The non-gravitational interactions of dark matter in colliding galaxy clusters

David Harvey[1,2*], Richard Massey[3], Thomas Kitching[4],
Andy Taylor[2], Eric Tittley[2]

[1]Laboratoire d'astrophysique, EPFL, Observatoire de Sauverny, 1290 Versoix, Switzerland
[2]Royal Observatory, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK
[3]Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK
[4]Mullard Space Science Laboratory, University College London, Dorking, Surrey RH5 6NT, UK
[*]To whom correspondence should be addressed; E-mail: david.harvey@epfl.ch

Collisions between galaxy clusters provide a test of the non-gravitational forces acting on dark matter. Dark matter's lack of deceleration in the 'bullet cluster collision' constrained its self-interaction cross-section $\sigma_{\mathrm{DM}}/m < 1.25 \, \mathrm{cm}^2/\mathrm{g}$ (68% confidence limit) for long-ranged forces. Using the Chandra and Hubble Space Telescopes we have now observed 72 collisions, including both 'major' and 'minor' mergers. Combining these measurements statistically, we detect the existence of dark mass at 7.6σ significance. The position of the dark mass has remained closely aligned within 5.8 ± 8.2 kpc of associated stars: implying a self-interaction cross-section $\sigma_{\mathrm{DM}}/m < 0.47 \, \mathrm{cm}^2/\mathrm{g}$ (95% CL) and disfavoring some proposed extensions to the standard model.

Many independent lines of evidence now suggest that most of the matter in the Universe is in a form outside the standard model of particle physics. A phenomenological model for cold dark matter (*1*) has proved hugely successful on cosmological scales, where its gravitational influence dominates the formation and growth of cosmic structure. However, there are several

challenges on smaller scales: the model incorrectly predicts individual galaxy clusters to have more centrally concentrated density profiles (*2*), larger amounts of substructure (*3, 4*), and the Milky Way to have more satellites able to produce stars (*5*) than are observed. These inconsistencies could be resolved through astrophysical processes (*6*), or if dark matter particles are either warm (*7*) or self-interact with cross-section $0.1 \leq \sigma_{\mathrm{DM}}/m \leq 1\,\mathrm{cm}^2/\mathrm{g}$ (*8–10*). Following (*11*), we define the momentum transfer per unit mass σ_{DM}/m, integrating over all scattering angles and assuming that individual dark matter particles are indistinguishable.

Self-interaction within a hidden dark sector is a generic consequence of some extensions to the standard model. For example, models of mirror dark matter (*12*) and hidden sector dark matter (*12–16*) all predict anisotropic scattering with $\sigma_{\mathrm{DM}}/m \approx 1\,\mathrm{barn}/\mathrm{GeV} = 0.6\,\mathrm{cm}^2/\mathrm{g}$, similar to nuclear cross-sections in the standard model. Note that couplings within the dark sector can be many orders of magnitude larger than those between dark matter and standard model particles, which is at most of order picobarns (*17*).

In terrestrial collider experiments, the forces acting on particles can be inferred from the trajectory and quantity of emerging material. Collisions between galaxy clusters, which contain dark matter, provide similar tests for dark sector forces. If dark matter's particle interactions are frequent but exchange little momentum (via a light mediator particle that produces a long-ranged force and anisotropic scattering), the dark matter will be decelerated by an additional drag force. If the interactions are rare but exchange a lot of momentum (via a massive mediator that produces a short-ranged force and isotropic scattering), dark matter will tend to be scattered away and lost (*11, 18, 19*).

The dynamics of colliding dark matter can be calibrated against that of accompanying standard model particles. The stars that reside within galaxies, which are visible in a smoothed map of their optical emission, have effectively zero cross-section because they are separated by such vast distances that they very rarely collide. The diffuse gas between galaxies, which is visible

in X-ray emission, has a large electroweak cross-section; it is decelerated and most is eventually stripped away by ram pressure (*20*). Dark matter, which can be located via gravitational lensing (*21*), behaves somewhere on this continuum (Fig. 1).

The tightest observational constraints on dark matter's interaction cross-section come from its behavior in the giant 'bullet cluster' collision 1E0657-558 (*22*). A test for drag yields $\sigma_{\mathrm{DM}}/m < 1.25\,\mathrm{cm}^2/\mathrm{g}$ (68% CL), and a test for mass loss yields $\sigma_{\mathrm{DM}}/m < 0.7\,\mathrm{cm}^2/\mathrm{g}$ (68% CL) (*18*). Half a dozen more galaxy cluster collisions have since been discovered, but no tighter constraints have been drawn. This is because the analysis of any individual system is fundamentally limited by uncertainty in the 3D collision geometry (the angle of the motion with respect to our line of sight, the impact parameter, and the impact velocity) or the original mass of the clusters.

