describe the projected mass distribution, we use the ellipticity

$$\epsilon = 1 - q_\perp \quad (33)$$

and the position angle ψ_ϵ of the projected major axis.

The projected surface mass density $\Sigma(\zeta)$ as a function of the elliptical radius ζ is related to the triaxial density profile $\rho(R)$ by (Stark 1977)

$$\Sigma(\zeta) = \frac{2}{\sqrt{f}}\int_\eta^\infty \frac{\rho(R)RdR}{\sqrt{R^2 - \zeta^2}} = \frac{2R_s}{\sqrt{f}}\int_{\zeta/R_s}^\infty \frac{\rho(R_s x)xdx}{\sqrt{x^2 - \left(\xi/\xi_s\right)^2}}, \quad (34)$$

where $\xi \equiv q_{\perp X}\zeta = \sqrt{X''^2 + Y''^2/q_\perp^2}$ is the *observable* elliptical radius, and $\xi_s = q_{\perp X}R_s$ is the observable scale length (semi-major axis) in the sky plane (Sereno 2007). The quantity $l_\parallel = R_s/\sqrt{f}$ represents the line of sight half length of the ellipsoid of radius $R = R_s$ (Sereno 2007). It is useful to introduce the dimensionless scale factor $e_\parallel$ that quantifies the extent of the cluster along the line of sight (Sereno 2007),

$$e_\parallel(a/c, b/c, \vartheta, \phi) = \frac{l_\parallel}{\xi_s} = \left(\frac{q_\perp}{q_a q_b}\right)^{1/2} f^{-3/4}. \quad (35)$$

The larger $e_\parallel$, the larger the elongation along the line of sight. The quantity $e_\parallel$ corresponds to the inverse of the elongation parameter e_Δ of Sereno (2007): $e_\parallel = 1/e_\Delta$.

[18] Note the projected axis ratio $q_\perp$ is equivalent to $1/e_P$ of Sereno (2007).

For a self-similar model $\rho(R) = \rho_s f_{3D}(R/R_s)$, the projected mass density profile is expressed as

$$\Sigma(\zeta) = \frac{2R_s\rho_s}{\sqrt{f}}\int_{\zeta/R_s}^\infty \frac{f_{3D}(x)xdx}{\sqrt{x^2 - \left(\xi/\xi_s\right)^2}} \equiv \Sigma_s f_{2D}\left(\xi/\xi_s\right), \quad (36)$$

where we have defined the scale surface mass density

$$\Sigma_s \equiv 2\rho_s R_s/\sqrt{f} = 2\rho_s\xi_s e_\parallel = 2f_{geo}\rho_s\sqrt{q_\perp}\xi_s \quad (37)$$

with $f_{geo} \equiv e_\parallel/\sqrt{q_\perp}$ (Sereno et al. 2010). Since $r_{s,2D} \equiv \sqrt{q_\perp}\xi_s$ is the geometric-mean scale radius in projection, the geometrical factor f_{geo} represents the degree of correction due to the line of sight elongation of the cluster. The halo mass, M_{200c}, can then be expressed as $M_{200c} = (4\pi/3)200\rho_c(c_{200c}r_{s,2D})^3 f_{geo}$. In this work, we employ the radial dependence of the projected NFW profile $f_{2D}(x)$ as given by Wright & Brainerd (2000). For $f_{geo} = 1$, this reduces to a projected (circular or elliptical) mass model. An elliptical mass density model can be described by $(M_{200c}, c_{200c}, \epsilon, \psi_\epsilon)$ (Oguri et al. 2010; Umetsu et al. 2012).

6.2. Intracluster Gas

Both observations and theory indicate that the ICM density is nearly constant on a family of concentric, coaxial ellipsoids (Kawahara 2010; Buote & Humphrey 2012a, 2012b). Although modeling both the gas and matter distributions as ellipsoids with constant axis ratios is not strictly valid for halos in HSE (Sereno et al. 2013), an ellipsoidal approximation for the ICM is suitable when systems with modest eccentricities are considered (Lee & Suto 2003).

Following Sereno et al. (2013), we make a few simplifying but non-informative working hypotheses to relate the matter and gas distributions. First, we assume that the matter and gas distributions in the cluster are ellipsoidal with constant but different axis ratios and co-aligned with each other. Second, the two distributions are assumed to have the same degree of triaxiality, that is, $\mathcal{T}(q_a, q_b) = \mathcal{T}^{ICM}(q_a^{ICM}, q_b^{ICM})$ with $\mathcal{T}^{ICM} \equiv (e_b^{ICM}/e_a^{ICM})^2 = [1 - (q_b^{ICM})^2]/[1 - (q_a^{ICM})^2]$ and $q_a^{ICM} \leqslant q_b^{ICM}$. If two ellipsoids have the same degree of triaxiality, then the misalignment angle between their major axes in the plane of the sky is zero (Romanowsky & Kochanek 1998), which is consistent with what has been observed in A1689 (Sereno & Umetsu 2011; Sereno et al. 2012). If $\mathcal{T} = \mathcal{T}^{ICM}$, we have the following relation for the ratio of eccentricities between ICM and matter (Sereno et al. 2013):

$$e_a^{ICM}/e_a = e_b^{ICM}/e_b \equiv e^{ICM}/e. \quad (38)$$

The intracluster gas in HSE is rounder than the underlying matter distribution: $e^{ICM}/e \simeq 0.7$ (Lee & Suto 2003).

With these assumptions, the number of independent axis ratios is reduced to three. Here we use q_a, q_b, and q_a^{ICM} as free parameters. Hence, the intermediate–major axis ratio q_b^{ICM} of the ICM is determined by $\mathcal{T}(q_a, q_b)$ and q_a^{ICM}:

$$q_b^{ICM} = \sqrt{1 - \frac{1 - \left(q_a^{ICM}\right)^2}{\mathcal{T}^2}}. \quad (39)$$

Finally, as supported by both theory and observations, we assume that the gas distribution is rounder than the matter distribution: $q_a \leqslant q_a^{\rm ICM}$.

Under these hypotheses, the projected matter and gas distributions of the cluster have different ellipticities ($\epsilon \neq \epsilon^{\rm ICM}$) and elongations ($e_\parallel \neq e_\parallel^{\rm ICM}$) but share the same orientation of the projected major axis, $\psi_\epsilon = \psi_\epsilon^{\rm ICM}$. There are a total of six parameters $(q_a, q_b, q_a^{\rm ICM}, \vartheta, \phi, \psi)$ needed to describe the intrinsic shape and orientation of the cluster system, compared to four observable geometric constraints, $(\epsilon, \epsilon^{\rm ICM}, \psi_\epsilon = \psi_\epsilon^{\rm ICM}, e_\parallel^{\rm ICM})$.

6.3. Bayesian 3D Inversion

In our analysis, the cluster model $\boldsymbol{p}$ is defined by seven *fundamental* parameters describing the total matter ellipsoid and one parameter determining the shape of the ICM halo:

$$\boldsymbol{p} = \left(M_{200c}, c_{200c}, q_a, q_b, \vartheta, \phi, \psi, q_a^{\rm ICM}\right). \tag{40}$$

Hence, the overall ellipsoidal model has eight free parameters. On the other hand, 2D lensing constraints reduce to four parameters (Sereno & Umetsu 2011), $(\kappa_{\rm s}, \xi_{\rm s}, \epsilon, \psi_\epsilon)$. A joint X-ray and SZE analysis of the ICM yields two additional constraints (Sereno et al. 2013), namely the ellipticity $\epsilon^{\rm ICM}$ of the ICM in projection and the elongation $e_\parallel^{\rm ICM}$ of the ICM along the line of sight. Accordingly, combined lensing and X-ray/SZE data sets effectively provide six observationally accessible parameters,

$$\boldsymbol{o} = \left(\kappa_{\rm s}, \xi_{\rm s}, \epsilon, \psi_\epsilon, \epsilon^{\rm ICM}, e_\parallel^{\rm ICM}\right). \tag{41}$$

That is, the problem is underconstrained.

To make robust inference on the intrinsic properties of the cluster, we use a forward modeling approach with Bayesian inference for this underconstrained inversion problem (Sereno et al. 2013). The observational parameters $\boldsymbol{o} = \boldsymbol{o}(\boldsymbol{p})$ can be uniquely specified by the intrinsic parameters $\boldsymbol{p}$. The total likelihood function of combined lensing and X-ray/SZE observations can be formally written as (Sereno et al. 2013)

$$\mathcal{L}[\boldsymbol{o}(\boldsymbol{p})] = \mathcal{L}_{\rm GL} \times \mathcal{L}_{\rm ICM} \tag{42}$$

with $\mathcal{L}_{\rm GL}$ the likelihood function of lensing observables and $\mathcal{L}_{\rm ICM}$ that of X-ray/SZE observables.

6.4. Priors

For our base model, we use uninformative priors for the intrinsic parameters $\boldsymbol{p}$. We adopt flat priors of $q_{\rm min} \leqslant q_a \leqslant 1$ and $q_a \leqslant q_b \leqslant 1$ for the intrinsic axis ratios of the matter distribution, where $q_{\rm min}$ is introduced to exclude models with extremely small axis ratios because such configurations would be dynamically unstable and not expected for cluster halos. The probability functions can then be expressed as $P(q_a) = 1/(1 - q_{\rm min})$ for $q_{\rm min} \leqslant q_a \leqslant 1$ and $P(q_b|q_a) = 1/(1 - q_a)^{-1}$ for $q_b \geqslant q_a$. In what follows, we fix $q_{\rm min} = 0.1$ (Oguri et al. 2005; Sereno et al. 2013). Alternatively, we may consider the axis-ratio priors that follow distributions obtained from ΛCDM N-body simulations (Jing & Suto 2002).

For the minor–major axis ratio of the ICM, we use a uniform distribution in the interval $q_a \leqslant q_a^{\rm ICM} \leqslant 1$ (see Section 6.2). The prior of $q_a^{\rm ICM}$, $P(q_a^{\rm ICM}|q_a)$ can then be defined in a similar way to that of q_b. For the orientation angles, we consider a population of randomly oriented halos with $P(\cos\vartheta) = 1$ for $0 \leqslant \cos\vartheta \leqslant 1$ and $P(\phi) = 1/\pi$ for $-\pi/2 \leqslant \phi \leqslant \phi/2$. Finally, we employ uniform priors for the remaining parameters.

7. MULTI-PROBE ANALYSIS OF A1689

Here we apply the Bayesian inversion method outlined in Section 6 to our multiwavelength observations of A1689. The results are discussed in Section 8.

7.1. Weak and Strong Lensing

A full 2D lensing analysis is crucial for comparison with predictions of the properties of aspherical clusters (Oguri et al. 2005). In this work, we have employed free-form methods for both weak- and strong-lensing mass reconstructions (Sections 4 and 5), which provide a pixelated Σ map and its covariance matrix in each regime.

In this subsection, we derive constraints on the projected halo properties (Section 6.1) from lensing data. We model the observed Σ field with a projected ellipsoidal NFW profile (Section 6.1), specified by $(\kappa_{\rm s}, \xi_{\rm s}, \epsilon, \psi_\epsilon)$. Additionally, we include the halo centroid $\boldsymbol{\theta}_{\rm c}$ as parameters to conservatively account for the degree of miscentering.

7.1.1. Weak-lensing Data

The χ^2 function for the Subaru weak-lensing observations is expressed as (Oguri et al. 2005)

$$\chi^2_{\rm WL} = \sum_{m,n=1}^{N_{\rm pix}} \left[\Sigma(\boldsymbol{\theta}_m) - \hat{\Sigma}(\boldsymbol{\theta}_m)\right]\left(C^{-1}\right)_{mn}\left[\Sigma(\boldsymbol{\theta}_n) - \hat{\Sigma}(\boldsymbol{\theta}_n)\right], \tag{43}$$

where $\boldsymbol{\Sigma} = \{\Sigma(\boldsymbol{\theta}_m)\}_{m=1}^{N_{\rm pix}}$ is the mass map from the 2D weak-lensing analysis (Section 4.2), C^{-1} is the inverse of the error covariance matrix, and the hat symbol denotes a modeled quantity. The corresponding likelihood is $\mathcal{L}_{\rm WL}(\kappa_{\rm s}, \xi_{\rm s}, \epsilon, \psi_\epsilon, \boldsymbol{\theta}_{\rm c}) \propto \exp(-\chi^2_{\rm WL}/2)$.

Figure 11 shows the results in terms of the marginalized posterior distribution for the scale convergence, $\kappa_{\rm s} = \Sigma_{\rm s}/\Sigma_{\rm c}$, and the scale radius, $\theta_{\rm s} = \xi_{\rm s}/D_{\rm l}$. Table 5 summarizes marginalized constraints on the individual parameters. In the present study, we employ the robust biweight estimators of Beers et al. (1990) for the central location (mean) and scale (standard deviation) of the marginalized posterior distributions (e.g., Sereno & Umetsu 2011; Umetsu et al. 2014).

7.1.2. Strong-lensing Data

Mass maps derived from strong lensing exhibit a high degree of correlation between adjacent regions. The problem is exacerbated for parametric methods, which model the total mass distribution by a superposition of lens components assuming parametric density profiles. This also persists in free-form modeling (Lubini et al. 2014), albeit to a lesser degree.

