Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **500,** 2627–2644 (2021) doi:10.1093/mnras/staa3193
Advance Access publication 2020 October 15

Reconciling galaxy cluster shapes, measured by theorists versus observers

David Harvey [1]★ Andrew Robertson [2,3] Sut-Ieng Tam,[2] Mathilde Jauzac,[2,3,4] Richard Massey [2,3] Jason Rhodes[5,6] and Ian G. McCarthy [7]

[1]*Lorentz Institute, Leiden University, Niels Bohrweg 2, Leiden, NL-2333 CA, the Netherlands*
[2]*Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK*
[3]*Centre for Extragalactic Astronomy, Department of Physics, Durham University, Durham DH1 3LE, UK*
[4]*Astrophysics and Cosmology Research Unit, School of Mathematical Sciences, University of KwaZulu-Natal, Durban 4041, South Africa*
[5]*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA*
[6]*California Institute of Technology, 1201 East California Blvd., Pasadena, CA 91125, USA*
[7]*Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK*

Accepted 2020 October 6. Received 2020 August 31; in original form 2020 July 17

ABSTRACT

If properly calibrated, the shapes of galaxy clusters can be used to investigate many physical processes: from feedback and quenching of star formation, to the nature of dark matter. Theorists frequently measure shapes using moments of inertia of simulated particles'. We instead create mock (optical, X-ray, strong-, and weak-lensing) observations of the 22 most massive ($\sim 10^{14.7}\,M_{\odot}$) relaxed clusters in the BAHAMAS simulations. We find that observable measures of shape are rounder. Even when moments of inertia are projected into 2D and evaluated at matched radius, they overestimate ellipticity by 56 per cent (compared to observable strong lensing) and 430 per cent (compared to observable weak lensing). Therefore, we propose matchable quantities and test them using observations of eight relaxed clusters from the *Hubble Space Telescope* (*HST*) and *C*handra X-Ray Observatory. We also release our *HST* data reduction and lensing analysis software to the community. In real clusters, the ellipticity and orientation angle at all radii are strongly correlated. In simulated clusters, the ellipticity of inner ($<r_{\rm vir}/20$) regions becomes decoupled: for example, with greater misalignment of the central cluster galaxy. This may indicate overly efficient implementation of feedback from active galactic nuclei. Future exploitation of cluster shapes as a function of radii will require better understanding of core baryonic processes. Exploitation of shapes on any scale will require calibration on simulations extended all the way to mock observations.

Key words: gravitational lensing: strong – gravitational lensing: weak – galaxies: clusters: general – dark matter – X-rays: galaxies: cluster.

1 INTRODUCTION

The Lambda cold dark matter (ΛCDM) concordant model of cosmology assumes that we are living in a Universe dominated by an unknown dark energy, accelerating the expansion of space–time and permeated by a dominant gravitating mass that we do not understand, dark matter (Parkinson et al. 2012; Kilbinger et al. 2013; Planck Collaboration XVI 2014; Hildebrandt et al. 2017; Abbott et al. 2018). In this framework, cosmological simulations predict that structure should form hierarchially, with the smallest haloes forming first (Springel et al. 2001; Sánchez et al. 2012; Vogelsberger et al. 2014; Schaye et al. 2015; McCarthy et al. 2017; Nelson et al. 2018). The resulting distribution of matter is in the form of a 'cosmic web' whereby filaments and sheets of mass funnel mass towards massive nodes or galaxy clusters.

Galaxy clusters are the largest known virialized structures in the Universe. They are dominated by dark matter, harbouring a halo of hot X-ray gas and in some cases thousands of galaxies (e.g. Smith et al. 2005; Postman et al. 2012; Lotz et al. 2017). As extreme peaks in the density field, clusters of galaxies are ideal laboratories to study dark matter (e.g. Kahlhoefer et al. 2014; Harvey et al. 2017a,b, 2019; Schwinn et al. 2017, 2018; Robertson et al. 2019; Banerjee et al. 2020) and constrain cosmology (e.g. Kratochvil, Haiman & May 2010; Marian et al. 2011; Cardone et al. 2013).

The mass of a galaxy cluster can exceed 10^{15} solar masses (e.g. von der Linden et al. 2012; Diego et al. 2015; Jauzac et al. 2015a, 2018; Merten et al. 2015; Zitrin et al. 2015). In these environments, space–time is deformed causing the observed image of distant galaxies that happen to align themselves with the cluster to be distorted. In extreme cases, the images can be stretched into arcs and split into multiple copies. Modelling strong gravitational lensing has become common place when measuring the mass distribution in clusters of galaxies (e.g. Merten et al. 2015; Zitrin et al. 2015; Jauzac et al. 2016a). However, strong lensing has its limitations, with the constraints limited to the core of the cluster, there is no information on substructures and mass in its outer regions. Weak gravitational lensing, where the effect of the cluster must be measured statistically, grants access to this missing information (e.g. Ragozzine et al. 2012; von der Linden et al. 2012). It is now normal to combine both weak

★ E-mail: harvey@lorentz.leidenuniv.nl (DH)

and strong gravitational lensing to get a full picture of the cluster (e.g. Cacciato et al. 2006; Diego et al. 2007; Merten et al. 2009; Merten et al. 2015; Zitrin et al. 2015). For a full review of mass mapping in clusters of galaxies, see Kneib & Natarajan (2011).

The exact form of galaxy clusters is still debated, however, it is generally accepted that they are triaxial in their shape with X-ray studies suggesting that roughly 70 per cent of all clusters tend to be prolate (Sereno et al. 2006). With X-ray shapes prone to non-thermal support, studies cited gravitational lensing as a more direct method to measure the shape of clusters. Oguri et al. (2010) studied the weak lensing properties of 25 clusters of galaxies, finding an average ellipticity, $\langle\epsilon\rangle = 0.46 \pm 0.04$. They interestingly found no correlation between both the lensing ellipticity and position angle of the member galaxies, suggesting no connection. More recently, Donahue et al. (2016) carried out a study where they identified the connection between the X-ray emission, the brightest cluster galaxy (BCG), and lensing (strong and weak combined). They found that there was a strong correlation between the position angle of the cluster at large radial distances ($r \sim 500$ kpc) and the BCG (inner 10 kpc), citing a coupling between the cluster and galactic star formation properties. Finally, a recent study of 20 relaxed and dynamically merging clusters looked at the misalignment of morphologies between the weak lensing and four probes, the Sunyaev–Zel'dovich effect, the X-ray morphology, the strong lensing morphologies, and the BCG (Umetsu et al. 2018).

Although many studies have measured the shapes of clusters, fewer studies exist comparing the shapes to hydrodynamical simulations (e.g. Umetsu et al. 2018). In these studies, it is normal to measure the inertia tensor directly from the particle data. By taking eigenvectors and eigenvalues, it is possible to extract directly the axial ratios and position angles. However, the radii in which the particles are chosen to measure this inertia tensor greatly impacts the inferred shape, and as such it is not clear how this shape relates to the weak or strong lensing (Velliscig et al. 2015).

In addition, sample matching is often overlooked. The majority of clusters that have been observed with deep enough imaging to measure strong lensing have extremely complicated selection functions, and the clusters selected are often highly irregular. For example, the Hubble Frontier Fields contain clusters that are incredibly complicated, with some studies even suggesting that they are in tension with ΛCDM (e.g. A2744; Jauzac et al. 2016b; Schwinn et al. 2017). As such, it raises the question of the comparability of the full sample of currently observed clusters with simulations. A recent study compared the complete sample of CLASH, HFF, and RELICS with simulations from the Horizon-AGN suite of simulations (Okabe et al. 2020). An interesting study that found the observed clusters appeared to have strong lensing regions significantly more elliptical than the BCG. However, many of the clusters used are massive merging clusters (for example, A2744, MACSJ0416, MACS1149.5) that have huge structures in-falling, which will bias the results should the simulations not reflect the identical sample.

In this paper, we study two key questions:

(i) Is the ellipticity calculated from the projected moment of inertia derived directly from particle data in simulations a good estimator of the shape derived from strong or weak lensing?

(ii) Is there any evidence for a radial dependent ellipticity in galaxy clusters, potentially signalling different physics acting on different regions of the cluster?

To answer these two questions, we carry out our investigation with two key differences to previous studies. First, we derive observationally matched products from the simulations so we can carry out a parallel analysis and make a direct comparison to observations, and secondly, we impose a strict selection cut on both the observed and simulated sample of clusters to ensure that they are analogous.

The paper is structured as follows: In Section 2, we outline the data that are used, including the reduction processes used and the simulations that have been adopted. In Section 3, we briefly review how we measure the ellipiticity from different probes (with a full description in Appendix A). In Section 4, we show our results and then in Section 5 we conclude.

2 DATA

2.1 *N*-body simulations of galaxy clusters

We use the BAHAMAS suite of simulations to compare the shapes of the observed clusters to clusters in a ΛCDM universe (McCarthy et al. 2017, 2018). BAHAMAS is a fully hydrodynamical simulation, with a baryonic feedback prescription based on the framework developed in the Overwhelmingly Large Simulations project (Schaye et al. 2015), that includes realistic feedback from active galactic nuclei (AGNs) and supernova, radiative cooling, star formation, chemodynamics, and stellar evolution. McCarthy et al. (2017) showed that the simulations can accurately reproduce the local stellar mass function and the gas mass–halo mass relation of galaxy clusters. We use fiducial resolution run of the BAHAMAS that simulated a large box of 400 Mpc h^{-1} with a plummer equivalent softening length of 4kpc h^{-1} with 1024^3 baryon and CDM particles in a WMAP9 cosmology (Bennett et al. 2013).

We use SUBFIND to find and extract the 40 largest galaxy clusters (and the subhaloes within these clusters) at a redshift of $z = 0.375$, cutting out boxes of 2 Mpc in the $x - y$ direction, and $\delta z = 10$ Mpc in projection. This depth is sufficient to encapsulate the additional mass from the environment of the clusters that will impact the shape of the cluster. Beyond this radius, the impact of line-of-sight structures will be minimal (see Section 4). The X-ray luminosity maps are derived according to section 5.1 of McCarthy et al. (2017). From these, we measure the X-ray concentration ratio, $\Gamma = S(< 100$ kpc$)/S(< 400$ kpc) and cut at $\Gamma > 0.2$ to ensure we have only relaxed clusters (Rasia, Meneghetti & Ettori 2013), leaving a total of 22 clusters. To derive the lensing observables, we first project these into a two-dimensional projected surface density and then carry out the following procedures:

(i) *Simulated strong lensing observable:* Our strong lensing catalogues are made from sources placed on to seven discrete source planes, equally spaced in redshift from $z_s = 1$ to $z_s = 4$. On each source plane, we place a regular grid of 128×128 sources, covering an area of 96×96 arcsec^2, centred behind the centre of the lensing cluster. Each source is modelled as an ellipse, with an area equal to that of a circle with radius 0.5 pkpc, and with an axial ratio drawn randomly from a uniform distribution between 0.5 and 1; a random position angle is also drawn for each source. For each of these sources, we calculate the locations and magnifications of all images of this source as they would appear in the (observed) lens plane, following the method in Robertson, Massey & Eke (2020). The deflection angles due to the lens are calculated on a regular grid in the lens plane from a pixelized map of the lensing convergence (on the same regular grid) using discrete Fourier transforms (see Robertson et al. 2019). Source-by-source, we find all points in the lens–plane grid that when mapped to the relevant source plane are enclosed by the boundary of the source. These points are split into contiguous sets in the lens plane (with steps in redshift of $\delta z = 0.5$), which are the individual lensed images. The magnification of an

Table 1. A summary of the galaxy cluster sample. Each cluster is confirmed relaxed from its X-ray observation. We show the following: Col 1. Name; Col 2. Survey; Col 3. Redshift; Col 4. Total, cleaned effective time *CXO* time; Col 5. Number of *HST* available filters; Col 6. Filter used for the weak lensing; Col 7. Exposure time for the weak lensing filter; Col 8. Total number of weak lensing sources available; Col 9. Number of strong lensing multiple images; Col 10. The root mean square of the strong lensing fit; Col 11. The virial mass estimated from strong lensing; Col 12. The concentration parameter, $c_{\rm vir}$ estimated from strong lensing. Note that for all clusters that are part of the CLASH sample, the publicly available photometric redshift catalogues are used.

Cluster	Survey	z	CXO (ks)	N_{HST}	WL filter	WL exp. (ks)	$N_{\rm WL}$	$N_{\rm I}$	RMS	$M_{200}(\times 10^{14}$ M$_\odot)$	$c_{\rm vir}$
A383	CLASH	0.19	53.449	16	*F*814*W*	4.243	1351	26	0.75	16.6 ± 5.1	4.0 ± 1.4
A2261	CLASH	0.22	37.526	16	*F*814*W*	4.099	1283	32	0.68	6.9 ± 0.6	9.0 ± 0.5
A1703	LOCUSS	0.28	82.295	7	*F*850*LP*	17.800	1738	42	1.03	13.5 ± 0.9	4.6 ± 0.2
A1835	LOCUSS	0.25	23.0781	4	*F*814*W*	2.360	1110	18	1.20	28.7 ± 1.7	3.7 ± 0.2
MACS0744	CLASH	0.69	96.111	16	*F*814*W*	8.893	762	20	1.52	9.9 ± 1.4	4.7 ± 0.8
MACS1206	CLASH	0.44	25.289	16	*F*814*W*	4.240	834	35	1.62	15.0 ± 0.2	4.8 ± 0.1
MACS1720	CLASH	0.39	71.764	16	*F*814*W*	3.988	960	17	1.61	9.8 ± 0.7	5.2 ± 0.3
MACS1931	CLASH	0.35	116.841	16	*F*814*W*	2.991	596	23	0.91	9.7 ± 0.3	5.0 ± 0.1

image is given by the number of lens–plane grid points that map into the source, divided by the expected number that would map into the source in the absence of lensing. The position of the image is given by the location of the lens–plane grid point that maps closest to the centre of the source in the source plane. The grid spacing in the lens plane is 0.02 arcsec.