The same dynamical effects are also predicted by simulations in collisions between low-mass systems (*11*). Observations of low-mass systems produce noisier estimates of their mass and position (*23–25*), but galaxy clusters continually grow through ubiquitous 'minor mergers', and statistical uncertainty can be decreased by building a potentially very large sample (*26, 27*). Furthermore, we have developed a statistical model to measure dark matter drag from many noisy observations, within which the relative trajectories of galaxies, gas, and dark matter can be combined in a way that eliminates dependence upon 3D orientation and the time since the collision (*28*).

We have studied all galaxy clusters for which optical imaging exists in the Hubble Space Telescope (Advanced Camera for Surveys) data archive (*29*) and X-ray imaging exists in the Chandra Observatory data archive (*30*). We select only those clusters containing more than one component of spatially extended X-ray emission. Our search yields 30 systems, mostly between redshift $0.2 < z < 0.6$ plus two at $z > 0.8$, containing 72 pieces of substructure in total (Table S1). In every piece of substructure, we measure the distance from the galaxies to

the gas δ_{SG}. Assuming this lag defines the direction of motion, we then measure the parallel δ_{SI} and perpendicular δ_{DI} distance from the galaxies to the lensing mass (Fig. 2).

We first test the null hypothesis that there is no dark matter in our sample of clusters (a similar experiment was first carried out on the Bullet Cluster, finding a 3.4 and 8σ detection (*31*)). Observations that do not presuppose the existence of dark matter (*32*) show that $10^{14}M_{\odot}$ clusters contain only 3.2% of their mass in the form of stars. We compensate for this mass, which pulls the lensing signal towards the stars and raised δ_{GI} by an amount typically $0.78 \pm 0.30\,\mathrm{kpc}$ (computed using the known distances to the stars δ_{SG}; see Materials and Methods). The null hypothesis is that the remaining mass must be in the gas. However, we observe a spatial offset between that is far from the expected overlap, even in the presence of combined noise from our gravitational lensing and X-ray observations (Fig. 3A). A Kolmogorov-Smirnov test indicates that the observed offsets between gas and mass are inconsistent with the null hypothesis at 7.6σ, a p-value of 3×10^{-14} (without compensation for the mass of stars, this is 7.7σ). This test thus provides direct evidence for a dominant component of matter in the clusters that is not accounted for by the luminous components.

Having reaffirmed the existence of dark matter, we attempt to measure any additional drag force acting upon it, caused by long-range self-interactions. We measure the spatial offset of dark matter behind the stars, compensating as before for the 16% of mass in the gas (*33*) by subtracting a small amount from δ_{SI} (on average $4.3 \pm 1.6\,\mathrm{kpc}$). We measure a mean dark matter lag of $\langle\delta_{\mathrm{SI}}\rangle = -5.8 \pm 8.2\,\mathrm{kpc}$ in the direction of motion (Fig. 3B), and $\langle\delta_{\mathrm{DI}}\rangle = 1.8 \pm 7.0\,\mathrm{kpc}$ perpendicularly. The latter is useful as a control test: symmetry demands that it must be consistent with zero in the absence of systematics. We also use its scatter as one estimate of observational error in the other offsets.

We interpret the lag through a model (*28*) of dark matter's optical depth (similarly to previous studies (*19, 23*)). Gravitational forces act to keep gas, dark matter and galaxies aligned,

4

while any extra drag force on dark matter induce a fractional lag

$$\beta \equiv \frac{\delta_{\mathrm{SI}}}{\delta_{\mathrm{SG}}} = B\left(1 - \mathrm{e}^{\left(\frac{-(\sigma_{\mathrm{DM}} - \sigma_{\mathrm{gal}})}{\sigma^\star/m}\right)}\right), \tag{1}$$

where σ_{gal} is the interaction cross-section of the galaxies, coefficient B encodes the relative behavior of dark matter and gas, and $\sigma^\star/m$ is the characteristic cross-section at which a halo of given geometry becomes optically thick. We assume that stars do not interact, so $\sigma_{\mathrm{gal}} \approx 0$. To ensure conservative limits on σ_{DM}/m, we also assume $B \approx 1$ and marginalize over $\sigma^\star/m \approx 6.5 \pm 3\,\mathrm{cm}^2/\mathrm{g}$, propagating this broad uncertainty to our final constraints (see Materials and Methods). Adopting the dimensionless ratio β brings two advantages. First, it removes dependence on the angle of the collision with respect to the line of sight. Second, it represents a physical quantity that is expected to be the same for every merger configuration, so measurements from the different systems can be simply averaged (with appropriate noise weighting, although in practice, the constraining power from weak lensing-only measurements comes roughly equally from all the systems).