The degree of correlation can be examined by an eigenvalue analysis. Let us decompose the C matrix as $C = U\Lambda U^{-1}$, with Λ the diagonal matrix of eigenvalues and U the unitary matrix of eigenvectors. The first few eigenvalues describe the principal modes of variation of the mass model (Lubini et al. 2014; Mohammed et al. 2014). Large eigenvalues correspond to

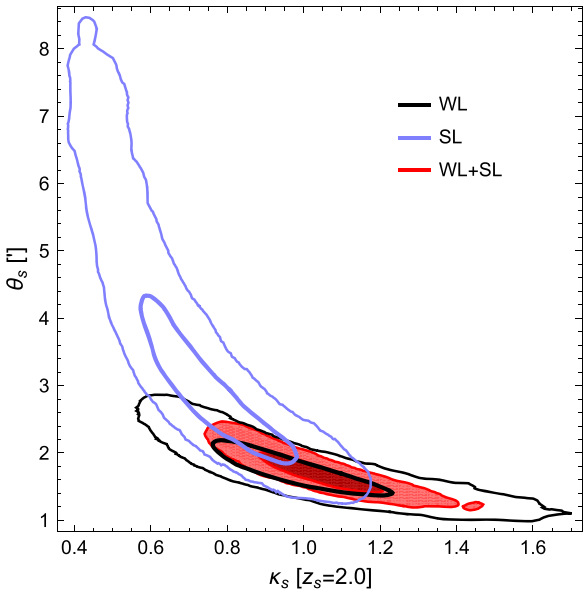

Figure 11. Marginalized posterior distribution for the projected NFW parameters (κ_s, θ_s) obtained from three different lensing data sets (see Table 5), namely weak-lensing-only (black; WL), strong-lensing-only (blue; SL), and combined weak and strong lensing (red shaded; WL+SL). For each case, the contour levels are at $\exp(-2.3/2)$ and $\exp(-11.8/2)$ of the maximum, corresponding to the 1σ and 3σ confidence levels, respectively, for a Gaussian distribution. The scale convergence $\kappa_s = \Sigma_s/\Sigma_c$ is normalized to a fiducial source redshift of $z_s = 2$.

massive pixels, namely, those composing the inner part of the mass distribution that is best constrained by strong lensing. The ordered list of eigenvalues progressively decreases with increasing rank and drops abruptly near the maximum rank, indicating a high degree of correlation (Figure 12).

Here, we employ a regularization approach to conservatively account for the high degree of correlation of the covariance matrix. This was first proposed by Umetsu et al. (2012) for the 1D analysis of strong-lensing mass profiles. If the covariance matrix C is not degenerate, we can construct a χ^2 function for each group of multiple images as

$$\chi^2_{\mathrm{SL},\alpha} = \sum_{m,n}\left[\Sigma_m - \hat{\Sigma}_m\right]\left(C^{-1}\right)_{mn}\left[\hat{\Sigma}_n - \hat{\Sigma}_n\right], \\ = \sum_m \frac{\left[\left(\Sigma_U\right)_m - \left(\hat{\Sigma}_U\right)_m\right]^2}{\Lambda_m}, \quad (44)$$

where Σ_m is the observed Σ value of the mth pixel, $(\Sigma_U)_m = \sum_l U_{ml}\Sigma_l$ is the projection onto the eigenbasis, α runs over the four groups of images (Section 5), and the hat symbol is used to denote a modeled quantity. Each group has its own $\mathbf{\Sigma}$, C, U, and Λ. Here we drop the index α on the right hand side to simplify the notation.

In this approach, we limit ourselves to the principal modes and truncate the summation at N_{max} largest eigenvalues as

$$\chi^2_{\mathrm{SL},\alpha} \approx \sum_{m=1}^{N_{\mathrm{max}}} \frac{\left[\left(\Sigma_U\right)_m - \left(\hat{\Sigma}_U\right)_m\right]^2}{\Lambda_m}. \quad (45)$$

A natural choice for N_{max} is the number of observational constraints. We thus set $N_{\mathrm{max}} = 2N_{\mathrm{im}}$ with N_{im} the number of multiple images used. The total χ^2 is given by

$$\chi^2_{\mathrm{SL}} = \sum_\alpha \chi^2_{\mathrm{SL},\alpha}. \quad (46)$$

We find that the eigenvalues before the drop range approximately between the minimum κ^2 value in the ensemble-averaged pixelated model and that found from the whole ensemble of models generated by PixeLens (Section 5.2). This is demonstrated in Figure 12. The $2N_{\mathrm{im}}$th eigenvalue lies approximately in the middle of this range and sets a conservative scale. We checked the reliability and performance of this regularization method using analytical models.

Some multiple image systems share very similar configurations (e.g., systems 1 and 2). Such a redundancy is valuable for determining cosmological parameters (Lubini et al. 2014), or for improving the sensitivity to local substructures. Assigning a full weight to systems having similar configurations would inflate the relative contribution of strong lensing with respect to weak lensing. To avoid this, we multiply χ^2_{SL} by a weighting factor w_{SL}, defined as the inverse of the geometrical average of the number of such redundant image systems. We find $w_{\mathrm{SL}} = 2/3$ for our analysis. The likelihood is then defined as $\mathcal{L}_{\mathrm{SL}}(\kappa_s, \xi_s, \epsilon, \psi_\epsilon, \boldsymbol{\theta}_c) \propto \exp(-w_{\mathrm{SL}}\chi^2_{\mathrm{SL}}/2)$.

The results are summarized in Table 5 and Figure 11.

7.1.3. Combining Weak and Strong Lensing

We now combine the weak- and strong-lensing likelihoods constructed in Sections 7.1.1 and 7.1.2, respectively, to jointly constrain the projected NFW parameters. The likelihood function $\mathcal{L}_{\mathrm{GL}}$ for the combined weak plus strong lensing data can be written as (Sereno & Umetsu 2011)

$$\mathcal{L}_{\mathrm{GL}} = \mathcal{L}_{\mathrm{WL}} \times \mathcal{L}_{\mathrm{SL}} \propto \exp\left[-\left(\chi^2_{\mathrm{WL}} + w_{\mathrm{SL}}\chi^2_{\mathrm{SL}}\right)/2\right], \quad (47)$$

where χ^2_{WL} and χ^2_{SL} are defined by Equations (43) and (46), respectively.

Figure 11 shows that the scale radius (θ_s) and the scale convergence (κ_s) are highly degenerate and anti-correlated. In particular, the scale radius is poorly constrained by strong lensing alone because of the limited coverage of multiple images, $\theta \lesssim 1\farcm1$ (Section 5.2). The allowed range of θ_s lies well outside the region where the multiple images are observed. Thus, the inference of parameters by strong lensing requires an extrapolation well beyond the observed region. For this reason, in the present study, we do not consider strong-lensing-only triaxial modeling (see Table 7). On the other hand, since the posterior distributions from the independent weak-lensing and strong-lensing analyses are compatible, combining weak lensing with strong lensing provides improved parameter constraints (Table 5).

7.2. Combined X-Ray Plus SZE Analysis

With a known halo geometry (e.g., sphericity) and under the ideal gas assumption, the thermodynamic quantities of the ICM are overconstrained by X-ray and SZE data. This is because the thermal pressure can be independently determined from thermal SZE data and X-ray spectroscopy/imaging data. We can therefore relax the assumption of spherical symmetry to solve for the elongation of the ICM distribution (Sereno et al. 2012). Combining gravitational lensing and X-ray/SZE observations with minimal geometric assumptions (Section 6.2) allows us to break the degeneracy between mass and elongation

Table 5
Parameters of the Projected NFW Model Constrained From Lensing Observations

Data[a]	κ_s[b]	ξ_s[c] ($'$)	ϵ[d]	ψ_ϵ[e] (deg)	$\boldsymbol{\theta}_c$[f] ($''$)
WL	0.97 ± 0.16	1.74 ± 0.27	0.29 ± 0.07	14.2 ± 8.4	−1.2 ± 3.0, 4.9 ± 4.1
SL	0.73 ± 0.14	3.00 ± 0.90	0.27 ± 0.09	13.0 ± 9.8	−0.8±1.9, −4.8 ± 2.3
GL	1.03 ± 0.11	1.70 ± 0.20	0.29 ± 0.05	11.4 ± 4.9	0.0 ± 1.3, −1.9 ± 1.4

Note.
[a] WL: weak lensing shear and magnification; SL: strong lensing; GL: combined strong lensing, weak-lensing shear and magnification.
[b] Scale convergence, $\kappa_s = \Sigma_s/\Sigma_c$, normalized to a reference source redshift of $z_s = 2$.
[c] Projected scale radius of the elliptical NFW model measured along the major axis.
[d] Projected mass ellipticity, $\epsilon = 1 - q_\perp$, with $q_\perp$ the projected minor–major axis ratio.
[e] Position angle of the major axis measured east of north.
[f] Halo centroid position relative to the BCG position.

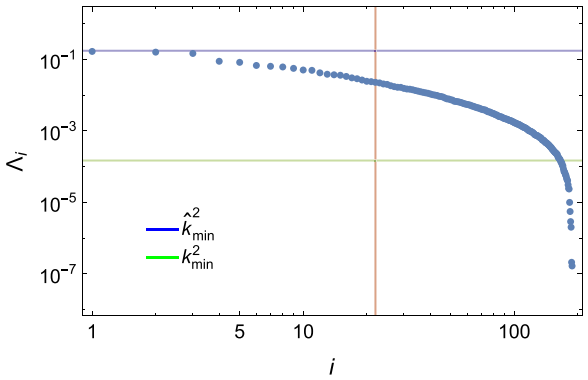

Figure 12. Ordered eigenvalues Λ of the covariance matrix for the PixeLens mass reconstruction. The vertical red line indicates the maximum rank considered for our analysis, $N_{\rm max} = 2N_{\rm im}$, i.e., the number of observational constraints on the image position. The blue horizontal line shows the minimum κ^2 value found in the ensemble-averaged pixelated model. The green horizontal line shows the minimum κ^2 value from the entire statistical ensemble of models generated by PixeLens. The results are shown for the covariance matrix as constrained by the systems 1, 5, and 11.

for the total matter distribution (Sereno et al. 2013). Such a multi-probe approach based on lensing and X-ray/SZE data is free from the assumption of HSE, compared to the lensing plus X-ray analysis, which relies on equilibrium assumptions between the gravitational potential and pressure components (see Limousin et al. 2013).

In our multi-probe approach, the ICM distribution is modeled with an ellipsoidal parametric profile which can fit X-ray surface-brightness (S_X) and temperature (T_X) distributions. Comparison with the SZE amplitude then determines the elongation $e_\parallel^{\rm ICM}$ For example, for an isothermal plasma (De Filippis et al. 2005), we have

$$1/e_\parallel^{\rm ICM} \propto D_1 \frac{S_X}{\Delta T_{\rm SZE}^2}\frac{T_X^2}{\Lambda_X} \tag{48}$$

with $\Delta T_{\rm SZE}$ the SZE temperature decrement and Λ_X the X-ray cooling function of the ICM. In this work, we rely on the X-ray data to constrain the ICM morphology in projection space; we use aperture-integrated constraints on the SZE signal (Table 6) to determine the line of sight elongation $e_\parallel^{\rm ICM}$.

Our X-ray data are taken from Sereno et al. (2012), who performed an X-ray analysis on *Chandra* and *XMM-Newton* observations. Here we briefly summarize essential results

Table 6
Integrated Comptonization Y Parameter Measured Interior to a Cylinder of Radius r

Instrument	r ($'$)	$Y(<r)$ (10^{-10} sr)
BIMA/OVRO	1.5	1.00 ± 0.28
BIMA/OVRO	3.0	2.64 ± 0.97
SZA	1.5	1.11 ± 0.10
SZA	3.0	2.83 ± 0.42
SZA	4.5	4.33 ± 0.81
SZA	6.0	5.50 ± 1.18

needed for this study. For details, we refer to Sereno et al. (2012). Sereno et al. (2012) showed that exposure corrected and point-source removed *Chandra* X-ray images in the 0.7–2.0 keV band are well described by concentric ellipses with ellipticity $\epsilon^X = 0.15 \pm 0.03$ and orientation angle $\psi_\epsilon^X = (12 \pm 3)$ degrees measured east of north. Following Sereno et al. (2012) and Sereno et al. (2013), we model the 3D electron density in the intrinsic coordinate system with the following parametric form (Vikhlinin et al. 2006; Ettori et al. 2009):

$$n_e = n_0 \left[1 + \left(\frac{R}{r_c}\right)^2\right]^{-3\beta/2} \left[1 + \left(\frac{R}{r_t}\right)^2\right]^{-\gamma/3}, \tag{49}$$

where n_0 is the central electron density, r_c is the ellipsoidal coreradius, $r_t (> r_c)$ is the ellipsoidal truncation radius, β is the slope in the intermediate density regions, and γ is the outer slope. The 3D gas density is parametrized as (Sereno et al. 2013)

$$T = \frac{T_0}{\left[1 + \left(R/r_T\right)^2\right]^{0.45}}, \tag{50}$$

where T_0 is the central gas temperature, and r_T describes a temperature decline at large cluster radii. The parametrizations of Equations (49) and (50) were motivated by the absence of cool-core features in our data. For further justification, see Section 5 of Sereno et al. (2012).