With a regular grid of lensed sources at different x–y positions and redshifts, we randomly select a source to enter the final image catalogue. To do this, we first create a magnification-biased luminosity function for each source position based on Oesch et al. (2010). We then generate a random luminosity and calculate the observed luminosity of all its respective images. Assuming a limiting magnitude of $m < 30$, we then determine which images would be observed. By adding the condition that we must observe counter images, we have our final image catalogue.

(ii) *Simulated weak lensing observable:* Kaiser & Squires (1993) showed that the normalized projected surface density (or convergence), κ can be related to the weak lensing shear, γ via:

$$\tilde{\gamma} e^{2i\theta} = \frac{l_1^2 - l_2^2 + i2l_1l_2}{l_1^2 + l_2^2}\tilde{\kappa}, \quad (1)$$

where the tildes denote Fourier transforms, l is the wavenumber, and the shear is in the form of the complex number $\gamma = \gamma_1 + i\gamma_2$. For more on the weak lensing shear and convergence, please see Appendix B. We therefore take the Fourier transform of the projected surface density and convert to a vector field. Then, by assuming a background galaxy distribution that is uniform across the sky, and a $z_{\rm source} = 1.0$ and a density of $n_{\rm gal}$/arcmin$^2 = 100$, we interpolate the shear field to individual galaxy positions and derive shear catalogues for each galaxy cluster. Since, we want to directly compare the estimated shape from this probe (and not the expected error bars), we assume a large background density and that each source galaxy is intrinsically circular (i.e. no intrinsic ellipticity).

2.2 Observational data of massive clusters

We use data from the Advanced Camera for Surveys (ACS) onboard the *Hubble Space Telescope* (*HST*) for both the strong and weak lensing analysis, the shape of the distribution of cluster members, and the shape of the BCG, and *Chandra X-Ray Observatory data* (*CXO*) for the X-ray analysis. We use the sample of galaxy clusters from Harvey et al. (2017b) consisting of 10 strong lensing clusters from the Local Cluster Substructure Survey (LoCuSS; Richard et al. 2010) and the Cluster Lensing And Supernova survey with Hubble (CLASH; Postman et al. 2012) that have at least 10 multiple images, ensuring sufficient constraints to measure the lensing parameters that govern the inner region of the cluster, and sufficient imaging by ACS to make a robust weak lensing measurement (i.e. one orbit in either *F*814*W* or *F*850*LP*). Furthermore, these 10 galaxy clusters are required to be relaxed, with no signature of dynamical disturbance as this could bias the shape of the clusters when trying to isolate the impact of physics in the core. We quantify this by measuring their X-ray isophotal concentration, $S = \Gamma_{100\ \rm kpc}/\Gamma_{400\ \rm kpc}$, where Γ is the integrated X-ray flux within a given radius. We apply a strict criteria that the cluster must have $S > 0.2$ (Rasia et al. 2013). Finally of the 10 clusters, Abell1413 does not have sufficient optical imaging in the ACS on *HST* for the weak lensing and AS1063 catalogue is currently in process from the BUFFALO survey (GO-15117; Steinhardt et al. 2020) and therefore will not be ready for another year. As such we have a final sample of eight galaxy clusters. A summary of these clusters can be found in Table 1.

2.2.1 *Strong lensing image selection*

For the strong lensing measurement, we adopt the published multiple image catalogues of *confirmed* images from Zitrin et al. (2015), Richard et al. (2010), and Limousin et al. (2008). The strong lensing model requires knowledge of the redshift of the source. As such, for those images that do not have spectroscopic redshifts we match the sources to publicly available photometric redshift catalogues (Jouvel et al. 2014) and add these source redshifts as parameters in the model with the 1σ error in the photometric redshift as a Gaussian prior.

2.2.2 *Cluster member selections*

In order to select cluster members, we also adopt the cluster member catalogues from Zitrin et al. (2015), Richard et al. (2010), and Limousin et al. (2008), who had identified the red sequence in order to classify the cluster members. For more, please see referenced papers.

2.2.3 *Weak lensing data reduction*

We obtained raw images of each cluster with the associated reference files and reanalyse the data. We first treat each individual exposure for radiation damage due to cosmic rays inducing charge 'traps' in the CCD. During read-out, the trapped charge is rereleased intermittently causing flux to erroneously appear along the read-out axis of the CCD. As a result, we must model this charge transfer inefficiency (CTI) and post-process the image in order to redistribute the charge (Anderson & Bedin 2010; Massey 2010; Massey et al. 2014).

Following this, we used the publicly available CALACS package[1] to recalibrate the individual raw images and then coadd them using the ASTRODRIZZLE package (Koekemoer et al. 2003), accounting for deformations induced by the telescope. During the drizzling process, we use a square kernel and a final pixel scale of 0.03 arcsec per pixel as recommended by Rhodes et al. (2007). For those exposures that are misaligned and taken at different epochs, we use SEXTRACTOR to extract sources from the image and then the TWEAKREG algorithm to realign to a common reference frame. We also use ASTRODRIZZLE on individual exposures to produce deformation free images required by the shape measurement process, PYRRG. Finally, we measure the shapes of all source galaxies using the publicly available code PYRRG (see Appendix B).

2.2.4 *CXO data reduction*

We reprocess all the available raw ACIS-I and ACIS-S *Chandra* data using the publicly available CIAO tools[2] as in Harvey et al. (2015) and the CHANDRA_REPRO script, creating new 'evt2' files (the raw photon tables). This ensures that the data are cleaned with the most up-to-date reference files. Following this, we extract a region of interest and combine the exposures using the MERGE_OBS script, which produces an estimate of the spatially varying point spread function (PSF). We use the wavelets source finder, WAVDETECT[3] with a filter size of 1 and 2 pixels to find point sources within the field that are extremely bright and not associated with the broad emission of the cluster halo. We remove the detected point sources from the field and then verify by eye, removing any residual point sources by hand. Finally, we weight each pixel by the estimated exposure map constructed during the MERGE_OBS script. This provides a clean, exposure time weighted flux map of the cluster.

2.3 Comparability of observations with simulations

The selection function of the CLASH and LoCUSS clusters, although X-ray selected, is complicated. We select a subsample of these clusters that are both relaxed and have a large number of multiple images, as such their typicality can be questioned. Although, we try to simplify the selection function by only choosing those clusters that have concentrated X-ray isophotes (i.e. relaxed), the fact that they have a larger number of multiple images may mean that they are more concentrated than the average cluster and more massive. Moreover, as Fig. 1 shows, the most relaxed massive clusters (solid grey histogram) in the BAHAMAS simulations (at $z = 0.375$) are less massive than the observed sample (red histogram) that may result in biases in the results (where we show the total mass enclosed within a radius at which the density is 200 times the critical density of the universe at that redshift). This should be taken into account when considering the generality of the conclusions (we also show the distribution of unrelaxed cluster in the dashed histogram).

3 METHODOLOGY

In order to investigate our two primary questions, we analyse both the 22 simulated clusters and the eight observed cluster with identical pipelines. Appendix A gives a detailed account of how we measure the shapes from each probe and how we calculate the associated

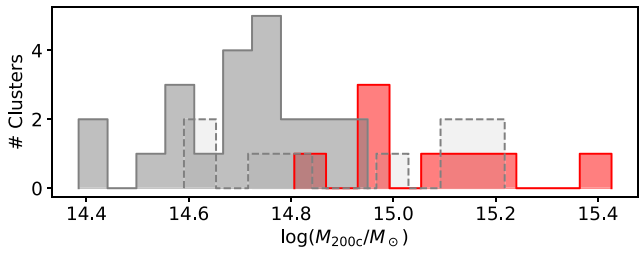

Figure 1. The distribution of M_{200c} of the observed clusters as estimated by strong gravitational lensing (red histogram) and the most massive relaxed (grey solid) and unrelaxed (grey dashed) simulated clusters in the BAHAMAS simulations at a redshift of $z = 0.375$.

errors, here we give a brief outline. In all cases the ellipticity is defined as

$$\epsilon = \frac{a^2 - b^2}{a^2 + b^2} = \frac{1 - q^2}{1 + q^2}, \quad \epsilon_1 = \epsilon \cos(2\theta) \quad \& \quad \epsilon_2 = \epsilon \sin(2\theta), \qquad (2)$$

where a and b are the semimajor and semiminor axes of the ellipse, respectively, $q = b/a$ is the axial ratio, θ is the angle of the major axis of the ellipse northwards from west.

(i) *Brightest cluster galaxy (BCG) (Appendix A1)* :We iteratively measure the flux weighted moment of inertia within a given cut radius. At each iteration, we calculate the ellipticity and position angle from all pixels within an ellipse defined by the previous iteration (initializing at an ellipticity of zero). We set the minimum radial cut to be $r_{\rm cut,\,min} = 20$ kpc to avoid reaching the plummer softening length of the simulation of $4\,h^{-1}$ kpc. Errors are calculated analytically from the fourth-order moments.

(ii) *Cluster member galaxies (Appendix A2)* :Derived from the mass weighted moments of the cluster member galaxies, we measure the shape using all cluster members with total stellar mass $\log(M_\star/M_\odot) > 10^{10}$, matching the mass resolution of the simulations. Since the cluster members are selected by their red sequence, we also select cluster members from the simulations that have no ongoing star formation. Error bars are derived via Monte Carlo realizations of the estimated shape through randomly distributing galaxy cluster members with a Poisson distribution. We measure the shape as a function of $r_{\rm cut}$, a cluster-centric radius within which we include galaxies, and discard any bin with less than 10 cluster members.

(iii) *X-ray isophote (Appendix A3)*: Similar to the BCG, we derive the shape of the X-ray isophote iteratively from the flux weighted image moments, measuring the ellipticity and position angle as function from all X-ray photons within some radius cut. Error bars are derived from 100 Monte Carlo realizations of the Poisson distributed data.

(iv) *Weak gravitational lensing (Appendix A4.1)*: We use the shape measurement code PYRRG and mass mapping algorithm, LENSTOOL to estimate the ellipticity of the cluster from weak gravitational lensing. LENSTOOL is a parametric fitting procedure, whereby we fit Navarro, Frenk, and White (Navarro, Frenk & White 1997) density profiles to the data. The error bars are derived from the width of the posterior during the MCMC fitting procedure.

(v) *Strong gravitational lensing (Appendix A4.2)*: The shape of the mass distribution as estimated from strong gravitational lensing using the mass mapping tool LENSTOOL. We fit an NFW profile for the cluster scale halo along with pseudo-isothermal elliptical mass distriubtions (PIEMD) to each cluster member. We also assume that the cluster members lie of the Fundamental Plane with a constant mass-to-light ratio. Error bars are derived through 20 Monte Carlo

[1] https://acstools.readthedocs.io/en/latest/calacs_hstcal.html
[2] http://cxc.harvard.edu/ciao/
[3] http://cxc.harvard.edu/ciao/threads/wavdetect/

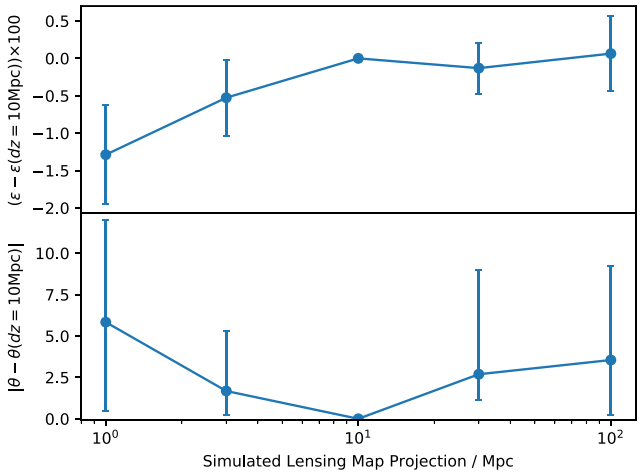

Figure 2. Impact of the cluster environment on the estimated ellipticity and position angle. We project the simulated lensing mass maps to varying depths and measure the weak lensing ellipticity (top panel) and position angle (bottom panel) relative to the fiducial projection of dz = 10 Mpc. We find that the shapes converge above at this fiducial value. We have multiplied the y-axis of the top panel by 100 for clarity.

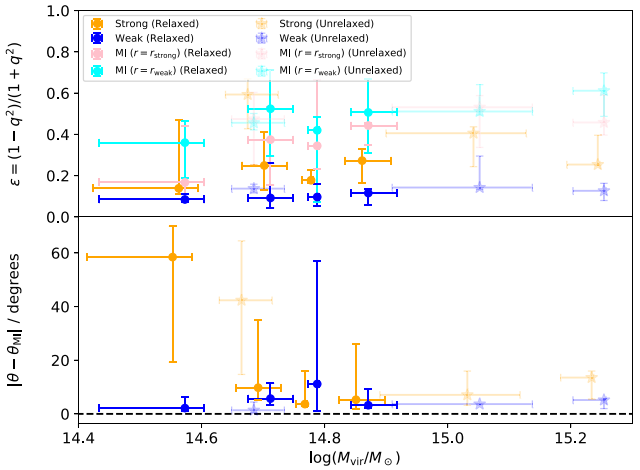

Figure 3. A direct comparison between the ellipticity (top panel) and position angle of the major axis (bottom panel) estimated from the projected moment of inertia (MI) calculated directly from the particle data at either the radius of the strong lensing (the mean radial distance of the multiple images, pink) or the weak lensing (the maximum distance of cluster members in the *HST* image, cyan) and the strong (orange) and weak (blue) lensing via mock observations. The solid spots represent the relaxed clusters with an X-ray concentration of $\Gamma > 0.2$ (see text) and the faded stars represent unrelaxed clusters with an X-ray concentration of $\Gamma < 0.2$. The bottom panel shows the misalignment of the strong and weak lensing mock observation estimates relative to the moment of inertia at the corresponding radius.

realizations of the estimated shape and the multiple image positions (since the width of the posterior was found to be a biased estimator of the error bar in Harvey et al. 2017b).