Combining measurements of all the colliding systems, we measure a fractional lag of dark matter relative to gas $\langle\beta\rangle = -0.04\pm0.07$ (68% CL). Interpreting this through our model implies that dark matter's momentum transfer cross-section is $\sigma_{\mathrm{DM}}/m = -0.25^{+0.42}_{-0.43}\,\mathrm{cm}^2/\mathrm{g}$ (68% CL, two-tailed), or $\sigma_{\mathrm{DM}}/m < 0.47\,\mathrm{cm}^2/\mathrm{g}$ (95%CL, one-tailed); the full PDF is shown in Fig. 4. This result rules out parts of model space of hidden sector dark matter models e.g. (*12,13,15,16*) that predict $\sigma_{\mathrm{DM}}/m \approx 0.6\,\mathrm{cm}^2/\mathrm{g}$ on cluster scales through a long-range force. The control test found $\langle\beta_\perp\rangle \equiv \langle\delta_{\mathrm{DI}}/\delta_{\mathrm{SG}}\rangle = -0.06 \pm 0.07$ (68% CL), consistent with zero as expected. This inherently statistical technique can be readily expanded to incorporate much larger samples from future all-sky surveys. Equivalent measurements of mass loss during collisions could also test dark sector models with isotropic scattering. Combining observations, these *astrophysically large particle colliders* have potential to measure dark matter's full differential scattering cross-
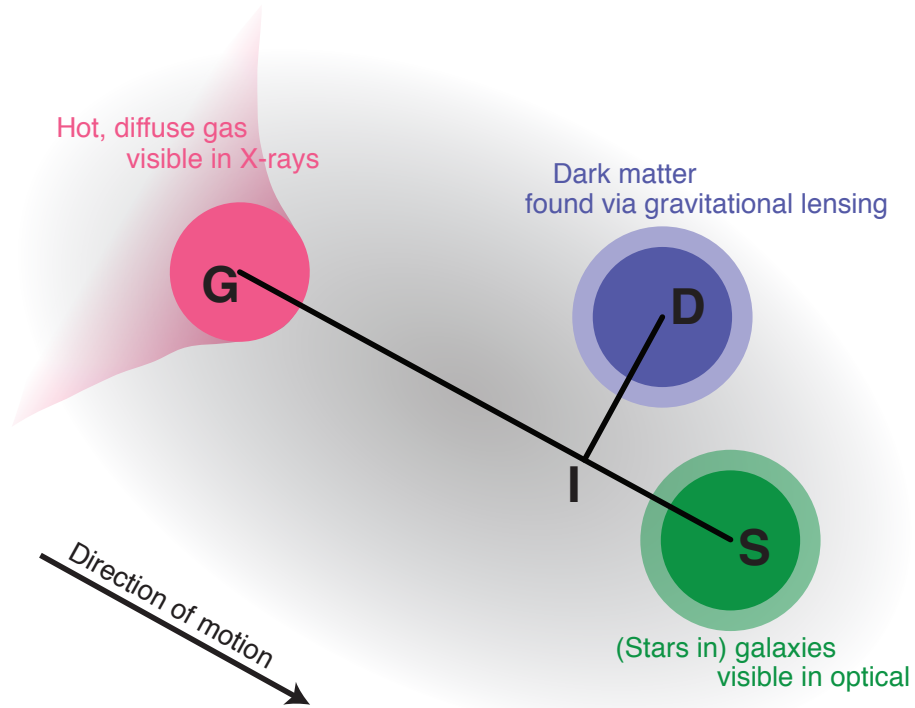
section.