The thermal SZE provides a complementary measure of the thermal energy content in a cluster. In this study, we perform a self-consistent multi-scale analysis of high-significance 30 GHz interferometric SZE observations of A1689 obtained with the

Berkeley-Illinois-Maryland Array (BIMA), the Owens Valley Radio Observatory (OVRO), and the SZA. The BIMA and OVRO observations of A1689 are presented in LaRoque et al. (2006), while the SZA observations of A1689 are presented in Gralla et al. (2011). Owing to the different scales probed by the instruments, we fit the OVRO/BIMA and SZA data separately using the spherical Arnaud et al. (2010) pressure profile. This profile is an adaptation of the generalized NFW pressure profile first proposed by Nagai et al. (2007), and first fitted to SZE observations in Mroczkowski et al. (2009). A joint fit to the OVRO, BIMA, and SZA data was also performed to determine the best-fit SZE centroid reported in Table 1.

As in Mroczkowski et al. (2009), a model for the cluster and contaminating radio sources is computed in the image plane, then Fourier transformed for comparison to the interferometric data. The best-fit model and 1σ confidence intervals are determined using a Markov chain Monte Carlo (MCMC) procedure. The OVRO and BIMA data measure radial scales from $0\farcm5$–$4'$, while the SZA data probe radial scales from $1'$–$6'$. Bonamente et al. (2012) showed that the adoption of the Arnaud et al. (2010) profile versus other non-isothermal pressure profiles accurate out to r_{500c} does not significantly impact the parameters derived from the fits when the radii for which the results are computed are at scales accessible to the instruments.

A summary of the SZE data used is given in Table 6. The integrated Comptonization parameter $Y(<r)$ interior to a cylinder of radius r is written in terms of the electron density and temperature profiles (Equations (49) and (50)) as

$$Y = \frac{\sigma_{\rm T} k_{\rm B}}{m_{\rm e} c^2} \int_{\Omega_r} d\Omega \int dl\, n_{\rm e} T \tag{51}$$

with $\sigma_{\rm T}$ the Thomson cross section, $k_{\rm B}$ the Boltzmann constant, $m_{\rm e}$ the electron mass, and c the speed of light in vacuum; Ω_r is the solid angle of the integration aperture.

The model profiles given by Equations (49)–(51) are then compared with combined X-ray surface brightness ($S_{\rm X}$), X-ray spectroscopic temperature ($T_{\rm X}$), and thermal SZE decrement (Y) observations. Briefly summarizing, the X-ray surface brightness profile $\{S_{{\rm X},i}\}_{i=1}^{N_{\rm S}}$ observed by *Chandra* was extracted from $N_{\rm S} = 68$ elliptical annuli out to an elliptical radius of $\xi = 900\,{\rm kpc}\, h_{70}^{-1}$ ($\sim 5'$), and the *XMM-Newton* temperature profile $\{T_{{\rm X},i}\}_{i=1}^{N_{\rm T}}$ was measured in $N_{\rm T} = 5$ elliptical annual bins out to $\xi = 900\,{\rm kpc}\, h_{70}^{-1}$ (Sereno et al. 2012). Thanks to the improved SZE analysis, the Y parameter is measured at several apertures from BIMA/OVRO and SZA data as summarized in Table 6. We find good consistency between the BIMA/OVRO and SZA results at $r = 1\farcm5$ and $3'$ where these independent data overlap. At an integration radius of $r = 3'$, our results are also in excellent agreement with $Y(<3') = (2.5 \pm 0.6) \times 10^{-10}$ sr from 94 GHz interferometric observations with the 7-element AMiBA (Umetsu et al. 2009, their Table 5).

The X-ray part of the χ^2 function can be written as (Sereno et al. 2012)

$$\chi^2_{\rm X} = \sum_{i=1}^{N_{\rm S}} \left(\frac{S_{{\rm X},i} - \hat{S}_{{\rm X},i}}{\sigma_{{\rm S},i}} \right)^2 + \sum_{i=1}^{N_{\rm T}} \left(\frac{T_{{\rm X},i} - \hat{T}_{{\rm X},i}}{\sigma_{{\rm T},i}} \right)^2 \tag{52}$$

with $(\hat{S}_{\rm X}, \hat{T}_{\rm X})$ model predictions for the corresponding X-ray observables and $(\sigma_{\rm S}, \sigma_{\rm T})$ their corresponding errors.

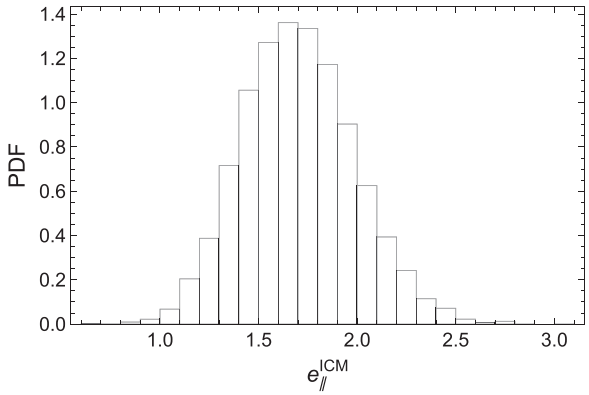

Figure 13. Marginalized posterior probability distribution of the elongation $e_\parallel^{\rm ICM}$ as derived from the combined X-ray plus SZE analysis (Section 7.2).

The χ^2 function for the SZE observations is written as

$$\chi^2_{\rm SZE} = \sum_j \sum_i \left(\frac{\Delta Y_{ji} - \Delta \hat{Y}_i}{\sigma_{\Delta,ji}} \right)^2, \tag{53}$$

where ΔY_{ji} is the differential Y parameter for the jth instrument (BIMA/OVRO or SZA) in the ith annular ring, $\Delta Y_{ji} \equiv Y_j(<r_{i+1}) - Y_j(<r_i)$, and $\sigma_{\Delta,ji}$ is its 1σ uncertainty. The Y values are sampled at every $1\farcm5$ (Table 6), which is sufficiently larger than the synthesized beam. Hence, differential ΔY measurements in adjacent annuli are approximately uncorrelated given the annulus size considered.

A combined analysis of the X-ray and SZE data is performed using the combined function $\chi^2 = \chi^2_{\rm X} + \chi^2_{\rm SZE}$. The parameter space is explored using an MCMC approach as described in Sereno et al. (2013). Since parameter constraints on the $n_{\rm e}$ and T models are dominated by the *Chandra* surface brightness and *XMM-Newton* temperature data, respectively, we find our results are fully consistent with those of Sereno et al. (2012) based on the same X-ray data. The best-fit central temperature ($T_0 = 9.8 \pm 0.2$ keV, Sereno et al. 2012) is in good agreement with the *Suzaku* X-ray results of Kawaharada et al. (2010). On the other hand, using the improved SZE data, we obtain tighter constraints on the elongation $e_\parallel^{\rm ICM}$. The resulting posterior distribution of $e_\parallel^{\rm ICM}$ is shown in Figure 13. The posterior mean and standard deviation are $e_\parallel^{\rm ICM} = 1.70 \pm 0.29$.

7.3. Multi-probe Deprojection

Here we perform joint likelihood analyses of combined lensing and X-ray/SZE data, using different combinations of lensing data sets (Section 7.1).

The likelihood $\mathcal{L}_{\rm ICM}$ of the X-ray/SZE data is written in terms of two observable ICM parameters (Section 6.3), namely, the ellipticity $\epsilon^{\rm ICM}$ and line of sight elongation $e_\parallel^{\rm ICM}$ of the ICM. Following Sereno et al. (2012, 2013), we include a nuisance parameter $\Delta e_\parallel^{\rm sys}$ that quantifies the additional uncertainty on $e_\parallel^{\rm ICM}$, accounting for potential calibration systematics in the X-ray/SZE measurements. It is assumed to follow a normal distribution with zero mean and standard deviation $\sigma_\parallel^{\rm sys} = 0.07$. Since the systematic uncertainty is quite small compared to the width of the marginalized posterior

Table 7
Intrinsic Parameters of the Total Matter Distribution Obtained Using Different Data Sets and Different Priors

Data[a]	Prior	M_{200c} ($10^{15}\,M_\odot h^{-1}$)	c_{200c}	q_a	q_b	$\cos\vartheta$[b]
WL	Spherical	1.31 ± 0.11	8.87 ± 1.11	1	1	⋯
WL	Flat	1.28 ± 0.26	10.70 ± 2.85	0.39 ± 0.18	0.77 ± 0.15	0.54 ± 0.29
WL	*N*-body	1.22 ± 0.23	9.15 ± 1.77	0.47 ± 0.08	0.66 ± 0.12	0.60 ± 0.30
SL	Spherical	1.79 ± 0.31	8.69 ± 1.26	1	1	⋯
GL	Spherical	1.32 ± 0.09	10.10 ± 0.82	1	1	⋯
GL	Flat	1.49 ± 0.25	10.30 ± 2.52	0.45 ± 0.20	0.77 ± 0.14	0.47 ± 0.29
GL	*N*-body	1.41 ± 0.19	9.65 ± 1.54	0.47 ± 0.08	0.66 ± 0.12	0.60 ± 0.29
WL + X/SZ	Flat	1.21 ± 0.19	7.91 ± 1.41	0.39 ± 0.16	0.56 ± 0.20	0.93 ± 0.06
WL + X/SZ	*N*-body	1.16 ± 0.17	7.42 ± 1.21	0.40 ± 0.08	0.52 ± 0.12	0.94 ± 0.05
GL + X/SZ	Flat	1.24 ± 0.16	8.36 ± 1.27	0.39 ± 0.15	0.57 ± 0.19	0.93 ± 0.06
GL + X/SZ	*N*-body	1.20 ± 0.13	7.89 ± 0.96	0.40 ± 0.08	0.52 ± 0.12	0.94 ± 0.05

Note. Intrinsic parameters of the total matter distribution of A1689 derived from a triaxial analysis of multiwavelength data sets, using spherical, flat, and *N*-body priors on the distribution of axis ratios (q_a, q_b).
[a] WL: weak-lensing shear and magnification; SL: strong lensing; GL: combined strong lensing, weak-lensing shear and magnification; X/SZ: combined X-ray and SZE measurements.
[b] Cosine of the angle between the major axis and the line of sight.

distribution $P(e_\parallel^{\rm ICM})$ (Figure 13), the impact on the final results is minor.

The X-ray/SZE part of the likelihood $\mathcal{L}_{\rm ICM}(\epsilon^{\rm ICM}, e_\parallel^{\rm ICM}; \Delta e_\parallel^{\rm sys})$ is written as (Sereno et al. 2013)

$$\mathcal{L}_{\rm ICM} = \frac{1}{\sqrt{2\pi}\,\sigma_{\epsilon,{\rm X}}} \exp\left[-\frac{\left(\epsilon^X - \epsilon^{\rm ICM}\right)^2}{2\sigma_{\epsilon,{\rm X}}^2}\right] \times P\left(e_\parallel^{\rm ICM} - \Delta e_\parallel^{\rm sys}\right) \times \frac{1}{\sqrt{2\pi}\,\sigma_\parallel^{\rm sys}} \exp\left[-\frac{1}{2}\left(\frac{\Delta e_\parallel^{\rm sys}}{\sigma_\parallel^{\rm sys}}\right)^2\right], \quad (54)$$

where ϵ^X and $\sigma_{\epsilon,{\rm X}}$ are the measured value of the ICM ellipticity and its uncertainty, respectively (Section 7.2).

To perform a joint analysis with the X-ray/SZE data, we consider three different likelihood functions for the lensing part, namely, $\mathcal{L}_{\rm WL}$, $\mathcal{L}_{\rm SL}$, and $\mathcal{L}_{\rm GL} = \mathcal{L}_{\rm WL}\mathcal{L}_{\rm SL}$, which are all functions of the projected NFW parameters $\kappa_{\rm s}$, $\xi_{\rm s}$, ϵ, ψ_ϵ, and $\boldsymbol{\theta}_{\rm c}$. Following Sereno et al. (2013), we exploit constraints from the X-ray analysis about the gas centroid $\boldsymbol{\theta}_{\rm c}^X$ and position angle ψ_ϵ^X (Section 7.2), which are used as priors for the centroid $\boldsymbol{\theta}_{\rm c}$ and position angle ψ_ϵ of the underlying halo (see Section 4 of Sereno et al. 2013). These priors are consistent with the geometric assumptions we have made in Section 6.2.

For our base model, we use flat priors for the intrinsic axis ratios of the underlying halo (Section 6.4). We also consider an alternative prior distribution predicted by cosmological *N*-body simulations of Jing & Suto (2002). For details, we refer to Sereno & Umetsu (2011) and Sereno et al. (2013).