4 RESULTS

Following the measurement of the ellipticity of each cluster in the observed and simulated sample, we present our results. Before we address our first question, we first test the environmental impact of substructures on the weak lensing shape estimate of the clusters. To do this, we project the extracted lensing maps from the simulation box to varying depths, dz. We show in Fig. 2, the estimated ellipticity relative to the fiducial depth of dz = 10 Mpc in the top panel and the misalignment angle of the cluster relative to the estimate at the fiducial projected depth of dz = 10 Mpc in the bottom panel. We find that at small projected depths, the ellipticity is underestimated and that the ellipticity estimates converge at the fiducial value of dz = 10 Mpc and thus this encapsulates the total information in the environment, and that the misalignment angles remain consistent within the uncertainty.

Having justified the projection depth of our simulations, we begin by addressing our first question: *'Is the ellipticity calculated from the projected moment of inertia derived directly from the particle data in simulations a good estimator of the shape derived from strong or weak lensing'*. Fig. 3 shows the mass-binned results from all 40 clusters in the simulated sample, each point showing the median, 16 per cent and 84 per cent of the distribution, the solid spots represent all those relaxed clusters that have an X-ray concentration $\Gamma > 0.2$ and the faded stars all unrelaxed clusters with $\Gamma < 0.2$. For more on how we calculate the moment of inertia, please see Appendix A5. The top panel shows the ellipticity from four estimators as a function of mass, they include: (1) *Pink,* the 2D ellipticity calculated from the projected moment of inertia (MI) measured on *all* particles within a mean radius at which the mock strong lensing is measured, (2) *Cyan,* the same ellipticity calculated from the moment of inertia using all particles within the outer most radius at which the mock weak lensing is measured (i.e. $r < 0.4r_{\rm vir}$), (3) *Orange,* the ellipticity of the cluster scale halo estimated by mock strong lensing observations, and (4) *Blue,* the ellipticity estimated by mock weak lensing observations. The bottom panel shows the misalignment of the position angle of the major axis between the two lensing estimates and the moment of inertia calculated at their respective radii (i.e. $\theta_{\rm strong} - \theta_{\rm MI,\,strong}$ and $\theta_{\rm weak} - \theta_{\rm MI,\,weak}$).

In general, we find that the ellipticity estimates from the strong and weak lensing differ from the estimate derived directly from the particle data, whereby the strong lensing and projected inertia tensor differ by a factor of $\langle \epsilon_{\rm S}/\epsilon_{\rm S,MI} \rangle = 0.64^{+0.05}_{-0.04}$ and the weak lensing underestimates the ellipticity by a factor of $\langle \epsilon_{\rm W}/\epsilon_{\rm W,MI} \rangle = 0.23^{+0.01}_{-0.01}$. However, the strong lensing estimate for the disturbed clusters seem to be more robust, with a smaller bias of $\langle \epsilon_{\rm S}/\epsilon_{\rm S,MI} \rangle = 0.8^{+0.2}_{-0.1}$ for the entire unrelaxed sample. This is counterintuitive since it would be easier to model a relaxed cluster yet they remain biased.

We find that the position angles in the bottom panel of the respective observable are well aligned, with a mean misalignment angle (in the relaxed sample) of $\langle |\theta_{\rm S} - \theta_{\rm S,MI}| \rangle = 8^{+13}_{-2}$ and $\langle |\theta_{\rm W} - \theta_{\rm W,MI}| \rangle = 4.5^{+2.0}_{-0.9}$ deg. However, the misalignment angle of the strong lensing is dominated by the first low mass, should we remove this we find that the strong lensing agrees much better with a misalignment angle of $\langle |\theta_{\rm S} - \theta_{\rm S,MI}| \rangle = 5^{+2}_{-1}$ deg. We find that for the unrelaxed sample, the strong lensing has a large misalignment angle $\langle |\theta_{\rm S} - \theta_{\rm S,MI}| \rangle = 13^{+11}_{-3}$ deg whereas the weak lensing, even in the unrelaxed sample is good agreement with the moment of inertia $\langle |\theta_{\rm W} - \theta_{\rm W,MI}| \rangle = 3.7^{+0.6}_{-0.9}$ deg. We verify by eye those clusters that have a large difference in the angle and find that massive structures in the core of the cluster can induce huge misalignments. For example, Fig. 4 shows an example cluster from the BAHAMAS simulations whereby the cluster scale dark matter halo as predicted by strong lensing (red contours) is misaligned with the moment of inertia (green dashed ellipse) by ∼80 deg, whereas the weak lensing (blue) is well aligned. We find that this is due to the distribution of strong lensing constraints (yellow stars) that do not fully probe the inner region, and

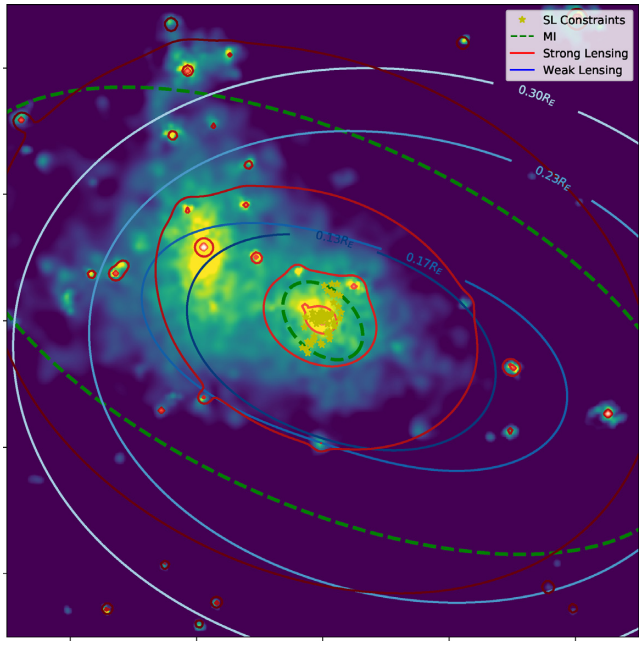

Figure 4. An example when the alignment of *the cluster scale halo* predicted from the strong lensing model (red contours) is misaligned with the moment of inertia (green dashed). In this cluster, the strong lensing constraints (yellow stars) are aligned along the 45 deg angle and will have some perturbation from the mass in the north-west of the strong lensing region, and while the model includes the subhalo to the east (through a scaling relation), the constraints are insensitive to this clump. As a result, the strong lensing predicts a dark matter cluster halo angle of 18 deg, whereas the moment of inertia (from the total matter) is closer to 100 deg resulting in an ∼80 deg misalignment. However, the weak lensing model (blue) is aligned well with the broad mass distribution, agreeing with the moment of inertia.

are likely perturbed by structures within the core. We see a halo in the north-west of the inner region that is likely biasing the strong lensing model, whereas the large halo to the east is outside the constraints and therefore not sensitive to this (although this halo will be included in the model through an assumed mass-to-light scaling relation). As such, it is important that when comparing strong lensing estimates to the moment of inertia, the entire distribution of mass must be taken into account and not just the cluster scale halo.

Finally, we estimate the cut radius for which the moment of inertia best matches the mock weak and strong lensing. To do this, we find that effective radii of the moment of inertia at which the misalignment angle with respect to the weak and strong lensing is the smallest, i.e.

$$r_{\rm eff} = r_{\rm cut}[\min\{\theta - \theta_{MI}(r)\}]. \quad (3)$$

We find that the radius at which the alignment with respect to the weak and strong lensing is a minimum, match the radius at which the weak and strong lensing is measured. i.e. the moment of inertia best matches the weak lensing when measured at the same radius as the weak lensing (and same with the strong lensing).

Clearly in the real, observed case, the unrelaxed clusters would be modelled by multihalo components, with each cluster carefully studied. In this case, we have simply modelled each one with a single cluster halo. However, these findings do still reinforce the need to carry out complete end-to-end comparisons between observations and data.

Following this investigation, we move to directly comparing the observed sample with the simulated sample. As an example of both pipelines, we show the results from a single cluster in each sample in Fig. 5. The top row shows a simulated cluster and the bottom the galaxy cluster A1703. The first column shows the estimated shape of the BCG from the stellar mass map (first row) and the *HST* optical image (second row). The second column shows the estimated shape of the cluster member galaxies, once again from the stellar mass map (first row) and the *HST* optical image (second row). The third column shows the estimated shape of the X-ray isophote overlaid on the X-ray emission map from the simulation (top) and the CXO (bottom), the fourth column shows the best-fitting weak lensing estimate overlaid on the total mass map (top) and the *HST* image (bottom), and the fifth column shows the best-fitting strong lensing model overlaid on the total mass map (top) and the *HST* image (bottom). Each panel has the *x*–*y* axes orientated to the north-west and the scale is given by the white bar. Each ellipse corresponds to a larger cut radius (except for the strong lensing that does not have any cut). We see that for these two examples, the halo shapes as measured from different observational probes correlate with one another.

With these models, we address our second question: *'Is there any evidence for a radial dependent ellipticity in relaxed clusters?'* We start by showing Fig. 6, which presents the ellipticity of each observational probe as a function of the radius cut (normalized to the virial radius of each cluster). The top panel of each figure shows each individually observed cluster (with the legend above the figure), and then the median, 16 per cent and 84 per cent percentile of the simulated sample in the grey shaded regions (with the dashed line showing the median value). The bottom panel of each figure shows the distributions of ellipticity that are relative to the individual cluster at a specific (arbitrary) radii. This allows us to compare trends in the changing shape of a cluster. The top of each panel denotes the respective probe, with BCG in the top left, cluster members in the top right, X-ray in the bottom left, and weak lensing in the bottom right. We do not show strong lensing since it has no measured radial dependence, plus the weak lensing and cluster member estimates for each cluster are offset from one another for clarity. We find in general that the ellipticities in the simulations match that of the observations, however, we do find that the clusters MACSJ1206 and MACSJ1931 seem to be outliers with respect to both the observed and simulated samples. Specifically, we find:

(i) *BCG, top left of Fig. 6*: The ellipticiites of the observed sample matches that of the simulated, however, MACS1206 remains outside the 64 per cent region of the simulations. The error bars of A383 and A1835 are a result of the large amounts of substructure within the BCG isophote. We also find that the central parts of the simulated BCGs tend to be slightly more elliptical than the outer regions.

(ii) *Cluster members, top right of Fig. 6*: We find that the simulated sample is on average consistent with the observed sample. Interestingly A1703, which exhibits no strong ellipticity in other probes has a elliptical distribution of cluster members, whereas MACS1206 shows the opposite with a circular distribution. The median of the simulated haloes suggests that the centre of the cluster has a more spherical distribution of cluster members.

(iii) *X-ray, bottom left of Fig. 6*: The third panel showing the X-rays shows how inner regions of the two clusters, MACSJ1931 and MACSJ1206, are both inconsistent with the simulated haloes, with MACS1206 showing a clear radial dependence not seen in the simulations. We also find that the observed X-ray distribution exhibits a stronger dependence on the cut radius than the simulated sample, with a more elliptical core.

(iv) *Weak lensing, bottom right of Fig. 6*: Finally, we find that the weak lensing shows a similar trend to that of the X-ray and BCG, with the two clusters, MACJ1931 and MACSJ1206, showing a clear

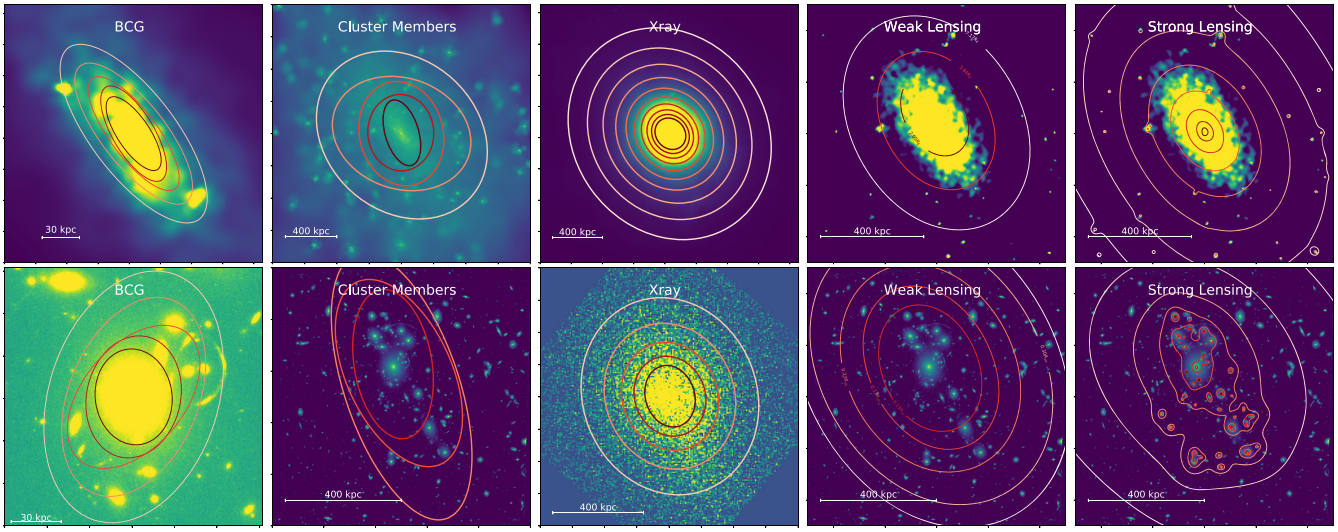

Figure 5. An example of a simulated cluster from the BAHAMAS simulations (top row) and the observed cluster, A1703 (bottom row) and their best-fitting ellipses. In each case, we show the estimated shape of the cluster for different radial cuts (decreasing shade of red). Each column shows the best-fitting shapes of the BCG overlaid on the simulated stellar mass map (top) and *HST* image (bottom), the cluster members overlaid on the stellar mass map (top) and the *HST* image (bottom), the X-ray shape overlaid on the X-ray emission map (top) and the CXO image (bottom), the weak lensing overlaid on the total mass map (top) and the *HST* image (bottom), and the strong lensing overlaid on the total mass map (top), and the *HST* image (bottom), respectively. Each panel has the x–y axes orientated to the north-west and the scale is given by the white bar.

upward trend towards to the inner regions of the halo and significant outliers from the simulated sample. Moreover, three of the observed cluster sample are more elliptical than the entire simulated sample.