Figure 1: Cartoon showing the three components in each piece of substructure, and their relative offsets, illustrated by black lines. The three components remain within a common gravitational potential, but their centroids become offset due to the different forces acting on them, plus measurement noise. We assume the direction of motion to be defined by the vector from the diffuse, mainly hydrogen gas (which is stripped by ram pressure) to the galaxies (for which interaction is a rare event). We then measure the lag from the galaxies to the gas δ_{SG}, and to the dark matter in a parallel δ_{SI} and perpendicular δ_{DI} direction.
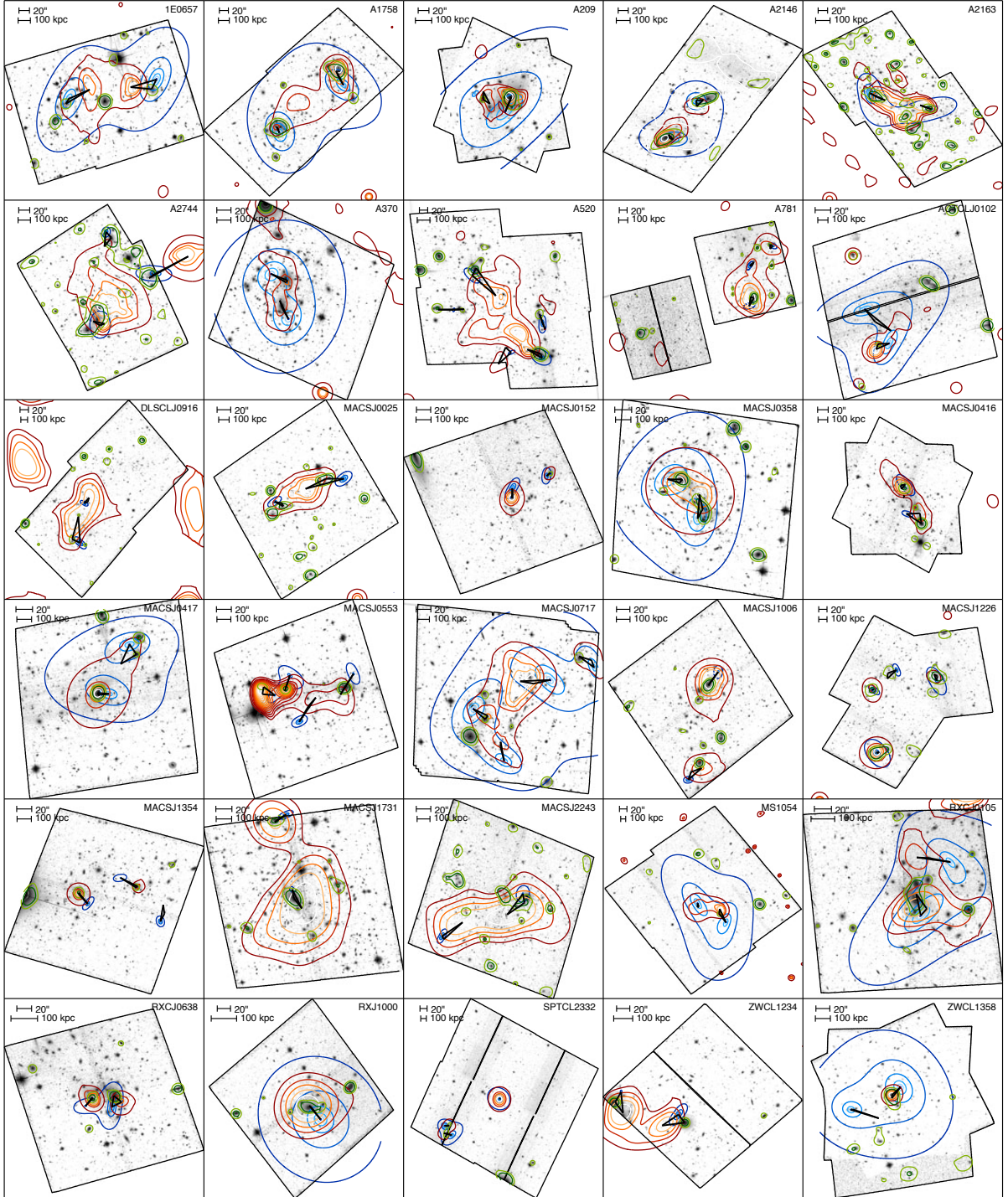
Figure 2: Observed configurations of the three components in the 30 systems studied. The background shows the HST image, with contours showing the distribution of galaxies (green), gas (red) and total mass, which is dominated by dark matter (blue).
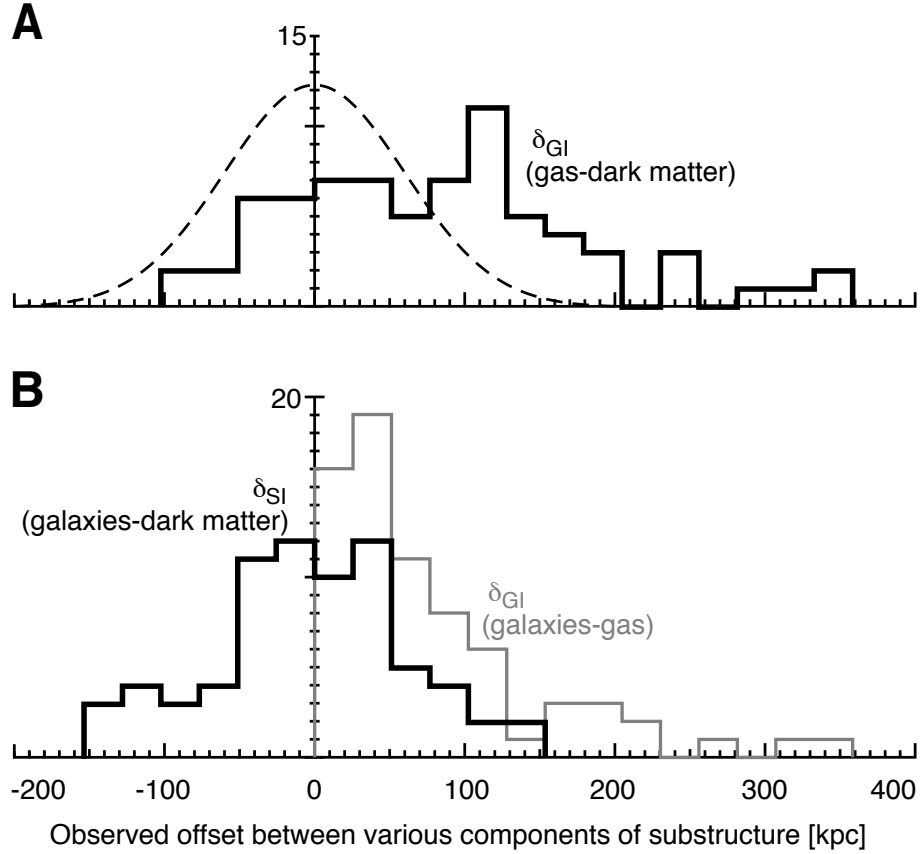
8