8. RESULTS AND DISCUSSIONS

The resulting constraints on the intrinsic parameters for the underlying halo (M_{200c}, c_{200c}, q_a, q_b, $\cos\vartheta$) are given in Table 7, for different combinations of data sets and three different priors on the axis-ratio distribution: (1) spherical prior ($q_a = q_b = 1$); (2) flat distribution of axis ratios and random distribution of halo orientations (Section 6.4); (3) *N*-body ΛCDM predictions (Jing & Suto 2002). The baseline results for the combined weak/strong-lensing and X-ray/SZE analysis obtained with flat priors are shown in Figure 14. Table 8 gives a summary of our baseline constraints on the intrinsic axis ratios of the ICM halo, ($q_a^{\rm ICM}$, $q_b^{\rm ICM}$), and on the ICM-to-matter ratio of halo eccentricities, $e^{\rm ICM}/e$. Table 9 lists the published (M_{200c}, c_{200c}) measurements for A1689 based on the combination of both weak and strong lensing. For previous compilations, see Comerford & Natarajan (2007, their Table A1) Limousin et al. (2007, their Table 4), Umetsu & Broadhurst (2008, their Table 5), Corless et al. (2009, their Table 4), and Coe et al. (2010, their Table 2).

8.1. Mass and Concentration

8.1.1. Spherical Modeling

The degree of concentration of A1689 has been a subject of controversy. Here we first compare the results obtained assuming a spherical NFW halo (Table 7) to those of previous work. Our full 2D weak-lensing analysis based on Subaru $BVR_{\rm C}i'z'$ data yields a projected concentration of $c_{200c} = 8.9 \pm 1.1$ ($c_{\rm vir} = 11.2 \pm 1.4$) at $M_{200c} = (1.31 \pm 0.11) \times 10^{15}\,M_\odot\,h^{-1}$. This is in excellent agreement with, and improved from, our earlier weak-lensing work: $c_{200c} = 10.7^{+4.5}_{-2.7}$ (Umetsu & Broadhurst 2008) and $c_{200c} = 10.2^{+2.5}_{-2.0}$ (Umetsu et al. 2011b), both of which are based on the joint analysis of shear and magnification data from Subaru *Vi′* imaging.[19] This accurate agreement comes in spite of using different data reduction procedures and mass reconstruction methods (Sections 2–4).

Combining weak and strong lensing reduces the uncertainties on the concentration. The *HST* strong-lensing data alone also favor a high degree of projected concentration, $c_{200c} = 8.69 \pm 1.26$, but with a somewhat higher halo mass, $M_{200c} = (1.79 \pm 0.31) \times 10^{15}\,M_\odot\,h^{-1}$. The combined weak

[19] Umetsu & Broadhurst (2008) derived a $\kappa(\boldsymbol{\theta})$ map for the cluster using an entropy-regularized maximum-likelihood combination of 2D shear and magnification maps. Umetsu et al. (2011b) derived a $\kappa(\theta)$ profile from a joint likelihood analysis of azimuthally averaged shear and magnification measurements.

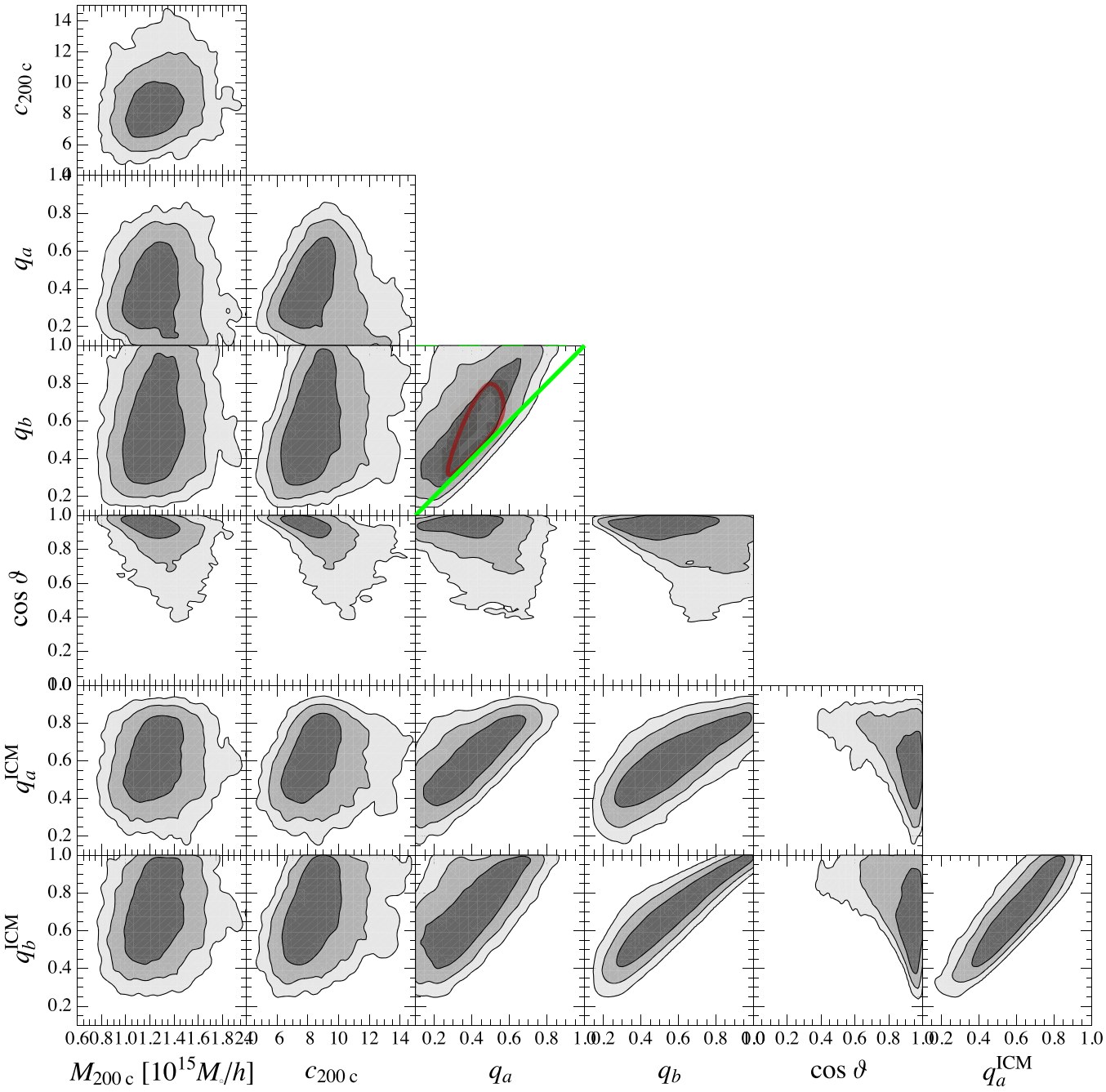

Figure 14. Marginalized posterior distributions for the intrinsic parameters of the triaxial cluster model obtained from a joint analysis of the weak/strong-lensing and X-ray/SZE data. In each panel, the contours levels are shown at $\exp(-2.3/2)$, $\exp(-6.17/2)$, and $\exp(-11.8/2)$ of the maximum, corresponding to the 1σ, 2σ, and 3σ confidence levels, respectively, for a Gaussian distribution. In the q_b vs. q_a plane, the green solid (diagonal) and dashed (horizontal) lines represent prolate ($q_a = q_b$) and oblate ($q_b = 1$) configurations, respectively, and the thick red line shows the 1σ contour for the axis-ratio distribution in ΛCDM N-body simulations of Jing & Suto (2002).

Table 8
Intrinsic Shapes of the ICM Distribution

Priors	$q_a^{\rm ICM}$	$q_b^{\rm ICM}$	$e^{\rm ICM}/e$
Flat	0.60 ± 0.14	0.70 ± 0.16	0.87 ± 0.07
N-body	0.58 ± 0.10	0.65 ± 0.11	0.89 ± 0.06

Note. Constraints on the intrinsic axis ratios ($q_a^{\rm ICM}$, $q_b^{\rm ICM}$) of the ICM distribution and the relation with the total matter distribution ($e^{\rm ICM}/e$), obtained from the full triaxial analysis of combined weak/strong-lensing and X-ray/SZE data sets (Section 7). $q_b^{\rm ICM}$ and $e^{\rm ICM}/e$ are derived parameters.

and strong lensing data yield $c_{200c} = 10.10 \pm 0.82$ at $M_{200c} = (1.32 \pm 0.09) \times 10^{15}\,M_\odot\,h^{-1}$, corresponding to the Einstein radius of $\theta_{\rm Ein} = 52^{+6}_{-7}{}''$ at $z_s = 2$. Our analysis thus reproduces the correct size of the observed Einstein radius (Table 1). These results are in good agreement with those of Umetsu & Broadhurst (2008) and Coe et al. (2010; Table 9), in spite of using completely independent approaches to strong lens modeling (Section 5). Most recent weak-and-strong lensing studies of A1689 appear to converge toward $c_{200c} \sim 9$–10 with a typical measurement uncertainty of 10%

Table 9
Published Mass and Concentration Measurements of A1689 From Combined Weak and Strong Lensing

Author	M_{200c} ($10^{15}\,M_\odot h^{-1}$)	c_{200c}	Prior[a]	External data[b]
Spherical modeling				
Broadhurst et al. (2005a)	1.20 ± 0.13	$10.9^{+1.1}_{-0.9}$	Spherical	⋯
Halkola et al. (2006)	1.58 ± 0.14	7.6 ± 0.5	Spherical	⋯
Umetsu & Broadhurst (2008) [c]	1.30 ± 0.11	$10.1^{+0.8}_{-0.7} \pm 2.2$	Spherical	⋯
Coe et al. (2010)	$1.3^{+0.3}_{-0.2}$	9.2 ± 1.2	Spherical	⋯
This work	1.32 ± 0.09	10.10 ± 0.82	Spherical	⋯
Triaxial modeling				
Oguri et al. (2005) [d]	$1.14^{+0.26}_{-0.51}$	$13.6^{+1.8}_{-10.5}$	Flat	⋯
Sereno & Umetsu (2011)	1.07 ± 0.23	9.3 ± 2.0	Flat	⋯
This work	1.49 ± 0.25	10.30 ± 2.52	Flat	⋯
With line of sight information				
Corless et al. (2009)	0.83 ± 0.16	12.2 ± 6.7	Flat + $\cos\vartheta$	⋯
Sereno & Umetsu (2011)	0.99 ± 0.17	7.7 ± 1.1	Flat + $\cos\vartheta$	⋯
Morandi et al. (2011)	1.81 ± 0.06	5.71 ± 0.47	Flat	X-ray
Sereno et al. (2013)	0.93 ± 0.12	7.8 ± 0.7	Flat	X-ray/SZE
This work	1.24 ± 0.16	8.36 ± 1.27	Flat	X-ray/SZE

Notes. The results based on the combination of both weak and strong lensing are summarized (converted from quoted values assuming an NFW density profile if necessary).
[a] Spherical: spherical prior on the intrinsic axis-ratios; Flat: flat prior on the intrinsic axis ratios; $\cos\vartheta$: ΛCDM-like prior on the biased orientation of strong-lensing cluster halos (Corless et al. 2009).
[b] External data sets used in combination with lensing for constraining the line of sight elongation.
[c] The weak-lensing mass map of Umetsu & Broadhurst (2008) was used in the triaxial analyses by Oguri et al. (2005), Sereno & Umetsu (2011), Morandi et al. (2011), and Sereno et al. (2013).
[d] NFW-equivalent of triaxial model parameters from Oguri et al. (2005).

(Table 9; with the spherical prior), thanks to the advanced analysis methods and greatly improved quality of data.

8.1.2. Triaxial Modeling

Including triaxiality weakens parameter constraints from lensing data (Oguri et al. 2005; Corless et al. 2009), compared to those derived assuming spherical symmetry. The parameter constraints become more degenerate and less restrictive because of the lack of information of the halo elongation along the line of sight (Table 7). These trends are also found in the posterior distributions from our data (Tables 7 and 9).

Now we consider the results from full triaxial analyses combining lensing with X-ray/SZE data. Table 7 shows that our posterior inference of the intrinsic parameters is insensitive to the assumed choice of priors ("Flat" or "*N*-body") when the line of sight information from X-ray/SZE data is combined with lensing, suggesting that the posterior constraints are dominated by the likelihood (i.e., information from data) rather than the prior (Sereno et al. 2013). Whatever the assumptions regarding the axis ratios, we find the posteriors (Table 7) to be statistically compatible with the predicted distribution $c(M)$ for the full population of halos in ΛCDM cosmological simulations (Bhattacharya et al. 2013; Meneghetti et al. 2014; Diemer & Kravtsov 2015).[20] This is demonstrated in Figure 15 for the weak-lensing plus X-ray/SZE analysis and for the weak/strong-lensing plus X-ray/SZE analysis, both based on the uninformative priors. Here we adopt the median *c*–*M* relation obtained by Diemer & Kravtsov (2015) as a reference model for comparison.

A1689 appears to be a high mass cluster of $M_{200c} \sim 10^{15}\,M_\odot\,h^{-1}$ in the high-concentration tail of the predicted $c(M)$ distribution (Figure 15). The posterior tail at lower concentrations of A1689 is only $\gtrsim 1\sigma$ away from the predicted median concentration ($\log_{10}\tilde{c}_{200c} \simeq 0.58 \pm 0.16$; Figure 15). Our results are also in agreement with those obtained by a multi-probe analysis of Sereno et al. (2013; see Table 9), who developed the triaxial inversion algorithm used in this work.