Following the measurement of the individual probes as function of radius, we collate each distribution and show the estimated radial dependence of each probe in Fig. 7. The top two panels in the left-hand column shows the ellipticities of the observed clusters and the top two panels in the right-hand column, the ellipticities of the simulated clusters. The top row shows the absolute distributions (i.e. each top panel in the four figures of Fig. 6), with the addition of the estimates from strong lensing (in orange) and the moment of inertia derived from the particle data in black. The second row shows the estimated shapes relative to the strong lensing estimate. The third row shows the misalignment angle of each probe with the strong lensing estimate.

In general, we find that the distributions of observed ellipticities are consistent with one another, with no significant evidence for any radial dependent change in the ellipticity. Moreover, we find that the simulations and observations roughly agree with one another. Finally, we find that the moment of inertia exhibits no radial dependence, with a constant ellipticity over the entire range.

The second row showing the results relative to the strong lensing shows that the observed ellipticities tend to be in agreement with the strong lensing estimate, apart from the BCG, which seems to be more elliptical for both the observed and simulated samples. Interestingly, we find the X-ray and weak lensing estimates are both more spherical than the strong lensing, while the cluster members, BCG, and the moment of inertia are much more elliptical. This is slightly inconsistent with Okabe et al. (2020), who found that the probes all tended to be less elliptical than the strong lensing estimate.

The third row shows the alignment of each probe's major axis with respect to the strong lensing. The dashed line shows the 45 deg line representing to expected mean from random misalignment. We find for both the observed and simulated clusters, all probes are aligned with the dark matter with a misalignment angle of ∼20 deg (depending on the cluster-centric radius). This is again consistent with Okabe et al. (2020), who found a mean misalignment angle of $\theta = 22.2 \pm 3.9$ deg. However, we do find that although the simulations are consistent with the observations, they do suggest a higher variance and random misalignments.

Finally, it would be interesting to study the correlation of each probe with each other and other probes. As such, we carry out a complete test of the correlation of all probes and radial cuts. We define the correlation with the standard Pearson's correlation coefficient

$$r_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x - \bar{x})^2}\sqrt{\sum(y - \bar{y})^2}}, \quad (4)$$

where barred quantities refer to the mean of the x and y samples. Fig. 8 shows the correlation matrix for the ellipticity of each component and radial bin. We divide the matrix along the diagonal between observations (top left) and simulations (bottom right). We have used solid black lines to denote each probe (i.e. X-ray, strong lensing, weak lensing, BCG, and cluster members, or 'gals') and below the figure the colour bar shows the correlation strength, with red showing an anticorrelation and green a positive one. We also correlate the moment of inertia, which is relevant only for the simulations, which is why this has the extra 5×5 matrix in the top left.

The major difference between the two is that the cross-correlations between each component are stronger in the observations than the simulations. This could be due to the fact that observed clusters are larger than the simulated ones and therefore any baryonic feedback in the simulations that could disrupt any correlation would have a larger (fractional) impact. Now specifically looking at each probe, the key points to highlight are

(i) We find tentative evidence that the inner of regions of the observed X-ray halo do not correlate with the outer regions. This is mostly driven by the two cluster MACS1206 and MACS1931 and as such requires a greater sample to confirm, whereas the simulations show strong a correlation throughout. The cross-correlation with other probes shows that the inner regions of the X-ray isophotes

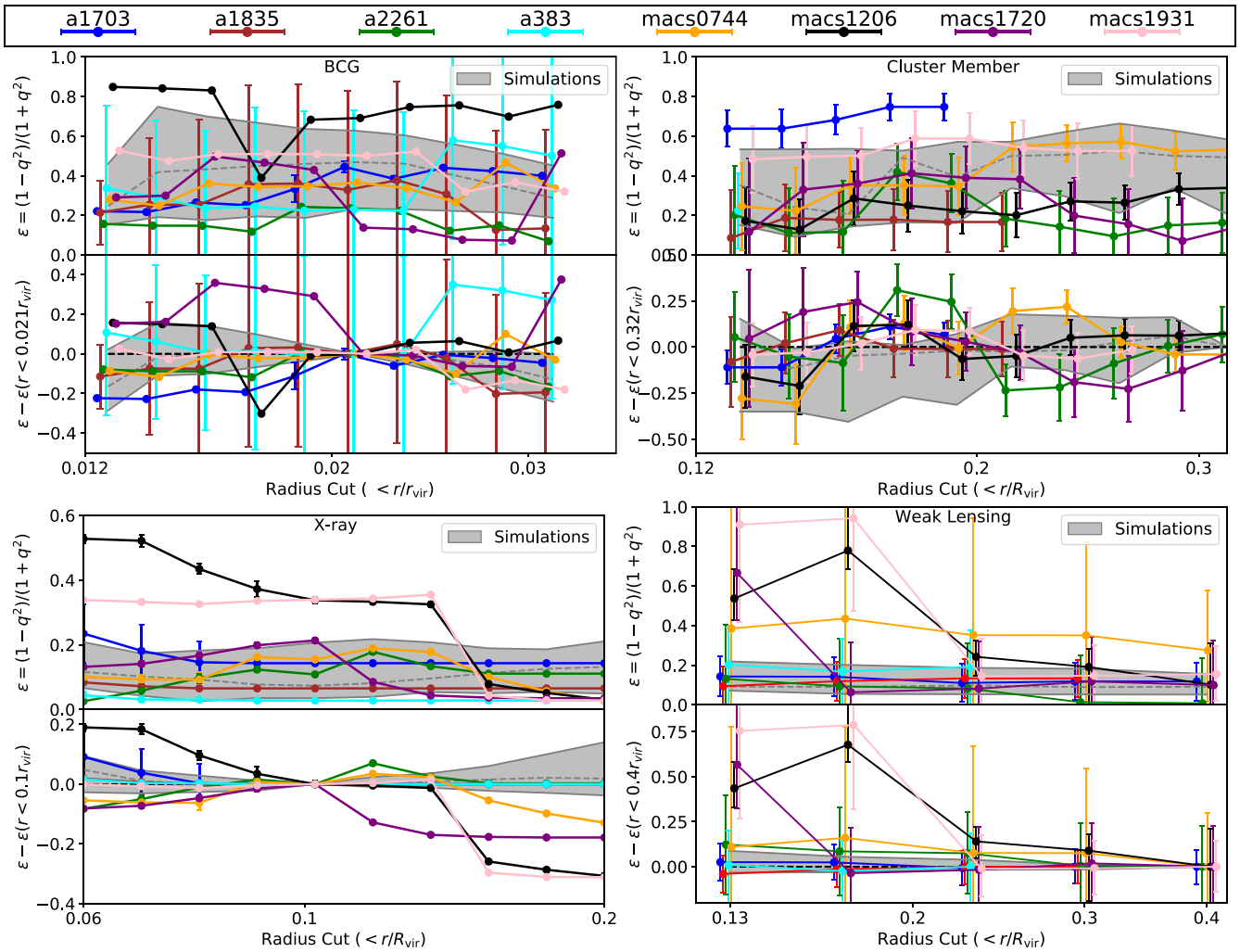

Figure 6. We show the comparison between the observed data set and the BAHAMAS simulations of the ellipticity ($\epsilon = (1 - b^2/a^2)/(1 + b^2/a^2)$) as a function of the radius cut. Each figure shows a different probe of the galaxy cluster (clockwise from top left: BCG; cluster members; weak lensing, X-ray). Within each figure, the top panel shows the absolute of individual clusters, with the legend for each cluster at the top of the main figure. The grey shaded region in each top panel shows the median and the 16 per cent and 84 per cent of the simulated. The bottom panel in each case shows the distributions of each cluster relative to some fixed radius.

correlate with the strong lensing and the smaller scales of the weak lensing. This is interesting if you compare to the simulated haloes, which exhibit the opposite effect, with strong and weak both correlating better with the outer regions of the X-ray. This could be caused by the fact that the outer regions in the observed sample have a large amount of sky background, however, the galaxy member shapes exhibit a strong correlation with the X-ray shapes at these scales, suggesting that this is not the case. Moreover, the simulated X-rays and galaxy distributions exhibit an anticorrelation. This could be due to the fact that the haloes in the observed sample have a cooler core than the simulations, with less thermal motions disrupting the halo or a systematic of the lower of cluster members in these bins. We also find that outer regions of the X-ray are well correlated with the cluster members, something which we also observe in the simulated sample.

(ii) We find that the inner regions of the observed weak lensing correlate strongly with the strong lensing and the BCG, whereas the outer regions do not. This is intuitive since these are sensitive only to the inner regions. We also find that the inner regions of the weak lensing correlate better with the inner regions of the cluster member shapes, however, this is only mild, whereas the simulated weak lensing exhibits a stronger correlation with the cluster member shapes, while a seemingly anticorrelation with the inner parts of the BCG.

(iii) The observed strong lensing shapes unsurprisingly have a strong correlation with the BCG at all scales (including the pre-mentioned correlations with the inner regions of the weak lensing and the X-rays). They are also mildly correlated with the distribution of cluster members. A similar trend is found with the simulations, whereby the strong lensing correlates well with all probes.

(iv) Other than the already mentioned correlations of the BCG (where its strongest correlation is with the inner regions of the clusters), we find that the outer regions of the observed BCG ellipticity is anticorrelated with the X-ray distribution. We do see signs of this anticorrelation with the inner regions of the simulated BCGs, with the outer regions of the BCG consistent with no correlation. What could drive this correlation is not clear and could be noise from the low ellipticity of the cluster members.

(v) Finally, we also correlate the two-dimensional moment of inertia (derived from the projected mass maps) with all simulated

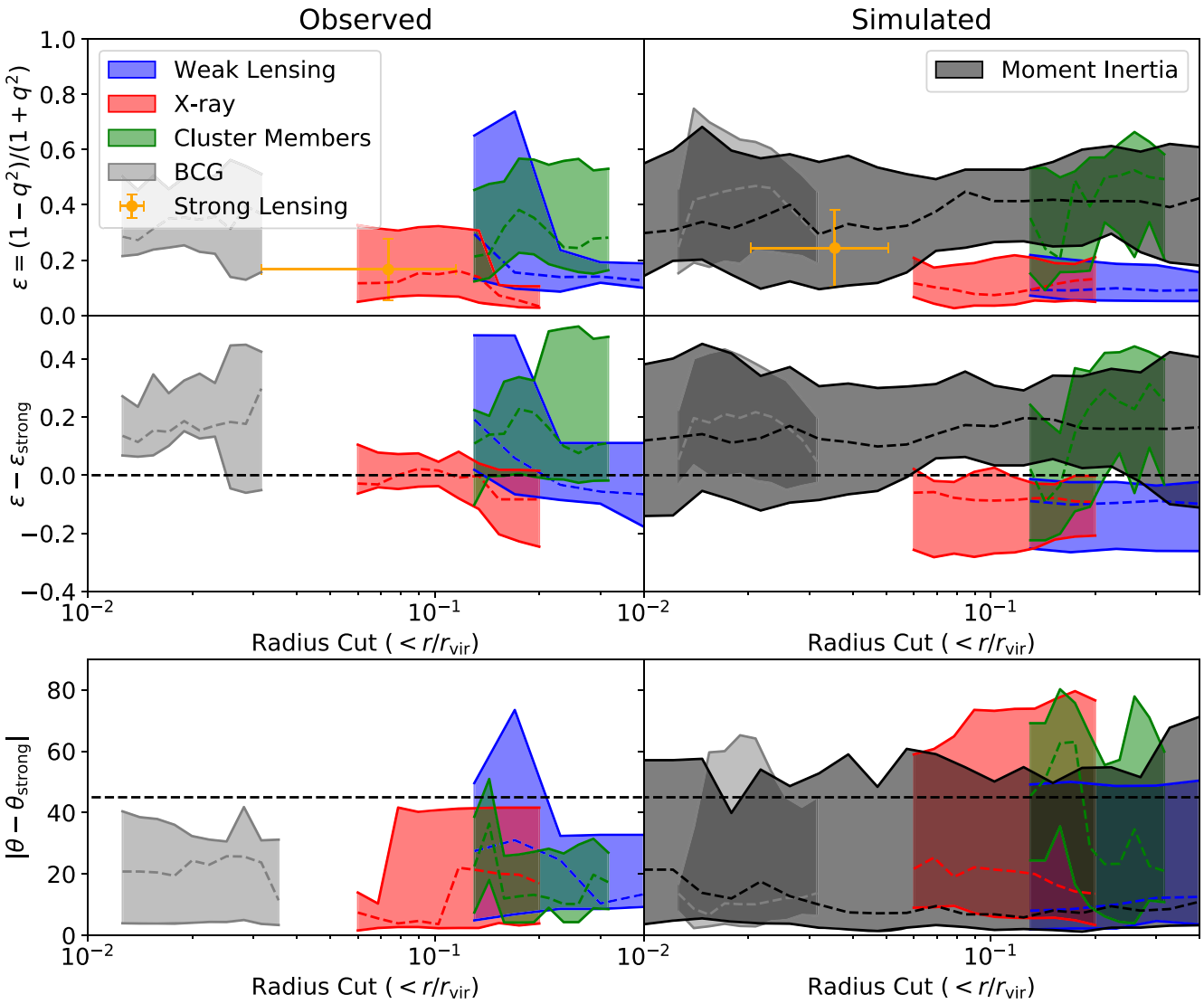

Figure 7. The left (right) hand columns of the top two figures show the median, 84 per cent, and 32 per cent distributions of the observed (simulated) ellipticity from the five different probes. The top row shows absolute values, with grey showing the shape of the BCG, the orange showing the strong lensing (with the width giving the range of virial radii this is measured at), the green shows the ellipticity derived from the cluster members, the blue gives the weak lensing shape, red gives the shape of the X-ray isophote, and we show the moment of inertia derived from the particle data in black. The middle row shows the same probes, except relative to the strong lensing ellipticity. The bottom panel shows the misalignment of each cluster relative to the strong lensing ellipticity estimate. The dashed line at 45 deg represents a random alignment.

probes. We find that as expected, the outer regions of the weak lensing and strong lensing do correlate well with inertia tensor. In addition, we find the outer regions of the X-ray isophote also agree with the moment of inertia. This could be evidence that the inner regions of the hot gaseous halo are dominated by thermal motions, whereas the outer regions do tend to correlate with the shape of the cluster. We also find a mild correlation with the galaxy cluster members. The cluster members being tracers for the underlying haloes should indeed have a correlation, however, this is not as strong as we naively expected.