Figure 3: Observed offsets between the three components of 72 pieces of substructure. Offsets δ_{SI} and δ_{GI} include corrections accounting for the fact that gravitational lensing measures the total mass, not just that of dark matter. (A) The observed offset between gas and mass, in the direction of motion. The smooth curve shows the distribution expected if dark matter does not exist; this hypothesis is inconsistent with the data at 7.6σ statistical significance. (B) Observed offsets from galaxies to other components. The fractional offset of dark matter towards the gas, $\delta_{\mathrm{SI}}/\delta_{\mathrm{SG}}$, is used to measure the drag force acting on the dark matter.
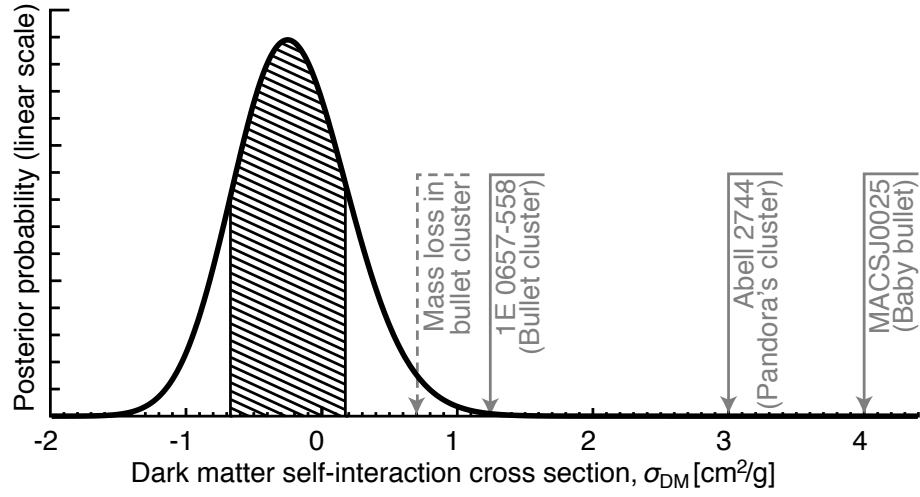
9

Figure 4: Constraints on the self-interaction cross-section of dark matter. These are derived from the separations $\beta = \delta_{\mathrm{SI}}/\delta_{\mathrm{SG}}$, assuming a dynamical model to compare the forces acting on dark matter and standard model particles (*28*). The hatched region denotes 68% confidence limits, to be compared to the 68% confidence upper limits from previous studies of the most constraining individual clusters in blue. Note that the tightest previous constraint is derived from a measurement of dark matter mass loss, which is sensitive to short range self-interaction forces; all other constraints are measurements of a drag force acting on dark matter, caused by long range self-interactions.

Acknowledgements

Supplementary materials for

The non-gravitational interactions of dark matter in colliding galaxy clusters

David Harvey[1,2], Richard Massey[3], Thomas Kitching[4], Andy Taylor[2] & Eric Tittley[2]

[1]Laboratoire d'astrophysique, EPFL, Observatoire de Sauverny, 1290 Versoix, Switzerland
[2]Royal Observatory, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK
[3]Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK
[4]Mullard Space Science Laboratory, University College London, Dorking, Surrey RH5 6NT, UK

Correspondence to: david.harvey@epfl.ch

This PDF file includes:

- Materials and Methods

- SupplementaryText

- Figs. S1 to S8

- References 32–47

Materials and methods

We followed overall procedures that we developed in blind tests on mock data (*24*), usually exploiting algorithms for high precision measurement that had been developed, calibrated and verified elsewhere. However, several custom adaptations were required to analyze the heterogeneous data from the Hubble Space Telescope (HST) and Chandra X-ray Observatory archives (Table S1 lists all the observed systems, and Figure S2 shows the offsets measured in each).