The halo concentration and orientation are strongly correlated (Sereno & Umetsu 2011; Sereno et al. 2013). For the posterior range $0° \leqslant \vartheta \leqslant 5°$ assuming a nearly perfect alignment between the halo major axis and the line of sight, we find $c_{200c} = 7.4 \pm 1.0$ (6.7 ± 1.1) from weak/strong lensing (weak lensing) combined with the X-ray/SZE data.

8.2. Intrinsic Shape and Orientation of A1689

We have obtained evidence for a triaxial mass distribution of A1689. The projected mass distribution derived from weak-lensing shear and magnification reveals a north–south elongation ($\psi_\epsilon = 14.^\circ 2 \pm 8.^\circ 4$ east of north, see Table 5 and Figure 1). We have determined the ellipticity of the projected mass distribution to be $\epsilon = 0.29 \pm 0.07$ (Table 5), which is typical for the population of collisionless CDM halos (Jing & Suto 2002) but slightly rounder than the standard CDM prediction for the mean halo ellipticity, $\langle\epsilon\rangle \sim 0.4$ (Oguri et al. 2010). The matter ellipticity is detected at the 4σ level from weak lensing alone, thanks to the greatly improved quality of Subaru data. Our free-form reconstruction from *HST* strong lensing gives a consistent estimate of $\epsilon = 0.27 \pm 0.09$.

[20] The theoretical predictions from Bhattacharya et al. (2013) and Diemer & Kravtsov (2015) are based on DM-only simulations, and those from Meneghetti et al. (2014) are based on nonradiative simulations of DM and baryons.

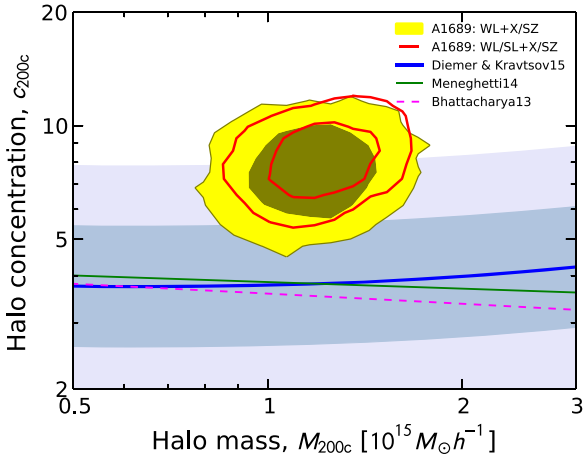

Figure 15. Marginalized constraints on the ellipsoidal NFW model parameters (M_{200c}, c_{200c}) for A1689 compared to the c–M relations predicted for the full population of halos in ΛCDM cosmological simulations (Bhattacharya et al. 2013; Meneghetti et al. 2014; Diemer & Kravtsov 2015). The yellow shaded regions show the results from weak lensing combined with X-ray/SZE data. The red contours are from the full analysis of weak/strong-lensing and X-ray/SZE data. For each case, the contours show the 68.3% and 95.4% confidence levels in the c–M plane. The light blue areas show the 1σ and 2σ ranges of intrinsic halo concentrations (with a 68% scatter of 0.16 dex), respectively, as obtained by Diemer & Kravtsov (2015). All model predictions are evaluated at the cluster redshift $z_l = 0.183$. Overall, the inferred range of c_{200c} is high but overlaps with the $\sim 2\sigma$ tail of the predicted distribution for high-mass cluster halos.

The ICM and matter distributions are co-aligned in projection ($\psi_\epsilon^X = 12° \pm 3°$) but with different ellipticities ($\epsilon_X = 0.15 \pm 0.03$), which is consistent with the geometric assumptions made (Section 6.2).

When combined with X-ray/SZE observations, our lensing data favor a triaxial geometry of the matter distribution with minor–major axis ratio $q_a \sim 0.4$ and major axis closely aligned with the line of sight ($\vartheta = 22° \pm 10°$, Table 7). These results are robust against the choice of priors and combinations of lensing data sets. Despite that the intermediate–major axis ratio q_b is less constrained, the data prefer prolate ($q_a = q_b$) over oblate ($q_b = 1$) configurations. A spherical configuration for A1689 is strongly ruled out. Overall, triaxial configurations fit the combined lensing and X-ray/SZE data much better than axially symmetric halos do (Sereno et al. 2013).

Our analysis shows that A1689 is elongated along the line of sight, as found by previous studies (Sereno et al. 2012, 2013; Limousin et al. 2013). From the posterior samples, we find $e_\parallel = 1.19 \pm 0.37$ (1.20 ± 0.34) and $e_\parallel^{\rm ICM} = 1.22 \pm 0.24$ (1.24 ± 0.25), as constrained by the combined weak/strong-lensing (weak lensing) and X-ray/SZE data sets. Such biased orientations are favored, although the intrinsic orientations are a priori assumed to be random. The a priori probability of a randomly oriented halo to have $\vartheta < 45°$ is ∼29% (Sereno et al. 2013). The a posteriori probability of such a configuration is found to be 96% (99%) assuming a flat (N-body-like) distribution of axis ratios. We emphasize that the use of X-ray plus SZE data is essential for obtaining data-driven constraints on the line of sight elongation. To break parameter degeneracies in a lensing-only triaxial analysis, one would have to assume informative priors on the halo shape and orientation (Corless et al. 2009; Sereno & Umetsu 2011).

We find that the ICM is mildly triaxial with $q_a^{\rm ICM} \sim 0.6$ and $q_b^{\rm ICM} \sim 0.7$ (Table 8). The ratio of ICM to matter eccentricities is $e^{\rm ICM}/e = 0.87 \pm 0.07$ (Table 8), supporting the theoretical assumption we have made that the shape of the gas distribution is rounder than the underlying matter (Section 6.2). On the other hand, we find that the gas distribution is more elongated than the gravitational potential ($e^{\rm ICM}/e \gtrsim 0.7$, Lee & Suto 2003), suggesting a deviation from HSE. These results are again insensitive to the choice of the priors. The inferred values of $q_a^{\rm ICM}$ and $q_b^{\rm ICM}$ are somewhat lower (more elongated) than, but consistent within errors with, the results of Sereno et al. (2012, 2013) based on the same X-ray data. The difference is mainly due to the improved, self-consistent SZE analysis.

Table 10
Ellipsoidal and Spherically Enclosed Mass Estimates For A1689

Overdensity[a]	Ellipsoidal[b]		Spherically Enclosed[c]	
Δ	R_Δ	$M(<R_\Delta)$	r_Δ	$M_{\rm sph}(<r_\Delta)$
500	1.89 ± 0.46	0.97 ± 0.13	1.08 ± 0.06	0.88 ± 0.13
200	2.79 ± 0.69	1.24 ± 0.16	1.60 ± 0.16	1.15 ± 0.16

Notes. The overdensity radii are given in units of Mpc h^{-1}. The enclosed masses are in units of $10^{15} M_\odot h^{-1}$.
[a] Mean interior overdensity with respect to the critical density ρ_c for closure of the universe at $z = 0.183$.
[b] Ellipsoidal overdensity radius R_Δ and total mass enclosed within R_Δ.
[c] Spherical overdensity radius r_Δ and spherically enclosed total mass within r_Δ.

8.3. Gas Mass Fraction

We compute the ratio of spherically enclosed gas mass $M_{\rm sph,gas}(<r)$ to total mass $M_{\rm sph,tot}(<r)$ using the posterior samples of the ellipsoidal cluster model:

$$f_{\rm gas}(<r) \equiv \frac{M_{\rm sph,gas}(<r)}{M_{\rm sph,tot}(<r)}, \tag{55}$$

where $M_{\rm sph}(<r)$ denotes the total mass enclosed within a sphere of radius r, $M_{\rm sph}(<r) = \int_{4\pi} d\Omega \int_0^r dr' r'^2 \rho(\boldsymbol{r}')$ with $d\Omega$ the solid angle. In Table 10, we list the values of ellipsoidal and spherical overdensity mass of the cluster evaluated at $\Delta = 200$ and 500.

The resulting $f_{\rm gas}$ profile is shown in Figure 16 as a function of integration radius r. The gas mass fraction within 0.9Mpc $\sim 1.2 r_{2500c}$ is estimated as $f_{\rm gas}(<0.9{\rm Mpc}) = 0.100^{+0.031}_{-0.016}$. When the gas mass measurements are extrapolated to r_{500c} (Table 10), we find $f_{\rm gas}(<r_{500c}) = 0.112^{+0.039}_{-0.020}$. When compared to the cosmic baryon fraction f_b inferred from Planck Collaboration et al. (2015b), $f_{\rm gas}(<r_{500c})/f_b = 0.71^{+0.25}_{-0.12}$. These are consistent with typical values observed for high-mass clusters (Allen et al. 2008; Umetsu et al. 2012; Okabe et al. 2014).

Previous studies based on X-ray and lensing data found relatively low $f_{\rm gas}$ values for A1689 using lensing total mass estimates, but assuming spherical symmetry[21]: $f_{\rm gas}(<0.25 r_{200c}) = (0.0557 \pm 0.0039) h_{70}^{-3/2}$ (Lemze et al. 2008); $f_{\rm gas}(<r_{2500c}) = 0.0552^{+0.0056}_{-0.0062}$, $f_{\rm gas}(<r_{500c}) = 0.0812^{+0.0145}_{-0.0157}$, and $f_{\rm gas}(<r_{200c}) = 0.1053^{+0.0227}_{-0.0246}$ (Okabe et al. 2014, see also Kawaharada et al. 2010).

[21] Lemze et al. (2008) found $r_{200c} = 1.71$ Mpc h^{-1} from their analysis.

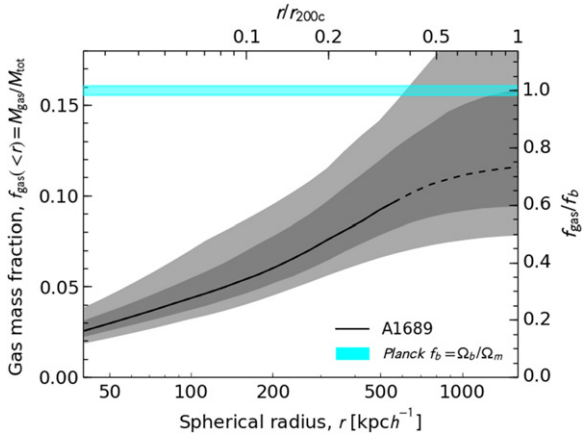

Figure 16. Ratio of spherically enclosed gas mass ($M_{\rm gas}$) to total mass ($M_{\rm tot}$) as a function of spherical radius r, derived from the full triaxial analysis of weak/strong-lensing and X-ray/SZE data. The middle line tracks the median. The gray shaded regions represent the 68.3% and 95.4% quantiles of the distribution. Portions of these lines are dashed to indicate extrapolations to larger cluster radii. The horizontal bar shows the cosmic baryon fraction $f_b = \Omega_b/\Omega_m$ determined by Planck Collaboration et al. (2015b).

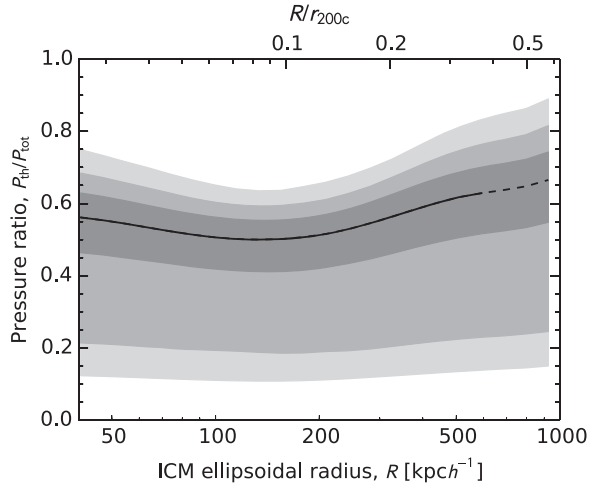

Figure 17. Ratio of the thermal gas pressure ($P_{\rm th}$) to the total equilibrium pressure ($P_{\rm tot}$) in A1689 as a function of the ellipsoidal radius R measured along the major axis of the ICM halo. The middle line tracks the median. The gray shaded regions show the 68.3%, 99.4%, and 99.7% quantiles of the distribution, respectively. Portions of these lines are dashed to indicate extrapolations to larger cluster radii.

Umetsu et al. (2009) measured gas fractions for a sample of four high-mass clusters including A1689 from a joint analysis of AMiBA SZE and Subaru weak-lensing observations, combined with published X-ray temperature measurements. Assuming spherical symmetry, they found for A1689 $f_{\rm gas}(<r_{2500c}) = 0.098^{+0.025}_{-0.026}$ and $f_{\rm gas}(<r_{500c}) = 0.115 \pm 0.029$, in excellent agreement with our results. Their gas fraction measurements are expected to be less sensitive to triaxiality because their $f_{\rm gas}$ estimator depends on the ratio of the SZE and lensing signals, which are subject to similar projection effects albeit with somewhat different degrees of impact.