5 DISCUSSION AND CONCLUSIONS

We have carried out an investigation into the shape of eight dynamically *relaxed* galaxy clusters using a combination the *HST* and the *CXO* and compare them to the 22 most massive clusters in the BAHAMAS simulations in a bid to answer two key questions: (1) Is the ellipticity calculated from the projected moment of inertia derived directly from the particle data in simulations a good estimator of the shape derived from strong or weak lensing? and (2) Is there any evidence for a radial dependent ellipticity in galaxy clusters, potentially signalling physics at different scales?

To answer the first question, we create mock strong and weak lensing observations from the BAHAMAS simulations and compare the ellipticity estimates to the projected moment of inertia calculated directly from the particle information. We find that in all cases, the mock strong and weak lensing observations of the relaxed sample underestimate the ellipticity with respect to the moment of inertia (at the same effective radius) by a factor $\langle \epsilon_{\rm S}/\epsilon_{\rm S,MI} \rangle = 0.64^{+0.05}_{-0.04}$ and $\langle \epsilon_{\rm W}/\epsilon_{\rm W,MI} \rangle = 0.23^{+0.01}_{-0.01}$, respectively. In addition, we find that the

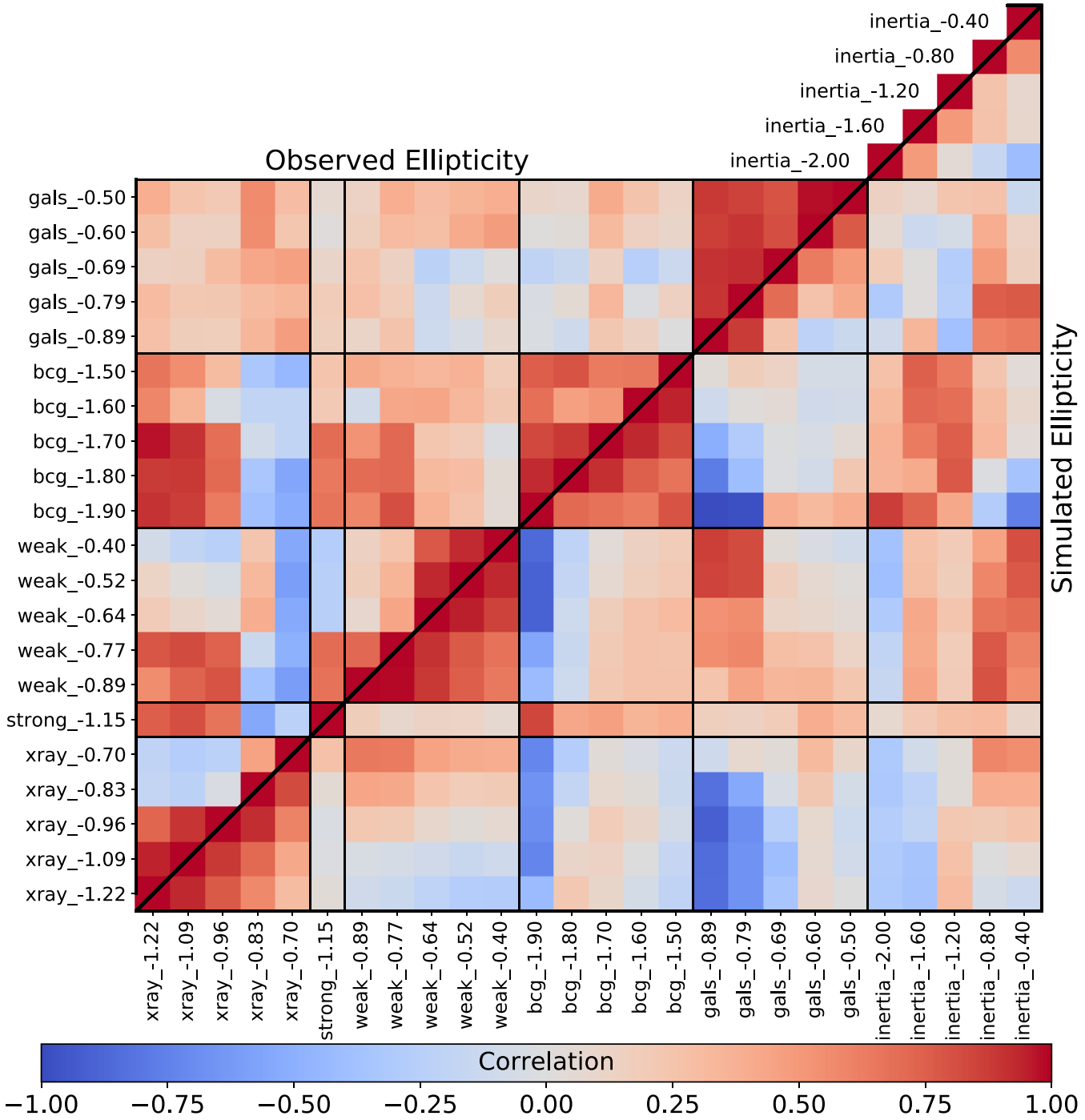

Figure 8. The correlation between each probe and radial cut. The correlation is divided along the diagonal with the top left for the observed clusters and the bottom right for the simulated. Each label shows the type of probe and the log of the radial cut normalized to the virial radius of the cluster. The colour bar based at the bottom shows the correlation range, with red representing an anticorrelation and green a positive correlation. We note that 'gals' corresponds to the shape derived from the distribution of cluster members and 'inertia' the projected ellipticity derived from moment of inertia of the particle data. We do not show the observed moment of inertia as this is not a directly observed quantity.

position angle estimated from the mock weak lensing is well aligned with the moment of inertia for both the relaxed and unrelaxed sample, yet the strong lensing is sensitive to substructures outside the critical curves, particularly for unrelaxed clusters, often inducing large misalignments. These results highlight the importance of deriving mock observations when comparing observations to simulations.

Following this, we then use a combination of BCG, X-ray emission, distribution of galaxy members, weak gravitational lensing, and strong gravitational lensing, creating parallel mock observations in all cases to answer our second primary question.

Studying the radial dependence of the ellipticity for each probe, we find that in general the broad distributions between the simulations and observations match well. The main discrepancies we find is that

the inner regions of MACSJ1206 and MACSJ1931 are much more elliptical than what is predicted by the simulations with the BCG, X-ray isophote, and weak lensing observations, all outliers to the expected distribution of the simulations.

Indeed, it is interesting that MACS1931 and MACS1206 are discrepant with respect to both the rest of the observed clusters and the simulated data set. This is particularly intriguing since they exhibit no special X-ray luminosity. MACS1931 does have a particularly disturbed BCG that could suggest ongoing feedback, however, the BCG of MACS1206 is extremely regular. Having said this, Caminha et al. (2017) found that despite MACS1206 being relaxed, there was significant asymmetry in the total mass profile, which again could suggest ongoing feedback that could impact the shape. This question of the validity of the selection criteria when determining whether a cluster is relaxed or merging and whether this needs to be more sophisticated than just the X-ray concentration, or (more importantly) it stresses the importance of matching selection criteria between simulated and observed samples since there are many microprocess in cluster that could be undetected, which have macroimpacts.

We study the auto- and cross-correlation between each probe and radial cut to understand how each probe relates to one another. We find in general all five mass components of the observed clusters trace the same underlying cluster shape, showing significant correlations between one another. In particular, we find the ellipticity at similar radial cuts strongly correlate. For example, the BCG ellipticity correlates with the strong lensing and the inner regions of the weak lensing, and the outer regions of the X-ray, galaxy members, and weak lensing all correlate with one another.

Interestingly, we do not find the same correlation with the simulated clusters, with trends differing from the observations. The small scales of each probe in these clusters exhibit much weaker correlations, with only the outer regions of the clusters showing the strong correlation seen in the observed clusters. We note that the cluster in the simulated sample is 0.5 dex less massive than the observed sample, which could be a cause of the weaker correlation. In these smaller haloes, the fractional impact of feedback may be larger, disrupting the inner regions and weakening observed correlations. We also find that the cluster member distribution correlates with the X-ray. Finally, we correlate the two-dimensional moment of inertia tensor derived from the projected particle data with each of the simulated probes and find that the weak and strong lensing strongly correlate and the X-ray emission in the outer regions also correlates. We find that the cluster member shapes, although tracers of the underlying structure, only mildly correlate with the shape of the cluster from the projected moment of inertia.

We conclude that weak and strong lensing is a good *proxy* for the moment of inertia derived from the particle data, however, they both significantly underestimate the ellipticity. Therefore, going forward whether weak and strong lensing studies of the shapes of clusters are to be compared to simulations (in a bid to make statements on the nature of dark matter or impact of baryons in clusters), mock observations must be generated and analysed.

Understanding the impact of baryons on massive clusters will be vital if we are to characterize how feedback alters the shape of clusters. Analysis directly comparing data to mock observations like this that probe different regions of the cluster will be important in this effort. Moreover, studies like these where exotic physics or modified gravity may change the shape of a cluster at all scales will also provide important tests of dark matter and general relativity.

This work has been presented in parallel with the public release of our shape measurement code PYRRG. Available to directly install from PyPi via https://pypi.org/project/pyRRG/, this PYTHON 3.7 code based upon Rhodes, Refregier & Groth (2000) is specifically designed for *HST* shape measurement. It is fitted with an automated star–galaxy classifier and outputs scientifically useful products such as catalogues for the mass reconstruction code LENSTOOL. For more, see https://github.com/davidharvey1986/pyRRG.

ACKNOWLEDGEMENTS

DH acknowledges support by the Delta Institute for Theoretical Physics. JR was supported by JPL, which is run under a contract for NASA by Caltech. MJ is supported by the United Kingdom Research and Innovation (UKRI) Future Leaders Fellowship 'Using Cosmic Beasts to uncover the Nature of Dark Matter' [grant number MR/S017216/1]. This project was also supported by the Science and Technology Facilities Council [grant number ST/L00075X/1]. SIT is supported by Van Mildert College Trust PhD Scholarships. RM is supported by the Royal Society. AR is supported by the European Research Council's Horizon 2020 project 'EWC' (award AMD-776247-6). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 769130).

DATA AVAILABILITY

Although data are private, it is available upon request.