Here we describe the methods we used to combine observations with different exposure times, filters, epochs and orientations – starting from the raw data and performing a full reduction to maximise data quality. To convert angular distances into physical distances, we assume a cosmological model derived from measurements of the Cosmic Microwave Background (*33*), $\Omega_{\mathrm{M}} = 0.31$, $\Omega_{\Lambda} = 0.69$, $H_0 = 67 \, \mathrm{km/s/Mpc}$.

Position of gas, seen in X-ray emission

We downloaded the raw event 1 files for all observations. To process these data, we used CIAO tools version 4.5, starting with basic reduction and calibration using the CIAO *repro* tool. In our analysis, it is particularly important to remove emission from point sources, and prevent X-ray bright Active Galactic Nuclei at the centers of clusters from biasing our position measurements. We therefore made a first pass at removing point sources using *celldetect*. We then filtered each event table for any potential spurious events such as solar flares by clipping the table at the 4σ below the mean flux level.

Having cleaned each exposure, we combined them using the *merge_obs* script from CIAO tools into a single exposure map corrected flux image, producing along with it exposure maps for each observation and the stacked image. We modeled the Chandra PSF at each position throughout the field, we created individual maps using *mkpsfmap* for each exposure at an effective energy of 1 keV, then combined each model weighting them by their respective exposure map. Figure S3 shows an example of the PSF map used for cluster A520.

To make a second pass to identify point sources, we passed the stacked image and PSF model through CIAO *wavdetect*, a wavelet smoothing algorithm that employs a 'Mexican hat' filter on a range of scales. This estimates the true size of each source, correcting for the size of the PSF. We used the smallest scales for the wavelet radii (1, 2 pixels) to identify point sources, and combined the larger scales (4, 8, 16, and 32 pixels) into a denoised version of the final image. We finally inspected every image by eye for any remaining point sources. We found that this double filter method proved very successful at removing point sources, with only AGN at the edge of the cluster remaining unflagged. Although their emission has extended wings, the cluster is usually in the center of the pointing, resulting in minimal contamination.

Finally, we measured the position of coherent substructure in the X-ray emission using *SExtractor* (*34*). This calculates positions from the first order moments of the light profile, which means that the returned position does not always coincide exactly with the brightest pixel. *SExtractor* does not report reliable errors in the positions but, since the dominant contribution of variation is the size of the smoothing kernel, we can estimate the robustness of our measurements by smoothing the image using different scales in *wavdetect*, and measure the rms across different scales. On average we found the rms error to be 4 arcseconds (roughly 30 kpc at redshift z=0.4).

Cluster	RA (deg)	DEC (deg)	z	ACS Filter	ACS (s)	Chandra (ks)
1E0657	104.612	-55.9477	0.296	F814W_F775W	15094.0	597.39
A1758	203.194	50.5426	0.2792	F814W	10000.0	216.00
A209	22.9728	-13.6127	0.206	F814W	4040.00	24.09
A2146	239.007	66.3725	0.234	F814W	9233.00	84.08
A2163	243.938	-6.14690	0.203	F814W	9192.00	444.59
A2744	3.58210	-30.3898	0.308	F814W	11980.0	133.12
A370	39.9627	-1.58000	0.373	F814W	3840.00	86.81
A520	73.5395	2.93110	0.202	F814W	18320.0	426.01
A781	140.149	30.4927	0.298	F814W	1620.00	49.54
ACTCLJ0102	15.7277	-49.2560	0.87	F814W	1916.00	359.16
DLSCLJ0916	139.046	29.8450	0.5343	F814W	9894.00	41.28
MACSJ0025	6.37460	-12.3818	0.5843	F814W	4200.00	168.61
MACSJ0152	28.1473	-28.8944	0.341	F606W	1200.00	20.04
MACSJ0358	59.7174	-29.9320	0.428	F814W	4620.00	65.74
MACSJ0416	64.0392	-24.0735	0.42	F814W	4037.00	57.50
MACSJ0417	64.3926	-11.9111	0.443	F814W	1910.00	95.92
MACSJ0553	88.3494	-33.7117	0.407	F814W	4572.00	88.74
MACSJ0717	109.389	37.7528	0.5458	F814W	8893.00	83.22
MACSJ1006	151.730	32.0198	0.359	F814W	1440.00	13.30
MACSJ1226	186.694	21.8673	0.37	F814W	5520.00	153.81
MACSJ1354	208.635	77.2528	0.3967	F814W	1200.00	35.46
MACSJ1731	262.913	22.8660	0.389	F814W	1440.00	22.28
MACSJ2243	340.837	-9.58910	0.447	F606W	1200.00	21.88
MS1054	164.245	-3.62000	0.826	F606W	8100.00	89.51
RXCJ0105	16.4096	-24.6801	0.23	F606W	1200.00	21.97
RXCJ0638	99.6953	-53.9735	0.1658	F606W	1200.00	21.78
RXJ1000	150.132	44.1491	0.154	F606W	1200.00	20.66
SPTCL2332	352.959	-50.8642	0.5707	F606W	7680.00	39.9
ZWCL1234	189.045	28.9929	0.2214	F814W	27632.0	51.75
ZWCL1358	209.951	62.5163	0.329	F850LP	13692.0	63.10