8.4. Degree of HSE

A quantitative assessment of the degree of equilibrium in the ICM is a critical issue for cluster cosmology based on hydrostatic mass estimates (e.g., Planck Collaboration et al. 2014; Sereno & Ettori 2014). A significant advantage of our method is the ability to determine the intrinsic structure, shape, and orientation of the cluster system without a priori assuming HSE (Sereno et al. 2013). This allows us to compare the ICM properties directly to the gravitating mass corrected for projection effects, and thus to quantify the contribution of the thermal gas pressure $P_{\rm th}$ to the total equilibrium pressure $P_{\rm tot}$ (Kawaharada et al. 2010; Molnar et al. 2010). Here $P_{\rm tot}$ is determined by the gravitational potential Φ through $\nabla P_{\rm tot} = -\rho_{\rm gas}\nabla\Phi$ with $\rho_{\rm gas}$ the gas mass density. A consequence of the pressure equilibrium is the *X-ray shape theorem* (Buote & Canizares 1994), namely, that the gas in strict HSE is expected to follow iso-potential surfaces of the underlying matter distribution. For A1689, we find that the gas is more elongated than the gravitational potential (see Section 8.2), which points to a deviation from equilibrium.

In Figure 17, we show the ratio of thermal to equilibrium gas pressure, $P_{\rm th}/P_{\rm tot}$, as a function of ellipsoidal radius R of the ICM distribution. To this end, we have recomputed the posterior probability distributions for the cluster parameters, by imposing a sharp prior of $e^{\rm ICM}/e = 0.7$ (see Sereno et al. 2013), corresponding to the assumption that the gas shape follows the gravitational potential. We find $P_{\rm th}/P_{\rm tot} \sim 0.6$ out to $\sim$0.9 Mpc ($\sim 0.4 r_{200c}$), indicating a significant level ($\sim$40%) of non-thermal pressure support. The results here are consistent with Sereno et al. (2013), although our analysis favors a slightly higher level of non-thermal pressure support. We find no significant radial trend in the $P_{\rm th}/P_{\rm tot}$ ratio profile.

Our results are in agreement with Molnar et al. (2010), who analyzed a simulated sample of massive regular clusters of $(1-2) \times 10^{15}\, M_\odot\, h^{-1}$ having a smooth density profile, drawn from high-resolution cosmological simulations. Their simulations are therefore highly relevant to interpreting the observations of A1689. They found a significant non-thermal contribution due to subsonic gas motions in the core region (20%–45%), a minimum contribution (5%–30%) at about $0.1 r_{\rm vir}$ (Lau et al. 2009), growing outward to about 30%–45% at the virial radius $r_{\rm vir}$ (Nelson et al. 2014).

Molnar et al. (2010) also tested the validity of HSE in A1689 using gravitational lensing (see Umetsu & Broadhurst 2008; Kawaharada et al. 2010) and *Chandra* X-ray observations under the assumption of spherical geometry, finding a non-thermal contribution of $\lesssim$40%. As discussed by Sereno et al. (2013), this however indicates that this test is highly sensitive to biases in the X-ray temperature measurements (Donahue et al. 2014). For the cluster, we find the *Chandra* temperatures are about 10% higher than the *XMM-Newton* results used here (Sereno et al. 2012), so that the thermal contribution $P_{\rm th}/P_{\rm tot} \gtrsim 0.6$ obtained by Molnar et al. (2010) could be correspondingly overestimated relative to our results based on the *XMM-Newton* temperatures.

By combining *Suzaku* X-ray observations with the same lensing data as used in Molnar et al. (2010), Kawaharada et al. (2010) showed, assuming spherical symmetry, that the thermal gas pressure within r_{500c} is at most 40%–60% of the equilibrium pressure and 30%–40% around the virial radius. Intriguingly, their *Suzaku* observations reveal anisotropic distributions of gas temperature and entropy in cluster outskirts at $\gtrsim r_{500c}$, correlated with large-scale structure of galaxies surrounding the cluster. The outskirt regions in contact with low-density void environments have low gas temperatures and

entropies, indicating that the outskirts of A1689 are in the process of being thermalized (Kawaharada et al. 2010). Their *Suzaku* temperature measurements are in agreement with the *XMM-Newton* results (Sereno et al. 2012).

Morandi et al. (2011, see also Limousin et al. 2013) obtained $M_{200c} = (1.81 \pm 0.06) \times 10^{15}\, M_\odot\, h^{-1}$, $c_{200c} = 5.71 \pm 0.47$, and $q_a \sim 0.5$ for A1689 from a joint analysis of *Chandra* X-ray, weak-lensing, and strong-lensing data (see Table 9). The inferred level of triaxiality is similar to what we have found (Table 7), whereas the concentration is somewhat smaller and the mass is significantly higher than our results. They found that about 20% of the total ICM pressure is in non-thermal form, by assuming that P_{th}/P_{tot} is constant with radius and the gas shape follows the form expected for HSE. We note again that the P_{th}/P_{tot} results are also sensitive to calibration biases in the X-ray temperature measurements.

The mass discrepancy between the present results and those by Morandi et al. (2011) can be explained by the difference in their relative weights assigned to the weak- and strong-lensing data sets. As we have seen in Section 8.1.1, the *HST* strong-lensing data favor higher values of $M(<r_{200c})$ (Table 7), although this represents a significant extrapolation beyond the radial range covered by the multiple images. Hence, if the parameter constraints are highly dominated by strong lensing, this could lead to an overestimate of M_{200c}.

8.5. Comparison With Planck Data

We compare the SZE measurements from the interferometric data presented in Section 7.2 with a total power estimate based on the recent *Planck* data (Planck Collaboration et al. 2015a). A1689 is detected by *Planck* with high significance (S/N > 15, Planck Collaboration et al. 2015d). We construct *Planck* SZE maps in two different ways with different assumptions, using the data in the 143, 217, and 353 GHz channels. The 217 and 353 GHz bands are used primarily to remove the CMB and Galactic foregrounds. The difference between the two maps accounts for different assumptions about the Galactic components: one is based on local estimates of the dust properties, and the other is on global properties. The resulting SZE maps are obtained at an effective resolution of 8′ FWHM. The SZE signal is integrated as a function of clustercentric radius. We obtain a direct estimate for the total Compton Y parameter of $Y_{Planck} = (3.8 \pm 0.8) \times 10^{-10}$ sr integrated out to a sufficiently large radius 13′ ($\sim r_{200c}$), beyond which the integrated SZE signal converges. Here the error is estimated from aperture photometry in the background regions

This direct *Planck* measurement of the total SZE signal can be compared to the results inferred from the interferometric SZA observations (Section 7.2). Taking $Y_{SZA}(<6')$ (Table 6) as a lower limit on the total SZE flux, we find $Y_{SZA}(<6')/Y_{Planck} = 1.45 \pm 0.44$. Hence, the results from two independent SZE instruments operating at different angular scales are compatible with each other at 1σ. The relatively low Y value derived from the *Planck* data could be understood in light of the low gas temperature and entropy at $\gtrsim r_{500c}$ observed by the *Suzaku* X-ray satellite (Section 8.4). The *Suzaku* X-ray observations are in agreement with the thermal pressure profile of A1689 obtained from *Planck* data out to $\sim 2r_{500c}$.

When compared to *Planck*'s hydrostatic mass estimate, $M_{500c} = (8.77 \pm 0.34) \times 10^{14}\, M_\odot\, h_{70}^{-1}$, our lensing mass measurements (Table 10) give a spherical mass ratio of $M_{Planck}/M_{GL} = 0.70 \pm 0.15$ and 0.58 ± 0.10 with and without corrections for lensing projection effects, respectively.

9. SUMMARY

We have carried out a 3D multi-probe analysis of the rich cluster A1689, one of the most powerful known lenses on the sky ($\theta_{Ein} = 47''.0 \pm 1''.2$ at $z_s = 2$, Table 1), by combining improved weak-lensing data from new wide-field $BVR_C i' z'$ Subaru/Suprime-Cam observations (Sections 3 and 4) with complementary strong-lensing (Section 5), X-ray and SZE (Section 7.2) data sets.

We have generalized the 1D weak-lensing inversion method of Umetsu et al. (2011b) to a 2D description of the mass distribution without assuming particular functional forms (Section 2). This free-form method combines the spatial shear pattern with azimuthally averaged magnification information, the combination of which breaks the mass-sheet degeneracy.

We have reconstructed the projected matter distribution from a joint weak-lensing analysis of 2D shear and azimuthally integrated magnification constraints (Section 4). The resulting mass distribution reveals elongation with an axis ratio of $q_\perp \sim 0.7$ in projection (Figures 1 and 8), aligned well with the distributions of cluster galaxies and ICM (see Kawaharada et al. 2010). When assuming a spherical NFW halo, our full weak-lensing analysis yields a projected halo concentration of $c_{200c}^{2D} = 8.9 \pm 1.1$ ($c_{vir}^{2D} \sim 11$), which is consistent with and improved from earlier weak-lensing work based on Subaru Vi' imaging (Umetsu & Broadhurst 2008; Umetsu et al. 2011b).

We obtain excellent consistency between weak and strong lensing in the region where these independent data overlap, $\lesssim$200 kpc (Figures 6 and 10). We also find an improved agreement between weak and strong lensing in terms of constraints on projected NFW parameters (Figure 11) relative to previous work (Sereno & Umetsu 2011). This is largely due to improved techniques for strong-lensing reconstruction and to careful regularization of the covariance matrix (Section 7.1.2).

In a parametric triaxial framework, we have determined the intrinsic structure, shape, and orientation of the matter and gas distributions of the cluster, by combining weak/strong lensing with X-ray/SZE data under minimal geometric assumptions (Section 7). We have shown that the data favor a triaxial geometry with minor–major axis ratio $q_a = 0.39 \pm 0.15$ and major axis closely aligned with the line of sight ($\vartheta = 22^\circ \pm 10^\circ$). A spherical configuration for A1689 has been strongly ruled out. We obtain a halo mass $M_{200c} = (1.24 \pm 0.16) \times 10^{15}\, M_\odot\, h^{-1}$ and a halo concentration $c_{200c} = 8.36 \pm 1.27$, which is higher than typical concentrations found for high-mass clusters ($3 \lesssim c_{200c} \lesssim 6$; e.g., Okabe et al. 2013; Merten et al. 2014; Umetsu et al. 2014), but overlaps well with the $\gtrsim 1\sigma$ tail of the predicted distribution (Figure 15; Bhattacharya et al. 2013; Meneghetti et al. 2014; Diemer & Kravtsov 2015).

We find that the ICM is mildly triaxial with $q_a^{ICM} = 0.60 \pm 0.14$ and $q_b^{ICM} = 0.70 \pm 0.16$ (Table 8). The gas distribution is rounder than the underlying matter, $e^{ICM}/e = 0.87 \pm 0.07$, but more elongated than the gravitational potential ($e^{ICM}/e \gtrsim 0.7$), suggesting a deviation from equilibrium. The gas mass fraction enclosed within a sphere of radius $r = 0.9$ Mpc $\sim 1.2 r_{2500c}$ is found to be $f_{gas} = 10.0^{+3.1}_{-1.6}\%$. When the gas mass measurements are extrapolated to r_{500c}, $f_{gas}(<r_{500c}) = 11.2^{+3.9}_{-2.0}\%$. When

compared to the cosmic baryon fraction f_b (Planck Collaboration et al. 2015b), we find $f_{\rm gas}(<r_{500c})/f_b = 0.71^{+0.25}_{-0.12}$ (Figure 16). These are consistent with typical values observed for high-mass clusters. The thermal gas pressure contributes to ∼60% of the total pressure out to ∼0.9 Mpc (Figure 17), indicating a significant level of non-thermal pressure support. The results are, however, sensitive to calibration biases in the X-ray temperature measurements (Donahue et al. 2014). When compared to *Planck*'s hydrostatic mass estimate, our lensing mass measurements yield a spherical mass ratio of $M_{\rm Planck}/M_{\rm GL} = 0.70 \pm 0.15$ and 0.58 ± 0.10 with and without corrections for lensing projection effects, respectively.

Extending this work to larger samples of clusters will enable us to recover intrinsic distributions of cluster structural properties (e.g., M_{200c}, c_{200c}) and axis ratios (q_a, q_b), for a direct statistical comparison with the standard ΛCDM paradigm and for a wider examination of alternative DM scenarios (e.g., Schive et al. 2014). The CLASH survey (Postman et al. 2012) provides such ideal multiwavelength data sets of high quality (Czakon et al. 2014; Donahue et al. 2014; Rosati et al. 2014; Umetsu et al. 2014; Zitrin et al. 2014), for a sizable sample of 25 high-mass clusters.