REFERENCES

Abbott T. M. C. et al., 2018, Phys. Rev. D, 98, 043526
Anderson J., Bedin L. R., 2010, PASP, 122, 1035
Arnouts S., Ilbert O., 2011, Astrophysics Source Code Library, record ascl:1108.009
Banerjee A., Adhikari S., Dalal N., More S., Kravtsov A., 2020, J. Cosmol. Astropart. Phys., 2020, 024
Bartelmann M., 2010, Class. Quantum Gravity, 27, 233001
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Bennett C. L. et al., 2013, ApJS, 208, 20
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Breiman L., 2001, Mach. Learn., 45, 5
Cacciato M., Bartelmann M., Meneghetti M., Moscardini L., 2006, A&A, 458, 349
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Caminha G. B. et al., 2017, A&A, 607, A93
Cardone V. F., Camera S., Mainini R., Romano A., Diaferio A., Maoli R., Scaramella R., 2013, MNRAS, 430, 2896
Diego J. M., Tegmark M., Protopapas P., Sandvik H. B., 2007, MNRAS, 375, 958
Diego J. M., Broadhurst T., Zitrin A., Lam D., Lim J., Ford H. C., Zheng W., 2015, MNRAS, 451, 3920
Donahue M. et al., 2016, ApJ, 819, 36
Harvey D., Courbin F., 2015, MNRAS, 451, L95
Harvey D., Massey R., Kitching T., Taylor A., Tittley E., 2015, Science, 347, 1462
Harvey D., Kneib J. P., Jauzac M., 2016, MNRAS, 458, 660
Harvey D., Robertson A., Massey R., Kneib J.-P., 2017a, MNRAS, 464, 3991
Harvey D., Courbin F., Kneib J. P., McCarthy I. G., 2017b, MNRAS, 472, 1972
Harvey D., Robertson A., Massey R., McCarthy I. G., 2019, MNRAS, 488, 1572
Hildebrandt H. et al., 2017, MNRAS, 465, 1454
Hoekstra H., Jain B., 2008, Annu. Rev. Nucl. Part. Sci., 58, 99
Jauzac M. et al., 2012, MNRAS, 426, 3369
Jauzac M. et al., 2015a, MNRAS, 446, 4132

Jauzac M. et al., 2015b, MNRAS, 452, 1437
Jauzac M. et al., 2016a, MNRAS, 457, 2029
Jauzac M. et al., 2016b, MNRAS, 463, 3876
Jauzac M. et al., 2018, MNRAS, 481, 2901
Jouvel S. et al., 2014, AAP, 562, A86
Jullo E., Kneib J.-P., Limousin M., Elíasdóttir Á., Marshall P. J., Verdugo T., 2007, New J. Phys., 9, 447
Kahlhoefer F., Schmidt-Hoberg K., Frandsen M. T., Sarkar S., 2014, MNRAS, 437, 2865
Kaiser N., Squires G., 1993, ApJ, 404, 441
Kilbinger M. et al., 2013, MNRAS, 430, 2200
Kneib J.-P., Natarajan P., 2011, A&AR, 19, 47
Koekemoer A. M., Fruchter A. S., Hook R. N., Hack W., 2003, in Arribas S., Koekemoer A., Whitmore B., eds, *HST* Calibration Workshop: Hubble after the Installation of the ACS and the NICMOS Cooling System. p. 337
Kratochvil J. M., Haiman Z., May M., 2010, Phys. Rev. D, 81, 043519
Krist J. E., Hook R. N., Stoehr F., 2011, in Kahan M. A., ed., Proc. SPIE Conf. Ser. Vol. 8127, Optical Modeling and Performance Predictions V. SPIE, Bellingham, p. 81270J
Lau E. T., Nagai D., Kravtsov A. V., Vikhlinin A., Zentner A. R., 2012, ApJ, 755, 116
Leauthaud A. et al., 2007, ApJS, 172, 219
Limousin M. et al., 2008, A&A, 489, 23
Lotz J. M. et al., 2017, ApJ, 837, 97
Marian L., Hilbert S., Smith R. E., Schneider P., Desjacques V., 2011, ApJ, 728, L13
Massey R., 2010, MNRAS, 409, L109
Massey R., Goldberg D. M., 2008, ApJ, 673, L111
Massey R., Kitching T., Richard J., 2010, Rep. Prog. Phys., 73, 086901
Massey R. et al., 2014, MNRAS, 439, 887
McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, MNRAS, 465, 2936
McCarthy I. G., Bird S., Schaye J., Harnois-Deraps J., Font A. S., van Waerbeke L., 2018, MNRAS, 476, 2999
Merten J., Cacciato M., Meneghetti M., Mignone C., Bartelmann M., 2009, A&A, 500, 681
Merten J. et al., 2015, ApJ, 806, 4
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Nelson D. et al., 2018, MNRAS, 475, 624
Oesch P. A. et al., 2010, ApJ, 725, L150
Oguri M., Takada M., Okabe N., Smith G. P., 2010, MNRAS, 405, 2215
Okabe T. et al., 2020, MNRAS, 496, 2591
Parkinson D. et al., 2012, Phys. Rev. D, 86, 103518
Planck Collaboration XVI, 2014, A&A, 571, 66
Postman M. et al., 2012, ApJS, 199, 25
Ragozzine B., Clowe D., Markevitch M., Gonzalez A. H., Bradač M., 2012, ApJ, 744, 94
Rasia E., Meneghetti M., Ettori S., 2013, Astron. Rev., 8, 40
Refregier A., 2003, ARA&A, 41, 645
Rhodes J., Refregier A., Groth E. J., 2000, ApJ, 536, 79
Rhodes J. D. et al., 2007, ApJS, 172, 203
Richard J. et al., 2010, MNRAS, 404, 325
Robertson A., Harvey D., Massey R., Eke V., McCarthy I. G., Jauzac M., Li B., Schaye J., 2019, MNRAS, 488, 3646
Robertson A., Massey R., Eke V., 2020, MNRAS, 494, 4706
Sánchez A. G. et al., 2012, MNRAS, 425, 415
Schaye J. et al., 2015, MNRAS, 446, 521
Schwinn J., Jauzac M., Baugh C. M., Bartelmann M., Eckert D., Harvey D., Natarajan P., Massey R., 2017, MNRAS, 467, 2913
Schwinn J., Baugh C. M., Jauzac M., Bartelmann M., Eckert D., 2018, MNRAS, 481, 4300
Sereno M., De Filippis E., Longo G., Bautz M. W., 2006, ApJ, 645, 170
Smith G. P., Kneib J.-P., Smail I., Mazzotta P., Ebeling H., Czoske O., 2005, MNRAS, 359, 417
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726
Steinhardt C. L. et al., 2020, ApJS, 247, 64
Teague M. R., 1980, J. Opt. Soc. Am., 70, 920
Umetsu K. et al., 2018, ApJ, 860, 104
Umetsu K. et al., 2020, ApJ, 890, 148
Velliscig M. et al., 2015, MNRAS, 454, 3328
Vogelsberger M. et al., 2014, MNRAS, 444, 1518
von der Linden A. et al., 2012, preprint(arXiv:e-prints)
Zitrin A. et al., 2015, ApJ, 801, 44

APPENDIX A: SHAPE MEASUREMENT OF DIFFERENT OBSERVATIONAL PROBES IN DATA

A1 The shape of the brightest cluster galaxy (BCG)

The shape of the BCG can be defined by its first-order image moment,

$$I = \int_0^R d^2\theta i(\theta), \tag{A1}$$

and its normalized quadrupole image moment,

$$J_{ij} = I^{-1} \int_0^R d^2\theta\ \theta_i\theta_j i(\theta), \tag{A2}$$

where $i(\theta)$ is the flux at position θ, $w(\theta)$. From this, the two components of ellipticity of a galaxy, χ_1 and χ_2, are

$$\chi_1 = \frac{J_{11} - J_{22}}{J_{11} + J_{22}}, \qquad \chi_2 = \frac{2J_{21}}{J_{11} + J_{22}}, \tag{A3}$$

where $\chi = (A^2 - B^2)/(A^2 + B^2) = \sqrt{(\chi_1^2 + \chi_2^2)}$ and the size of the object, d, is given by the combination of the quadrupole moments,

$$d = \sqrt{\frac{1}{2}(J_{11} + J_{22})}. \tag{A4}$$

We then find that the error in χ_1 and χ_2,

$$\sigma^2_{\chi_1} = \frac{\sigma^2_{xx}(1 - \chi_1)^2 + \sigma^2_{yy}(1 + \chi_1)^2 - 2(1 - \chi_1^2)\sigma_{xxyy}}{(J_{xx} + J_{yy})^2} \tag{A5}$$

and

$$\sigma^2_{\chi_2} = \frac{(\sigma^2_{xx} + \sigma^2_{yy} + 2\sigma_{xxyy})\chi_2^2 + 4(\sigma^2_{xy} - \chi_2(\sigma_{xxxy} * \sigma_{xyyy}))}{(J_{xx} + J_{yy})^2}, \tag{A6}$$

where σ_{ij} is the error in the given second-order moment,

$$\sigma_{ij} = \frac{\int d^2\theta \sigma_i^2(\theta_i\theta_j - J_{ij})}{\int d^2\theta\ i(\theta)}, \tag{A7}$$

and σ_{ijkl} is the error in the fourth-order moment, J_{ijkl}, where

$$J_{ijkl} = I^{-1} \int d^2\theta\ \theta_i\theta_j\theta_k\theta_l i(\theta), \tag{A8}$$

and

$$\sigma_{ijkl} = \frac{\int d^2\theta \sigma_i^2(\theta_i\theta_j - J_{ij}) * (\theta_k\theta_l - J_{kl})}{\int d^2\theta\ i(\theta)}, \tag{A9}$$

and the error in $i(\theta)$, σ_i, is

$$\sigma_i(\theta)^2 = \sigma^2_{\rm sky} + \frac{i(\theta) - b(\theta)}{t_{\rm exp}}, \tag{A10}$$

where $\sigma_{\rm sky}$ is the estimated variance in the background and $b(\theta)$ is the estimated absolute background. To measure the radial dependence of the shape of the BCG, we measure the total flux within some cut radius. To avoid bias due to the aperture when measuring the shape, we carry out an iterative method whereby we calculate an initial

ellipticity and then moderate the radial distances of each pixel to match the shape, such that

$$r' = \sqrt{\theta_1^2 + \frac{\theta_2^2}{(1-\chi)}}. \quad \text{(A11)}$$

We continue to iterate until the ellipticity converges to within a 1 per cent error. We limit the inner radius to a minimum of $r_{\rm cut} > 20$ kpc to avoid reaching the softening scale of the simulations ($4h^{-1}$ kpc) in both the simulations and the observations, ensuring an equal comparison.

A2 The shape of the distribution of cluster member galaxies

We follow a similar method to estimate the shape of the BCG whereby we measure the intensity weighted image moment, however, with two key differences. To probe the radial dependence, we measure the cumulative number of galaxies inside an cluster-centric cut radius R, and we measure the mass weighted moment similar to the moment of inertia tensor, such that

$$I = \int_0^R d^2\theta M_*(\theta), \quad \text{(A12)}$$

and its normalized, weighted moment quadrupole mass weighted moment,

$$J_{ij} = I^{-1} \int_0^R d^2\theta\ \theta_i\theta_j M_*(\theta), \quad \text{(A13)}$$

where $M_*(\theta)$ is the total stellar mass and R is the cluster-centric radius. We then calculate the two components of ellipticity using equation (A5). For the simulations $M_* = M_*(< 100\text{kpc})\ \text{M}_\odot$, for the observations where the complete spectral energy distribution (SED) is well sampled, i.e. the CLASH clusters (Jouvel et al. 2013), we use LePhare to estimate the stellar mass of cluster members (Arnouts & Ilbert 2011). LePhare is photometric redshift estimator that uses a SED fitting method. Here, we fix the redshift to that of the cluster and constrain only the stellar mass and adopt a Calzetti extinction law (Calzetti et al. 2000). We compare our stellar masses with publicly available catalogues from the CLASH website and find in general a tight correlation, however, due to spurious outliers (stellar masses $>10^{14}$), in the public catalogue we decide to use the derived masses here. Where we do not have the spectral information for the cluster members (a1703 and a1835), we fit an empirical absolute magnitude (in the $F814W$ band) – stellar mass relation and predict these. The top panel of Fig. A1 shows this relation. We see that the six CLASH clusters show a tight correlation and the black dashed line the fitted correlation. The bottom panel shows a histogram of the estimated stellar masses from the CLASH clusters. We see that they span below the mass resolution of the simulations. Therefore, to match the shapes in the simulations and observations, we only estimate the shape from cluster members with $\log(M_*/M_\odot) > 10$. In addition, since we are selecting cluster members from the red sequence, we select only galaxies in the simulations that have negligible star formation.

We derive error bars empirically (since Poisson noise would prevent an accurate estimate using analytical estimates using the fourth-order moments). To do this, we measure the observed shape and then populate a field with the same ellipticity and angle and sample this randomly and remeasure the shape. We do this 100 times and find the upper and lower 34 percentiles around the observed estimate.

Finally, we note that the mass of the clusters in the observed sample is larger than that of the simulated sample and as such we would expect there to be a difference in the number density of cluster members. We hypothesize that the result would be a biased estimator of the shape of the distribution of cluster members in the event there are not many cluster members. As such, we first measure the difference in the subhalo mass functions of the two samples. The top panel of Fig. A2 shows how the number of cluster members in the observed sample is a factor of ∼4 greater than the simulated sample. Moreover, close to the core of the cluster the number of members is small. We test the impact this may have on estimation of the shape. To do this, we carry out a Monte Carlo simulations of a mock sample of cluster members. We distribute a sample of members assuming a Gaussian distribution (with a mean of zero and width of 10 arcsec, although this choice has no impact on the result), with an ellipticity of $\epsilon = 0.4$ and recalculate the ellipticity of the members for different sample sizes. The bottom panel of Fig. A2 shows that below 100 galaxies the shape becomes biased, however, our relatively large statistical error of ∼0.1, means that this does not become significant until ∼10 galaxies. We therefore limit our ellipticity measurements to bins that have at least 10 cluster members.

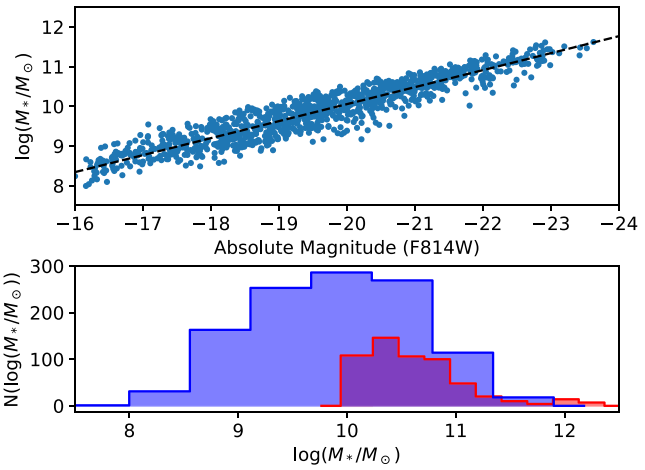

Figure A1. We estimate the stellar mass from the CLASH photometric catalogues (Jouvel et al. 2013) and the publicly available LePhare photometric-redshift estimator in order to estimate the shape of the distribution of cluster members using the moment of inertia tensor (top panel blue dots). For those clusters without high-fidelity photometric catalogues, we fit an empirical relation between $F814W$ absolute magnitude and estimate the stellar mass from their observed magnitudes (top panel dash line). The bottom panel shows the distribution of cluster member stellar masses from observations in blue and from simulations in red. We measure the moment of inertia tensor only from those cluster members with a mass greater than the mass threshold of the simulation ($\log(M_*/M_\odot) > 10$) in order to match the two samples.