1

Figure S1: The full sample of 30 merging complexes, and their locations on the sky. The columns show, from left to right: the name of the cluster, its right ascension, declination, and redshift, the HST/ACS filter used for our lensing analysis, and the total exposure time for that particular filter, and the (cleaned) exposure time of the Chandra X-ray image.
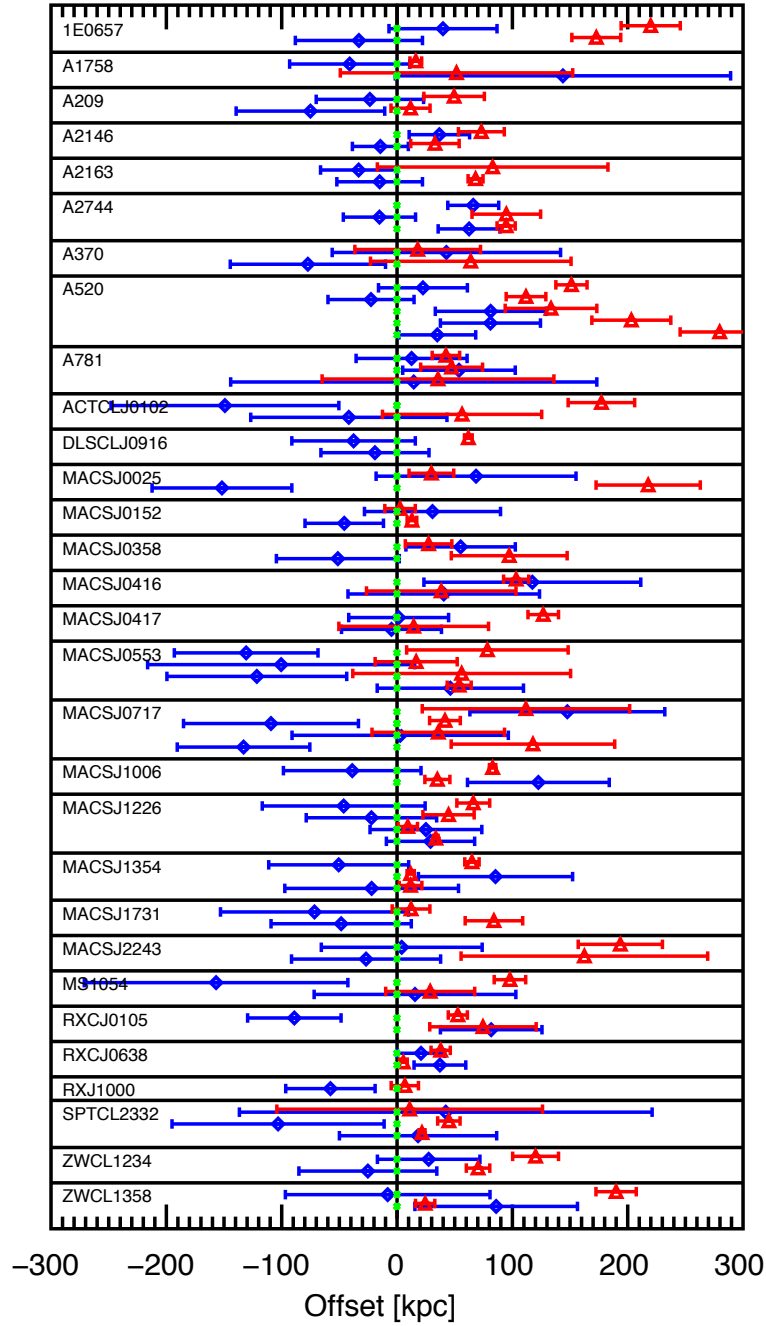
Figure S2: Observed offsets between galaxies, gas and dark matter in 72 components of sub-structure. In each case, the green triangle, at the centre of the coordinate system, denotes the position of the galaxies. The separation between galaxies and gas, δ_{SG}, is shown in red. The separation of the dark matter with respect to the galaxies, projected onto the SG vector, δ_{SI}, is shown in blue. The error bars show the locally estimated 1σ errors.