We thank the anonymous referee for constructive comments and suggestions. We are grateful for discussions with Sherry Suyu, Adi Zitrin, and Radek Wojtak. We acknowledge the Subaru Support Astronomers plus Kai-Yang Lin and Hiroaki Nishioka, for assistance with our Subaru observations. We thank Nick Kaiser for making the IMCAT package publicly available. We thank Oliver Czoske for providing the redshift survey information for A1689. The work is partially supported by the Ministry of Science and Technology of Taiwan under the grant MOST 103-2112-M-001-030-MY3. M. S. acknowledges financial contributions from contracts ASI/INAF I/023/12/0, by the PRIN MIUR 2010–2011 "The dark universe and the cosmic evolution of baryons: from current surveys to Euclid" and by the PRIN INAF 2012 "The universe in the box: multiscale simulations of cosmic structure." M. N. acknowledges financial support from PRIN INAF 2014. J. M. D. acknowledges support of the consolider project CSD2010-00064 and AYA2012-39475-C02-01 funded by the Ministerio de Economia y Competitividad. N. O. is supported by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (26800097). This work was partially supported by "World Premier International Research Center Initiative (WPI Initiative)" and the Funds for the Development of Human Resources in Science and Technology under MEXT, Japan. This research was performed while T. M. held a National Research Council Research Associateship Award at the Naval Research Laboratory (NRL). We thank John Carlstrom, Megan Gralla, Marshall Joy, Dan Marrone, and the entire SZA and OVRO/BIMA teams for providing the SZA and OVRO/BIMA data used in this study. Support for the SZA observations presented in this work was provided by NSF through award AST-0838187 and PHY-0114422 at the University of Chicago. The OVRO and BIMA observations presented here were supported by National Science Foundation grants AST 99-81546 and 02-28963.

APPENDIX A
NONLINEAR EFFECT ON THE SOURCE-AVERAGED LENSING FIELDS

A.1 Reduced Gravitational Shear

The reduced shear, $g = \gamma/(1-\kappa)$, is nonlinear with κ, so that the averaging operator with respect to the source redshift acts nonlinearly on κ. In general, a spread of the source redshift distribution, in combination with the single source-plane approximation, may lead to an overestimation of the gravitational shear in the nonlinear regime.

Let us expand the reduced shear $g = g(z)$ with respect to $\kappa(z) = W(z)\kappa_\infty$ and $\gamma(z) = W(z)\gamma_\infty$ as

$$g = \gamma/(1-\kappa) = W\gamma_\infty\left(1 - W\kappa_\infty\right)^{-1} = W\gamma_\infty \sum_{k=0}^{\infty}\left(W\kappa_\infty\right)^k. \tag{56}$$

The reduced shear averaged over the source redshift distribution is expressed as

$$\langle g\rangle = \gamma_\infty \sum_{k=0}^{\infty}\left\langle W^{k+1}\right\rangle_g \kappa_\infty^k, \tag{57}$$

where the angular brackets represent an ensemble average over the redshift distribution of background sources. In the weak-lensing limit where $\kappa_\infty \ll 1$, $\langle g\rangle \approx \langle W\rangle_g\gamma_\infty \equiv \langle\gamma\rangle$. The next order of approximation is

$$\langle g\rangle \approx \gamma_\infty\left(\langle W\rangle_g + \left\langle W^2\right\rangle_g \kappa_\infty\right) \approx \frac{\langle W\rangle_g\gamma_\infty}{1 - \kappa_\infty\left\langle W^2\right\rangle_g\Big/\langle W\rangle_g}. \tag{58}$$

Seitz & Schneider (1997) showed that Equation (58) yields an excellent approximation in the mildly nonlinear regime with $\kappa_\infty \lesssim 0.6$. Defining $f_{W,g} \equiv \langle W^2\rangle_g\big/\langle W\rangle_g^2$, we have the following expression for the source-averaged reduced shear valid in the mildly nonlinear regime:

$$\langle g\rangle \approx \frac{\langle\gamma\rangle}{1 - f_{W,g}\langle\kappa\rangle}, \tag{59}$$

with $\langle\kappa\rangle = \langle W\rangle_g\kappa_\infty$. For a lens at relatively low redshift, $\langle W^2\rangle_g \approx \langle W\rangle_g^2$ and $f_{W,g} \approx 1$, leading to the single source-plane approximation: $\langle g\rangle \approx \langle\gamma\rangle/(1-\langle\kappa\rangle)$. The level of bias introduced by this approximation is $\Delta g/g \approx (f_{W,g}-1)\langle\kappa\rangle$. In typical ground-based deep observations of $z_l \lesssim 0.5$ clusters, $\Delta f_W \equiv f_W - 1$ is found to be of the order of several percent (Umetsu et al. 2014), so that the relative error is negligibly small in the mildly nonlinear regime.

A.2 Magnification Bias

Let us consider a maximally depleted sample of background sources with $\alpha = -d\log\overline{N}_\mu(>F)/dF = 0$, for which the effect of magnification bias is purely geometric, $b_\mu = \mu^{-1}$, and insensitive to the intrinsic source luminosity function. In the nonlinear subcritical regime, the source-averaged

magnification bias is expressed as (Umetsu 2013; Umetsu et al. 2014)

$$\left\langle \mu^{-1} \right\rangle = \left(1 - \langle\kappa\rangle\right)^2 - \left|\langle\gamma\rangle\right|^2 + \left(f_{W,\mu} - 1\right)\left(\langle\kappa\rangle^2 - \langle\gamma\rangle^2\right) \approx \left(1 - \langle\kappa\rangle\right)^2 - \left|\langle\gamma\rangle\right|^2, \tag{60}$$

where $f_{W,\mu} \equiv \langle W^2\rangle_\mu / \langle W\rangle_\mu^2$ is of the order of unity, $\langle\kappa\rangle = \langle W\rangle_\mu \kappa_\infty$, and $\langle\gamma\rangle = \langle W\rangle_\mu \gamma_\infty$. Hence, the error associated with the single source-plane approximation is $\langle\Delta\mu^{-1}\rangle = (f_{W,\mu} - 1)(\langle\kappa\rangle^2 - \langle\gamma\rangle^2) \equiv \Delta f_{W,\mu}(\langle\kappa\rangle^2 - \langle\gamma\rangle^2)$, which is much smaller than unity for background populations of our concern ($\Delta f_{W,\mu} \sim O(10^{-2})$) in the mildly nonlinear subcritical regime where $\langle\kappa\rangle \sim |\langle\gamma\rangle| \sim O(10^{-1})$. It is therefore reasonable to use the single source-plane approximation for calculating the magnification bias of depleted source populations with $\alpha \ll 1$.

APPENDIX B
DISCRETIZED EXPRESSIONS FOR CLUSTER LENSING PROFILES

First, we derive a discrete expression for the mean interior convergence $\kappa_\infty(<\theta)$ as a function of clustercentric radius θ using the azimuthally averaged convergence $\kappa_\infty(\theta)$. In the continuous limit, the mean convergence $\kappa_\infty(<\theta)$ interior to radius θ can be expressed in terms of $\kappa_\infty(\theta)$ as

$$\kappa_\infty(<\theta) = \frac{2}{\theta^2}\int_0^\theta d\,\ln\theta'\theta'^2\kappa_\infty(\theta'). \tag{61}$$

For a given set of $(N_{\rm bin}+1)$ concentric radii θ_i $(i = 1, \ldots, N_{\rm bin}+1)$, defining $N_{\rm bin}$ radial bands in the range $\theta_{\rm min} \equiv \theta_1 \leqslant \theta \leqslant \theta_{N_{\rm bin}+1} \equiv \theta_{\rm max}$, a discretized estimator for $\kappa_\infty(<\theta)$ can be written in the following way:

$$\kappa_\infty(<\theta_i) = \left(\frac{\theta_{\rm min}}{\theta_i}\right)^2 \kappa_\infty(<\theta_{\rm min}) + \frac{2}{\theta_i^2}\sum_{j=1}^{i-1}\Delta\ln\theta_j\bar{\theta}_j^2\kappa_\infty(\bar{\theta}_j), \tag{62}$$

with $\Delta\ln\theta_i \equiv (\theta_{i+1} - \theta_i)/\bar{\theta}_i$ and $\bar{\theta}_i$ the area-weighted center of the ith annular bin defined by $[\theta_i, \theta_{i+1}]$. In the continuous limit, we have

$$\bar{\theta}_i = 2\int_{\theta_i}^{\theta_{i+1}} d\theta'\theta'^2 \Big/ \left(\theta_{i+1}^2 - \theta_i^2\right) = \frac{2}{3}\frac{\theta_i^2 + \theta_{i+1}^2 + \theta_i\theta_{i+1}}{\theta_i + \theta_{i+1}}.$$

Next, we derive discretized expressions for the tangential reduced shear $g_+(\theta)$ and the inverse magnification $\mu^{-1}(\theta)$ in terms of the binned convergence $\kappa_\infty(\bar{\theta}_i)$, using the following relations:

$$g_+(\bar{\theta}_i) = \frac{\langle W\rangle_g\left[\kappa_\infty(<\bar{\theta}_i) - \kappa_\infty(\bar{\theta}_i)\right]}{1 - f_{W,g}\langle W\rangle_g\kappa_\infty(\bar{\theta}_i)}, \tag{63}$$

$$\mu^{-1}(\bar{\theta}_i) = \left[1 - \langle W\rangle_\mu\kappa_\infty(\bar{\theta}_i)\right]^2 - \langle W\rangle_\mu^2\left[\kappa_\infty(<\bar{\theta}_i) - \kappa_\infty(\bar{\theta}_i)\right]^2, \tag{64}$$

where both the quantities depend on the mean convergence interior to the radius $\bar{\theta}_i$, $\kappa_\infty(<\bar{\theta}_i)$. By assuming a constant density in each radial band, we find the following expression for $\kappa_\infty(<\bar{\theta}_i)$:

$$\kappa_\infty(<\bar{\theta}_i) = \frac{1}{2}\left[\left(\theta_i/\bar{\theta}_i\right)^2\kappa_\infty(<\theta_i) + \left(\theta_{i+1}/\bar{\theta}_i\right)^2\kappa_\infty(<\theta_{i+1})\right], \tag{65}$$

where $\kappa_\infty(<\theta_i)$ and $\kappa_\infty(<\theta_{i+1})$ can be computed using Equation (62).

Accordingly, all relevant cluster lensing observables, $g_+(\theta)$ and $n_\mu(\theta)$, can be uniquely specified by the binned convergence profile $\{\kappa_{\infty,{\rm min}}, \kappa_{\infty,i}\}_{i=1}^{N_{\rm bin}}$ with $\kappa_{\infty,{\rm min}} \equiv \kappa_\infty(<\theta_{\rm min})$ and $\kappa_{\infty,i} \equiv \kappa_\infty(\bar{\theta}_i)$.

APPENDIX C
TWO-DIMENSIONAL TO ONE-DIMENSIONAL PROJECTION

To make a direct comparison between the results from 1D and 2D weak-lensing analyses, we construct a projected mass profile $\Sigma(\theta)$ from an optimally weighted radial projection of the $\Sigma(\boldsymbol{\theta})$ field as (Morandi et al. 2011)

$$\boldsymbol{\Sigma}_{(1)} = \left[A^t C_{(2)}^{-1} A\right]^{-1} A^t C_{(2)}^{-1}\boldsymbol{\Sigma}_{(2)} \tag{66}$$

where $\boldsymbol{\Sigma}_{(2)} = \{\Sigma(\boldsymbol{\theta}_m)\}_{m=1}^{N_{\rm pix}}$ is a pixelated mass map, $C_{(2)}$ is the pixel–pixel covariance matrix of $\boldsymbol{\Sigma}_{(2)}$, $\boldsymbol{\Sigma}_{(1)}$ is a vector of radially binned Σ values, and A is a mapping matrix whose elements A_{mi} represent the fraction of the area of the mth pixel lying within the ith clustercentric radial bin (Section 2.3.2). The covariance matrix for $\boldsymbol{\Sigma}_{(1)}$ is given by

$$C_{(1)} = \left[A^t C_{(2)}^{-1} A\right]^{-1}. \tag{67}$$

REFERENCES

Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., et al. 2008, ApJS, 175, 297
Allen, S. W., Rapetti, D. A., Schmidt, R. W., et al. 2008, MNRAS, 383, 879
Andersson, K. E., & Madejski, G. M. 2004, ApJ, 607, 190
Arnaud, M., Pratt, G. W., Piffaretti, R., et al. 2010, A&A, 517, A92
Bartelmann, M., & Schneider, P. 2001, PhR, 340, 291
Becker, M. R., & Kravtsov, A. V. 2011, ApJ, 740, 25
Beers, T. C., Flynn, K., & Gebhardt, K. 1990, AJ, 100, 32
Benítez, N. 2000, ApJ, 536, 571
Benitez, N., Dupke, R., Moles, M., et al. 2014, arXiv:1403.5237
Benítez, N., Ford, H., Bouwens, R., et al. 2004, ApJS, 150, 1
Bertin, E. 2006, in ASP Conf. Ser. 351, Astronomical Data Analysis Software and Systems XV ed. C. Gabriel et al. (San Francisco, CA: ASP), 112
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bertin, E., Mellier, Y., Radovich, M., et al. 2002, in in ASP Conf. Ser. 281, Astronomical Data Analysis Software and Systems XI, ed. D. A. Bohlender, D. Durand & T. H. Handley (San Francisco, CA: ASP), 228
Bhattacharya, S., Habib, S., Heitmann, K., & Vikhlinin, A. 2013, ApJ, 766, 32
Bleem, L. E., Stalder, B., de Haan, T., et al. 2015, ApJS, 216, 27
Bonamente, M., Hasler, N., Bulbul, E., et al. 2012, NJPh, 14, 025010
Broadhurst, T., Benítez, N., Coe, D., et al. 2005b, ApJ, 621, 53