A3 The shape of the X-ray isophote

To measure the shape of the X-ray isophote, we adopt the same technique as in the previous section for the galaxy shapes, however, instead of weighting by the mass we weight by the exposure weighted flux. To measure the shape as a function of radial distance, we follow the method laid out in Lau et al. (2012). At each cumulative radial bin, we iteratively estimate the ellipticity by including all pixels inside the elliptical radius,

$$r' = \sqrt{\theta_1^2 + \frac{\theta_2^2}{(1-\chi)}}. \quad \text{(A14)}$$

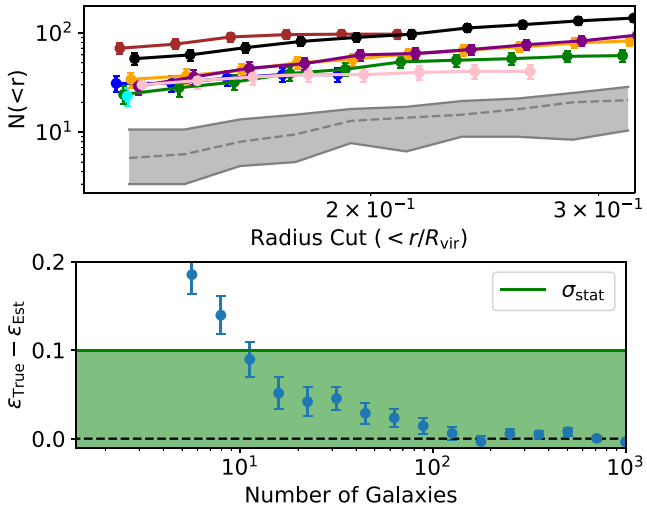

Figure A2. Top Panel: The radial distribution of subhaloes for each cluster (coloured lines) and the simulations (grey region representing the 16 per cent and 84 per cent distribution around the median value, dashed line). We find that the number of subhaloes is a factor of ∼4 greater in the observed sample than the simulated sample attributed to the difference in virial mass. Bottom Panel: The bias on the measured ellipticity of the cluster member distribution. We mock a simulated distribution of cluster members and re-estimate the ellipticity. We show the difference between the input and estimated as a function of the number of galaxy members available to measure the shape off. We find that below 10, the systematic bias becomes larger than the estimated, statistical bias of $\sigma_\epsilon = 0.1$ (green region). We therefore only select bins where there are $n > 10$ galaxies.

At each iteration, we estimate the elliptcity and angle and then re-estimate the shape including those pixels from the previous ellipticity estimate. We stop iterating when the ellipticity converges within an 1 per cent error.

To estimate the error in the ellipticity, we Monte Carlo each cluster image whereby we randomly resample each pixel assuming a Poisson distribution around the true value. We then re-estimate the shape of the cluster. We do this 100 times and take the 32 per cent percentiles around the observed value.

A4 The shape of the cluster from gravitational lensing

We use the publicly available LENSTOOL software that is a mass modelling algorithm that fits realistic parametric models to both strong and weak lensing observables to constrain the properties of the lensing potential (Jullo et al. 2007). It has become commonplace to use LENSTOOL for combined strong and weak lensing analyses (e.g. Jauzac et al. 2016a), however, here we want to analyse the difference in ellipticity between the core and the outer regions of the cluster. In this way, we treat the strong and weak lensing reconstructions completely independently.

A4.1 Weak lensing mass mapping

To derive the weak lensing mass maps, we must first measure the weak lensing shear of all the background source galaxies in the cluster field. We do this via the publicly available code PYRRG (see Section B). With the catalogues, we use LENSTOOL to estimate the weak lensing parameters. It does this by first projecting the observed ellipticities in the image plane to the source plane and then compares the subsequent source plane ellipticities with that expected from a Gaussian distribution with a width equal to the ellipticity dispersion of the sample, i.e.

$$\mathrm{loss} = \sum_{i=1}^{2} \frac{\chi_{i,s}^2}{\sigma_\epsilon}, \tag{A15}$$

where

$$\chi_s = \frac{\chi_- 2g + g^2\chi^\star}{1 + |g|^2 - 2\mathcal{R}(g\chi^\star)}, \tag{A16}$$

where we have assumed that the ellipticity can be written as a complex number in the form $\chi_1 = \chi + i\chi_2$, as produced from PYRRG, the subscript 'S' denotes the ellipticity of the source and the star denotes the complex conjugate. We also adopt the mass and concentration from the strong lensing as a Gaussian prior on the weak lensing mass reconstruction.

A4.2 Strong lensing mass mapping

We follow the same procedure as used in Harvey et al. (2017b) whereby we choose to model the global total matter halo with a Navarro, Frenk and White profile (Navarro et al. 1997) (and hence do not explicitly model the intracluster gas) and model each member galaxy, including the BCG as a PIEMD

$$\rho_{\mathrm{NFW}} \propto \frac{1}{x_{\mathrm{NFW}}(1 + x_{\mathrm{NFW}})^2} \tag{A17}$$

and

$$\rho_{\mathrm{PIEMD}} \propto \frac{1}{(1 + x_{\mathrm{core}}^2)(1 + x_{\mathrm{cut}}^2)}, \tag{A18}$$

where $x_{\mathrm{NFW}} = r/r_{\mathrm{s}}$, where $r_{\mathrm{s}} = r_{\mathrm{vir}}/c_{\mathrm{vir}}$, $x_{\mathrm{core}} = r/r_{\mathrm{core}}$, $x_{\mathrm{cut}} = r/r_{\mathrm{cut}}$ and r is the three-dimensional cluster centric radius. Indeed, LENSTOOL models the two-dimensional NFW potential, Ψ and not the mass distribution, Σ and then assumes that $\epsilon_\Sigma = 3\epsilon_\Psi$. All ellipticities we report here are ellipticities of the mass distribution. For more, we direct the reader to Jullo et al. (2007). We also assume that the member galaxies fall on the Fundamental Plane (including the BCG) following a consistent mass-to-light ratio in order to reduce the parameter space such that for the ith cluster member,

$$r_{\mathrm{core,i}} = r^\star_{\mathrm{core}} \left(\frac{L}{L^\star}\right)^{1/2}, \tag{A19}$$

and

$$r_{\mathrm{cut,i}} = r^\star_{\mathrm{cut}} \left(\frac{L}{L^\star}\right)^{1/2}, \tag{A20}$$

and that the velocity dispersion of the galaxy is

$$\sigma_i = \sigma^\star \left(\frac{L}{L^\star}\right)^{1/4}. \tag{A21}$$

As is common amongst strong lensing reconstructions, we assume $r^\star_{\mathrm{core}} = 0.15$ kpc and we have a tight Gaussian prior of $\sigma^\star = 158 \pm 26$ km s^{-1} and $r^\star_{\mathrm{cut}} = 45 \pm 1$ kpc. Following this, we have six free NFW parameters from the main halo, and then two free parameters for the galaxy members. In rare cases, we model individual galaxies as not doing this has shown to potentially bias mass reconstructions (Harvey, Kneib & Jauzac 2016).

A4.3 Ellipticity error validation

Harvey et al. (2017b) found that the error estimate from LENSTOOL on the Cartesian position of a dark matter halo using strong gravitational

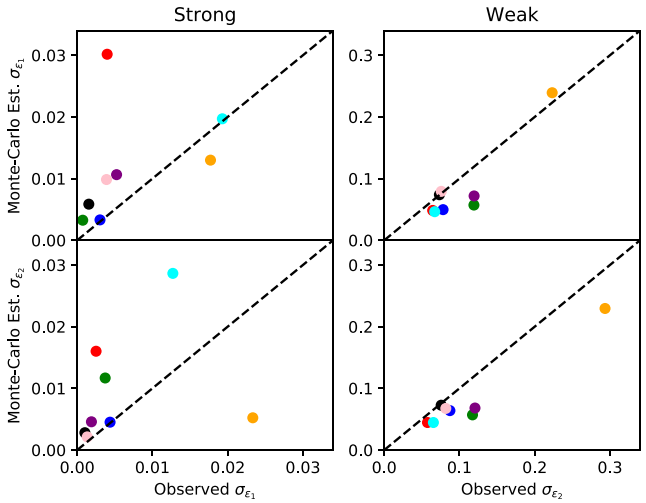

Figure A3. Gravitational lensing error validation. We Monte Carlo the best-fitting strong (left column) and weak (right column) lensing model 20 and 100 times respectively. We show the error from the estimated posterior on the x-axis and the variance in the Monte Carlo on the y-axis for χ_1 in the top panel and χ_2 in the bottom panel.

was underestimated by a minimum of an order of magnitude. In order to quantify whether or not the error reported using the width of the posterior distribution reasonably reflects the true error in the ellipticity we carry out two tests; one for the strong lensing observables and one for the weak lensing.

To do this, we follow the same method as Harvey et al. (2017b). We mock up 20 simulations based on the true data. Using the source positions from the data (for both weak and strong lensing), we use the best-fitting mass model from the strong lensing reconstruction and project the sources to image positions to give a catalogue of weak and strong lensing image positions. In the case of the weak lensing, we add noise through the random distribution of galaxy shapes, modelled by a Gaussian, with a mean of zero and a width equal to that of the true cluster. For the strong lensing, we randomly shift the position, with the shift drawn from a Gaussian with a zero mean and a width of 0.5 arcsec. We then reconstruct the mass distribution using LENSTOOL. We Monte Carlo each cluster 100 times for the weak lensing and 10 for the strong, since the strong lensing reconstructions take significantly longer to converge.

We find that in the weak lensing case, the broad posterior is biased when ellipticity is sampled in polar coordinates, as such we sample the data in Cartesian coordinates; ϵ_1 and ϵ_2, where we defined $\epsilon = (a^2 - b^2)/(a^2 + b^2)$ and a and b are the semimajor and minor axes of the mass distribution, respectively. We then convert the results back to polar coordinates to be consistent with other probes. Fig. A3 shows the results of the strong and weak lensing reconstructions. Each panel shows the estimate of the variance in the Monte Carlo tests as a function of the direct estimate of the error from the width of the posterior sampled during the MCMC of the true data. The left (right) hand column shows the results from the strong (weak) lensing. We find that the estimate of the error from the posterior in the strong lensing significantly underestimates the true error from the Monte Carlo tests, whereas the weak lensing the posteriors slightly overestimates the error. As such during the analysis we take the error in the strong lensing from the variance in the Monte Carlo tests, and the weak lensing directly from the posterior.

A5 Shape from the moment of inertia using simulated particle data

Unlike some of our observational shape definitions, our moment of inertia calculation does not iteratively fit for the centre of the halo. Instead, the 2D centre of the halo is defined as the location (in projection) of the particle with the most negative gravitational potential energy, and this is kept fixed throughout the calculation. The moment of inertia calculation is done in 2D, having first projected all particles within 5 r_{200} of the cluster centre (in 3D) along the relevant line of sight. For a given radius, $r_{\rm cut}$, we begin by finding all particles in a circle of radius $r_{\rm cut}$. The inertia tensor

$$I_{ij} \equiv \sum_n x_{i,n}\, x_{j,n}\, m_n \Big/ \sum_n m_n \tag{A22}$$

is calculated for this distribution of particles, where $(x_{1,n}, x_{2,n})$ are the coordinates of the nth particle, which has mass m_n. We label the eigenvalues of I_{ij} as a^2 and b^2 (with $a \geq b$), and the corresponding eigenvectors as $\boldsymbol{e_1}$ and $\boldsymbol{e_2}$. The axial ratio is $q = b/a$.

Our process is iterative, and in each iteration we calculate the inertia tensor for the particles within an ellipse, where the axial ratio of the ellipse is determined by q from the previous iteration, and the major axis of the ellipse is along the $\boldsymbol{e_1}$ direction. The area of the ellipse is kept constant, meaning that it has a semimajor axis of length $r_{\rm cut}/\sqrt{q}$ and a semiminor axis of length $r_{\rm cut}\sqrt{q}$. We continue this iterative process until consecutive iterations agree on q to better than 1 per cent.

APPENDIX B: SHAPE MEASUREMENT ALGORITHM: PYRRG

In this section, we briefly outline the publicly available shape measurement software, specifically designed for images from the ACS on the *HST*. We begin by first outlining the theoretical basics. For a review, please see Bartelmann & Schneider (2001), Massey, Kitching & Richard (2010), Hoekstra & Jain (2008), and Refregier (2003). For a review on strong and weak gravitational lensing, please see (Bartelmann 2010).

Gravitational lensing is simply a distortion of a background source at a position, β by foreground matter, inducing a shift of $\hat{\alpha}$ to the observed position θ, i.e.

$$\beta = \theta - \frac{D_{\rm LS}}{D_{\rm S}}\hat{\alpha} = \theta - \alpha, \tag{B1}$$

where we have introduced the reduced deflection angle, α. This deflection angle is related to the potential causing the deflection, which is the projected three-dimensional Newtonian potential, Φ,

$$\nabla\Psi = \alpha, \tag{B2}$$

where

$$\Psi = \frac{D_{\rm s}}{D_{\rm L} D_{\rm S}}\frac{2}{c^2}\int \Phi(D_{\rm L}, \theta, z) dz. \tag{B3}$$

From this we can derive the distortion matrix by examining how a change in the source position effects the change in the image position, i.e. $\partial\theta/\partial\beta$, also known as the lensing Jacobian,

$$A_{ij} = \frac{\partial\beta}{\partial\theta} = \delta_{ij} - \frac{\partial^2\Psi}{\partial\theta_i\theta_j} = \delta_{ij} - \Psi_{ij}, \tag{B4}$$

where we have denoted the second derivative of Ψ by the subscript, i and j. The second derivative of the lensing potential gives the two

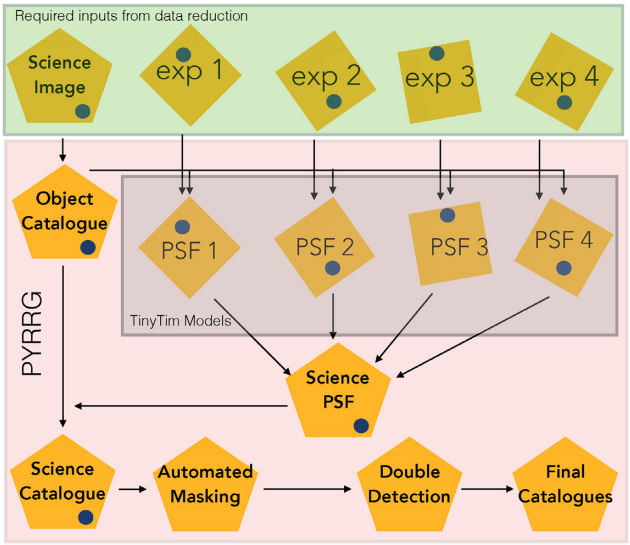

Figure B1. An overview of the PYRRG algorithm. It requires the input science image and associated weight file from the data reduction pipeline, plus all the associated exposures. From this it chooses the 'best' PSF from the TinyTim models and combines them to produce a PSF at the position of each galaxy in the catalogue. It then corrects the galaxies, calculate the shears and outputs a science catalogue. It then carries out post-processing procedures to create a final 'clean' catalogue.

observables, the convergence, which is the trace of A,

$$\kappa = \frac{1}{2}(\Psi_{11} + \Psi_{22}), \tag{B5}$$

and corresponds to a scalar increase or decrease in the size of a distorted background source and the shear, γ is a two-component vector field given by

$$\gamma_1 = \frac{1}{2}(\Psi_{11} - \Psi_{22}) \quad \text{and} \quad \gamma_2 = \Psi_{12} = \Psi_{21}, \tag{B6}$$

corresponding to a stretch along the x-axis for γ_1 and 45° for γ_2. We now have a relation between the observable distortion and the lensing potential. Here we limit the expansion of the Jacobian to first order, and hence assume a weak lensing limit. Indeed the shear and convergence are coupled and one cannot be observed without the other, this is known as the reduced shear,

$$g = \gamma/(1 - \kappa). \tag{B7}$$

B1 Shape measurement: PYRRG

The weak lensing shape measurement consists of six key sections. An overview of the PYRRG algorithm can be found in Fig. B1.

(i) Source finding
(ii) Moment measuring
(iii) Star–galaxy classification
(iv) PSF estimation
(v) Shear estimation
(vi) Catalogue cleaning and masking

B2 Source finding

PYRRG employs the 'hot and cold' method that was originally developed in Leauthaud et al. (2007) to extract sources from the image and then extended to studies used in Jauzac et al. (2012, 2015b, 2016a, 2018) and Harvey & Courbin (2015). Using the open source program SEXTRACTOR (Bertin & Arnouts 1996), PYRRG carries out two scans of the image. The first, 'hot' scan, uses a smaller minimum number of pixels to count as a source, thus finding smaller objects. The second, 'cold' scan, uses a larger number of pixels to classify a source. We then use the publicly available STILTS[4] software to combine the two catalogues in to one final catalogue.

B3 Moment measurement

Following the source detection, PYRRG measures the weighted multipole moments of each object in order to characterize the shape. For a full description please see Rhodes et al. (2000), however here we outline the basics. We define the zeroth order multiple moment of a two-dimensional image in θ, with an intensity distribution i,

$$I = \int d^2\theta\, w(\theta) i(\theta), \tag{B8}$$

then the quadruple *normalized*, weighted moment is

$$J_{ij} = I^{-1} \int d^2\theta\ \theta_i \theta_j\, w(\theta) i(\theta), \tag{B9}$$

followed by the fourth order,

$$J_{ijk} = I^{-1} \int d^2\theta\ \theta_i \theta_j \theta_k\, w(\theta) i(\theta), \tag{B10}$$

where the weight is simply a Gaussian with a width w, where $w = \sqrt{(A_{\rm SEX} B_{\rm SEX}/\pi)}$, AND A and B are the semimajor and semiminor axes as estimated by SEXTRACTOR. From this we can define the two components of ellipticity of a galaxy, χ_1 and χ_2, as

$$\chi_1 = \frac{J_{11} - J_{22}}{J_{11} + J_{22}}, \qquad \chi_2 = \frac{2J_{21}}{J_{11} + J_{22}}, \tag{B11}$$

and the size of the object, d, is given by the combination of the quadrupole moments,

$$d = \sqrt{\frac{1}{2}(J_{11} + J_{22})}. \tag{B12}$$

B4 Star–galaxy classification: random forest

Following the measurement of the normalized image moments, we classify objects in to three distinct categories, stars (both saturated and not), galaxies and noise. Given that it is a simple classifying problem, we adopt a Random Forest to automatically classify this.

A Random Forest is a supervized machine learning tool that generates an ensemble of decision trees that are then trained on known data to produce predictions for unknown data (Breiman 2001). It generates a single tree by randomly subsampling the data and carrying out simple regression to create an estimator for the subsample of data. It then resamples randomly with replacement to generate another tree. Given that each tree is a poor unbiased estimator of the truth, the aggregated estimator should be the correct one. The number of trees defines how good the overall estimator is, but also how long it takes to train and how large the classifier is.

We generate a range of data to train the Random Forest. We use data from *HST* including a sample of 21 SLACS galaxies, 29 galaxy clusters, all at a range of depths. This way we try to span the entire range of parameter space including, object magnitudes, environment, and signal to noise. We generate the ground truth by manually

[4]http://www.star.bris.ac.uk/~mbt/stilts/

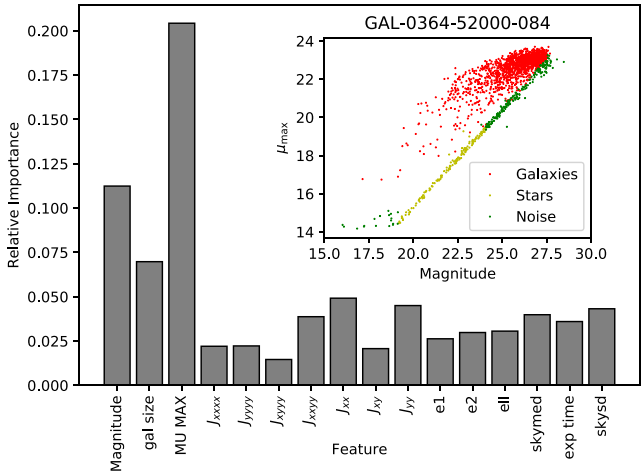

Figure B2. Results of the automated star–galaxy classifier. We use a Random Forest trained on data from the *HST* to predict the classification of each object. The main panel shows the relative importance of each feature in the classifier, the inset shows the results of classifying objects for the object *HST* field with GAL-0364-52000-084. We find that the Random Forest has a 93 per cent rate of correctly classifying stars, 99 per cent rate of classifying galaxies and 83 per cent of characterizing noise (including saturated stars).

classifying stars and galaxies from their magnitudes, $\mu_{\rm max}$, and size for individual exposures. We then aggregate this data and parse through the Random Forest all information regarding these objects including the object magnitude, the size of the object, the brightest pixel in the object ($\mu_{\rm max}$), the second and fourth order moments, the uncorrected ellipticities, the median sky background and the variance around this, and finally the exposure time of the image. The main panel of Fig. B2 shows the relative importance of each feature in the classifier, where relative importance is a unit-less coefficient that defines the importance of each feature in discriminating between classifiers, i.e. when the data are subsampled, each decision tree will define the most important classifier for that subsample. Over the entire ensemble of decision trees, this importance reflects at which level in the decision tree this feature lies. The inset shows the result on a blind test galaxy, GAL-0364-52000-084. We find that the Random Forest has a 93 per cent rate of correctly classifying stars, 99 per cent rate of classifying galaxies and 83 per cent of characterizing noise (including saturated stars).[5]

B5 Point spread function measurement

Having classified the stars, PYRRG then estimates the impact of the telescope on the image, i.e. the PSF. *HST* warms up and cools down due to the heating of the Sun and therefore the focus of the telescope changes over time. We therefore estimate the impact of this by

(i) Taking the known positions of stars from the drizzled science image and finding the corresponding position in the individual exposures that make up that image.

(ii) Measuring the second and fourth order moments of the stars in each of the *individual exposures*

[5]PYRRG allows manual selection of galaxies through an interactive region selecting scheme, however, the default and what is used for this work is the Random Forest.

(iii) Having measured the moments, we compare to the various Tiny Tim models of the PSF (Krist, Hook & Stoehr 2011). We then interpolate this model to the known positions of the galaxies.

(iv) Combining the PSFs from each individual exposure at the position of the galaxy by rotating each PSF moment through an angle ϕ to the same reference frame as the drizzled science image in order to find the new rotated moment, $J^{'}$, (Teague 1980),

$$J'_{jk} = \sum_{r=0}^{j}\sum_{s=0}^{k}(-1)^{k-s}\binom{j}{k}\binom{k}{s} \times(\cos\phi)^{j-r+s}(\sin\phi)^{k+r-s}(J_{j+k-r-s,r+s}), \tag{B13}$$

and then summing the moments for a given position of the galaxy to acquire the final PSF.

B6 Shear estimation

Following the estimation of the PSF, we then correct the galaxy moments and calculate the shear. As shown in Rhodes et al. (2000), the final estimated shear is given by

$$\gamma_i = \langle\chi_i\rangle/G, \tag{B14}$$

where χ_i is given by equation (A3) and with

$$G = 2 - \langle\chi^2\rangle - \frac{1}{2}\langle\lambda\rangle - \frac{1}{2}\langle\chi\cdot\mu\rangle, \tag{B15}$$

where $\langle\chi\cdot\mu\rangle = \chi_1\mu_1 + \chi_2\mu_2$,

$$\lambda = (J_{1111} + 2J_{1122} + J_{2222})/(2d^2w^2), \tag{B16}$$

and the two components of the spinor, μ,

$$\mu_1 = (-J_{1111} + J_{2222})/(2d^2w^2),$$
$$\mu_2 = -2(J_{1112} + J_{1222})/(2d^2w^2), \tag{B17}$$

where w is the size of the weight function $w(\theta)$ in equation (B8). From this we have a final estimator of the shear, γ.

B7 Catalogue cleaning and masking

Having measured the shear we go through a series of cleaning operations including,

(i) *Automatic masking:* Using the known position of stars and saturated stars, we generate polygons that have the same size as stars and mask any object that lies within these polygons,

(ii) *Removal of double detections:* We remove double detections whereby removing objects that lie within the isophote of a larger object.

(iii) *Creation of a* LENSTOOL*catalogue:* Creates a catalogue to be parsed in to the mass mapping algorithm LENSTOOL.

Finally we match the catalogues with the CLASH catalogues, which have accurate photometric redshifts (Jouvel et al. 2013) in order to remove contaminations. This includes cluster members and sources that are in the strong lensing regime (and hence cannot be used for the weak lensing estimates).

We first remove all galaxies that lie within the cluster. To do this we remove all galaxies that lie at $z > z_{\rm cl} + \delta z$. The value of δz depends on the cluster, Umetsu et al. (2019) used a value of $\delta z = 0.2$. Here we test the impact of this cut on our sample and cluster shapes. Fig. B3 shows the estimated ϵ_1 (ϵ_2) shown in the top (bottom) panel relative to the estimated shape with a cut of $\delta z = 0.05$. We find that in general the clusters are not sensitive to this cut except MACSJ1206, which exhibits significant sensitivity. We find that a small cluster of

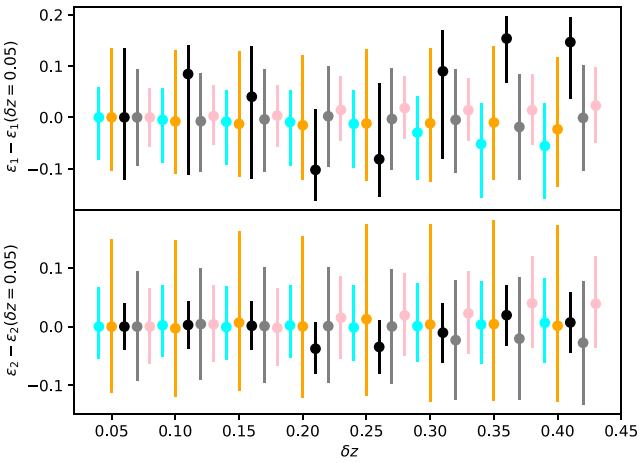

Figure B3. The dependence of estimate of the cluster shape from weak lensing on the source galaxy redshift cut. We cut the galaxy such that only those galaxies in the catalog are $z_s > z_{\rm cluster} + \delta z$. The top panel shows the first component of ellipticity ϵ_1 and the bottom panel ϵ_2. We show for each cluster (different colours) with respect to the shape estimate for $\delta z = 0.05$. We offset the clusters at each cut value for clarity. We find that the only cluster that is sensitive to this choice is MACSJ1206, where a small cluster in redshift space at ∼0.7 drives the fit.

galaxies at $z \sim 0.7$ is pushing the fit to negative ϵ_1. We therefore select a cut of $\delta z = 0.05$ for clusters except MACSJ1206 where we choose $\delta z = 0.35$. We then remove all source galaxies that are within the strong lensing region. Massey & Goldberg (2008) found that moment based methods such as RRG are accurate to within 7 per cent beyond the Einstein radius of the cluster at which point by galaxies being to be strongly lensed and weak lensing assumptions break down. We therefore conservatively extract all galaxies up to 10 per cent beyond the estimated Einstein radius of the cluster to ensure we are not including arcs and flexed galaxies.

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