Broadhurst, T., Takada, M., Umetsu, K., et al. 2005a, ApJL, 619, L143
Broadhurst, T. J., & Barkana, R. 2008, MNRAS, 390, 1647
Broadhurst, T. J., Taylor, A. N., & Peacock, J. A. 1995, ApJ, 438, 49
Buote, D. A., & Canizares, C. R. 1994, ApJ, 427, 86
Buote, D. A., & Humphrey, P. J. 2012a, MNRAS, 420, 1693
Buote, D. A., & Humphrey, P. J. 2012b, MNRAS, 421, 1399
Clowe, D., Bradač, M., Gonzalez, A. H., et al. 2006, ApJL, 648, L109
Coe, D., Benítez, N., Broadhurst, T., & Moustakas, L. A. 2010, ApJ, 723, 1678
Coe, D., Benítez, N., Sánchez, S. F., et al. 2006, AJ, 132, 926
Coe, D., Umetsu, K., Zitrin, A., et al. 2012, ApJ, 757, 22
Coles, J. 2008, ApJ, 679, 17
Comerford, J. M., & Natarajan, P. 2007, MNRAS, 379, 190
Connolly, A. J., Szalay, A. S., & Brunner, R. J. 1998, ApJL, 499, L125
Corless, V. L., King, L. J., & Clowe, D. 2009, MNRAS, 393, 1235
Covone, G., Sereno, M., Kilbinger, M., & Cardone, V. F. 2014, ApJL, 784, L25
Czakon, N. G., Sayers, J., Mantz, A., et al. 2014, arXiv:1406.2800
Czoske, O. 2004, in IAU Colloq. 195, Outskirts of Galaxy Clusters: Intense Life the Suburbs, ed. A. Diaferio (Cambridge: Cambridge Univ. Press), 183
De Filippis, E., Sereno, M., Bautz, M. W., & Longo, G. 2005, ApJ, 625, 108
Diego, J. M., Broadhurst, T., Benitez, N., et al. 2015, MNRAS, 446, 683
Diemer, B., & Kravtsov, A. V. 2015, ApJ, 799, 108
Donahue, M., Voit, G. M., Mahdavi, A., et al. 2014, ApJ, 794, 136
Du, W., & Fan, Z. 2014, ApJ, 785, 57
Ettori, S., Morandi, A., Tozzi, P., et al. 2009, A&A, 501, 61
Ford, J., Hildebrandt, H., van Waerbeke, L., et al. 2012, ApJ, 754, 143
Ford, J., Hildebrandt, H., van Waerbeke, L., et al. 2014, MNRAS, 439, 3755
Gralla, M. B., Sharon, K., Gladders, M. D., et al. 2011, ApJ, 737, 74
Grillo, C., Eichner, T., Seitz, S., et al. 2010, ApJ, 710, 372
Gruen, D., Seitz, S., Becker, M. R., Friedrich, O., & Mana, A. 2015, MNRAS, 449, 4264
Gruen, D., Seitz, S., Brimioulle, F., et al. 2014, MNRAS, 442, 1507
Haiman, Z., Mohr, J. J., & Holder, G. P. 2001, ApJ, 553, 545
Halkola, A., Seitz, S., & Pannella, M. 2006, MNRAS, 372, 1425
Hamana, T., Miyazaki, S., Kashikawa, N., et al. 2009, PASJ, 61, 833
Heymans, C., Van Waerbeke, L., Bacon, D., et al. 2006, MNRAS, 368, 1323
Hildebrandt, H., Muzzin, A., Erben, T., et al. 2011, ApJL, 733, L30
Hjorth, J., & Williams, L. L. R. 2010, ApJ, 722, 851
Hoekstra, H., Herbonnet, R., Muzzin, A., et al. 2015, MNRAS, 449, 685
Hu, W., & Kravtsov, A. V. 2003, ApJ, 584, 702
Hwang, H. S., Geller, M. J., Diaferio, A., Rines, K. J., & Zahid, H. J. 2014, ApJ, 797, 106
Ilbert, O., Salvato, M., Le Floc'h, E., et al. 2010, ApJ, 709, 644
Israel, H., Schellenberger, G., Nevalainen, J., Massey, R., & Reiprich, T. H. 2015, MNRAS, 448, 814
Jauzac, M., Clément, B., Limousin, M., et al. 2014, MNRAS, 443, 1549
Jimeno, P., Broadhurst, T., Coupon, J., Umetsu, K., & Lazkoz, R. 2015, MNRAS, 448, 1999
Jing, Y. P., & Suto, Y. 2002, ApJ, 574, 538
Johnston, D. E., Sheldon, E. S., Wechsler, R. H., et al. 2007, arXiv:0709.1159
Kaiser, N., & Squires, G. 1993, ApJ, 404, 441
Kaiser, N., Squires, G., & Broadhurst, T. 1995, ApJ, 449, 460
Kasun, S. F., & Evrard, A. E. 2005, ApJ, 629, 781
Kawahara, H. 2010, ApJ, 719, 1926
Kawaharada, M., Okabe, N., Umetsu, K., et al. 2010, ApJ, 714, 423
LaRoque, S. J., Bonamente, M., Carlstrom, J. E., et al. 2006, ApJ, 652, 917
Lau, E. T., Kravtsov, A. V., & Nagai, D. 2009, ApJ, 705, 1129
Lee, J., & Suto, Y. 2003, ApJ, 585, 151
Lemze, D., Barkana, R., Broadhurst, T. J., & Rephaeli, Y. 2008, MNRAS, 386, 1092
Lemze, D., Broadhurst, T., Rephaeli, Y., Barkana, R., & Umetsu, K. 2009, ApJ, 701, 1336
Lilly, S. J., Le Fèvre, O., Renzini, A., et al. 2007, ApJS, 172, 70
Limousin, M., Morandi, A., Sereno, M., et al. 2013, SSRv, 177, 155
Limousin, M., Richard, J., Jullo, E., et al. 2007, ApJ, 668, 643
Lubini, M., & Coles, J. 2012, MNRAS, 425, 3077
Lubini, M., Sereno, M., Coles, J., Jetzer, P., & Saha, P. 2014, MNRAS, 437, 2461
Massey, R., Heymans, C., Bergé, J., et al. 2007, MNRAS, 376, 13
Medezinski, E., Broadhurst, T., Umetsu, K., Benítez, N., & Taylor, A. 2011, MNRAS, 414, 1840
Medezinski, E., Broadhurst, T., Umetsu, K., et al. 2007, ApJ, 663, 717
Medezinski, E., Broadhurst, T., Umetsu, K., et al. 2010, MNRAS, 405, 257
Medezinski, E., Umetsu, K., Nonino, M., et al. 2013, ApJ, 777, 43
Meneghetti, M., Rasia, E., Vega, J., et al. 2014, ApJ, 797, 34
Merten, J., Cacciato, M., Meneghetti, M., Mignone, C., & Bartelmann, M. 2009, A&A, 500, 681
Merten, J., Meneghetti, M., Postman, M., et al. 2014, arXiv:1404.1376
Miyazaki, S., Hamana, T., Ellis, R. S., et al. 2007, ApJ, 669, 714
Miyazaki, S., Komiyama, Y., Sekiguchi, M., et al. 2002, PASJ, 54, 833
Mohammed, I., Saha, P., & Liesenborgs, J. 2015, PASJ, 67, 219
Molnar, S. M., Chiu, I.-N., Umetsu, K., et al. 2010, ApJL, 724, L1
Morandi, A., Limousin, M., Rephaeli, Y., et al. 2011, MNRAS, 416, 2567
Mroczkowski, T., Bonamente, M., Carlstrom, J. E., et al. 2009, ApJ, 694, 1034
Nagai, D., Kravtsov, A. V., & Vikhlinin, A. 2007, ApJ, 668, 1
Narikawa, T., Kobayashi, T., Yamauchi, D., & Saito, R. 2013, PhRvD, 87, 124006
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. 2014, ApJ, 792, 25
Nieuwenhuizen, T. M., & Morandi, A. 2013, MNRAS, 434, 2679
Nonino, M., Dickinson, M., Rosati, P., et al. 2009, ApJS, 183, 244
Oguri, M., Bayliss, M. B., Dahle, H., et al. 2012, MNRAS, 420, 3213
Oguri, M., Takada, M., Okabe, N., & Smith, G. P. 2010, MNRAS, 405, 2215
Oguri, M., Takada, M., Umetsu, K., & Broadhurst, T. 2005, ApJ, 632, 841
Okabe, N., Smith, G. P., Umetsu, K., Takada, M., & Futamase, T. 2013, ApJL, 769, L35
Okabe, N., Takada, M., Umetsu, K., Futamase, T., & Smith, G. P. 2010, PASJ, 62, 811
Okabe, N., & Umetsu, K. 2008, PASJ, 60, 345
Okabe, N., Umetsu, K., Tamura, T., et al. 2014, PASJ, 66, 99
Park, C.-G., Ng, K.-W., Park, C., Liu, G.-C., & Umetsu, K. 2003, ApJ, 589, 67
Peng, E., Andersson, K., Bautz, M. W., & Garmire, G. P. 2009, ApJ, 701, 1283
Planck Collaboration, Adam, R., Ade, P. A. R., et al. 2015a, arXiv:1502.01582
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, A&A, 571, A20
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2015b, arXiv:1502.01589
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2015c, arXiv:1502.01597
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2015d, arXiv:1502.01598
Postman, M., Coe, D., Benítez, N., et al. 2012, ApJS, 199, 25
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, in Numerical Recipes in FORTRAN. The Art of Scientific Computing (2nd ed.; Cambridge: Cambridge Univ. Press)
Rapetti, D., Allen, S. W., Mantz, A., & Ebeling, H. 2010, MNRAS, 406, 1796
Rasia, E., Meneghetti, M., Martino, R., et al. 2012, NJPh, 14, 055018
Rines, K., Geller, M. J., Diaferio, A., & Kurtz, M. J. 2013, ApJ, 767, 15
Romanowsky, A. J., & Kochanek, C. S. 1998, ApJ, 493, 641
Rosati, P., Balestra, I., Grillo, C., et al. 2014, Msngr, 158, 48
Saha, P., & Williams, L. L. R. 2004, AJ, 127, 2604
Schive, H.-Y., Chiueh, T., & Broadhurst, T. 2014, NatPh, 10, 496
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Schneider, P., King, L., & Erben, T. 2000, A&A, 353, 41
Schneider, P., & Seitz, C. 1995, A&A, 294, 411
Seitz, C., & Schneider, P. 1997, A&A, 318, 687
Seljak, U. 1998, ApJ, 506, 64
Sereno, M. 2007, MNRAS, 380, 1207
Simet, M., Battaglia, N., Mandelbaum, R., & Seljak, U. 2015, arXiv:1502.01024
Sereno, M., & Ettori, S. 2014, arXiv:1407.7868
Sereno, M., Ettori, S., & Baldi, A. 2012, MNRAS, 419, 2646
Sereno, M., Ettori, S., & Moscardini, L. 2014a, arXiv:1407.7869
Sereno, M., Ettori, S., Umetsu, K., & Baldi, A. 2013, MNRAS, 428, 2241
Sereno, M., Lubini, M., & Jetzer, P. 2010, A&A, 518, A55
Sereno, M., & Umetsu, K. 2011, MNRAS, 416, 3187
Sereno, M., Veropalumbo, A., Marulli, F., et al. 2014b, arXiv:1410.5438
Sereno, M., & Zitrin, A. 2012, MNRAS, 419, 3280
Stark, A. A. 1977, ApJ, 213, 368
Taylor, A. N., Dye, S., Broadhurst, T. J., Benitez, N., & van Kampen, E. 1998, ApJ, 501, 539
Tyson, J. A., & Fischer, P. 1995, ApJL, 446, L55
Taylor, J. E., & Navarro, J. F. 2001, ApJ, 563, 483
Umetsu, K. 2013, ApJ, 769, 13
Umetsu, K., Birkinshaw, M., Liu, G.-C., et al. 2009, ApJ, 694, 1643
Umetsu, K., & Broadhurst, T. 2008, ApJ, 684, 177
Umetsu, K., Broadhurst, T., Zitrin, A., et al. 2011a, ApJ, 738, 41
Umetsu, K., Broadhurst, T., Zitrin, A., Medezinski, E., & Hsu, L. 2011b, ApJ, 729, 127
Umetsu, K., Medezinski, E., Broadhurst, T., et al. 2010, ApJ, 714, 1470
Umetsu, K., Medezinski, E., Nonino, M., et al. 2012, ApJ, 755, 56

Umetsu, K., Medezinski, E., Nonino, M., et al. 2014, ApJ, 795, 163
Umetsu, K., Tada, M., & Futamase, T. 1999, PThPS, 133, 53
Van Waerbeke, L., Mellier, Y., Erben, T., et al. 2000, A&A, 358, 30
Vikhlinin, A., Kravtsov, A., Forman, W., et al. 2006, ApJ, 640, 691
von der Linden, A., Mantz, A., Allen, S. W., et al. 2014, MNRAS, 443, 1973
Watson, W. A., Iliev, I. T., Diego, J. M., et al. 2014, MNRAS, 437, 3776
White, M., Carlstrom, J. E., Dragovan, M., & Holzapfel, W. L. 1999, ApJ, 514, 12
Wright, C. O., & Brainerd, T. G. 2000, ApJ, 534, 34
Zitrin, A., Fabris, A., Merten, J., et al. 2015, ApJ, 801, 44