4

Figure S3: An example model of the size of the Chandra X-ray telescope's Point Spread Function (PSF). The model PSF is used to identify and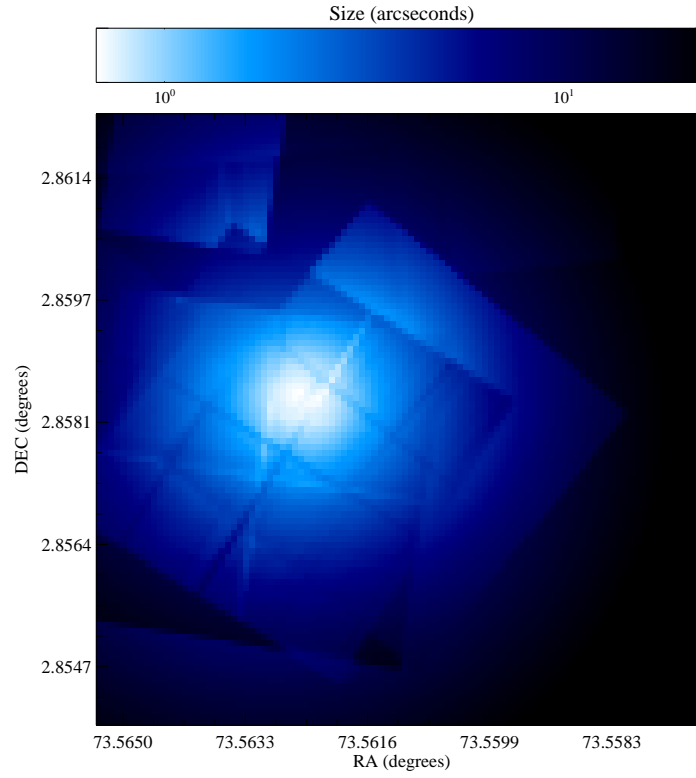 remove point sources, e.g. Active Galactic Nuclei – and to thereby identify extended X-ray emission from hot gas within the cluster. The image shows a combined, exposure map weighted, PSF map stacked for the various observations of galaxy cluster A520.

Position of galaxies, seen in optical emission

We searched the HST archive for data acquired with the Advanced Camera for Surveys (ACS) instrument, which has the largest field of view. We considered only filters F606W, F814W and F850LP, whose high throughput ensures deep imagining, and whose red wavelengths ensure (a) that the optical emission samples the old stars that dominate the mass content of these systems and (b) a high density of high redshift galaxies visible behind the cluster, to provide sufficient lensing signal. Some clusters had been observed in more than one wavelength band. We used only a single band for all the clusters to further homogenize the data, but have compared a subset of our results in different bands to check for systematic errors. For our main analysis, we selected the broad F814W band, unless there are significantly more exposures in another.

We corrected the raw, pixellated data for charge transfer inefficiency (*35*), then performed basic data reduction and calibration using the standard *Calacs* pipeline. We used *tweakReg* to orient and align individual exposures, then stacked them using *MultiDrizzle* (*36*) with a Gaussian convolution kernel and PIXFRAC=0.8 (*37*) to produce a deep, mosaicked image with a pixel scale of 0.03 arcseconds. In the process, *MultiDrizzle* also output a reoriented image of each individual exposure, which we used for star/galaxy identification and PSF estimation.

We estimated the distribution of mass in galaxies via the proxy of the light emitted by their stars. In our single-band imaging, we were able to identify and mask foreground stars in the Milky Way (which appear pointlike), but assumed any foreground or background galaxies to be randomly positioned and thus merely add shot noise to our measurements. We smoothed the masked image using *wavdetect*, and measured the position of coherent substructure using *SExtractor* (*34*). This calculates positions from the first order moments of the light profile, which means that the returned position does not always coincide exactly with the brightest pixel. *SExtractor* does not report reliable errors in the positions. However, since the dominant contribution of noise is inclusion or omission of galaxies inside the smoothing kernel, we estimated the robustness of our measurements by smoothing the image using different scales in *wavdetect*, and compared the resulting positions. On average, we found an rms error in the position of the extracted halos of 0.6 arcseconds (roughly 4.5 kpc at redshift z=0.4).

We also tried two other ways to quantify the position of the galaxies. First, we measured the smoothed distribution of galaxies in the image, with all galaxies weighted equally (this represents the opposite – and least realistic – assumption of galaxies' mass/light ratio). To do this in practice, we passed the galaxy catalogue through the X-ray data reduction pipeline, as if each galaxy were a single X-ray photon. This created a smoothed image, in which we identified substructure using *SExtractor*. Since the same galaxies contributed both to the flux-weighted and galaxy-weighted positions, the two measurements are correlated. We measure the uncertainty on the galaxy weighted positions to be 5 kpc, about the same as the flux-weighted positions. We obtain consistent values of $\langle \beta \rangle = 0.054 \pm 0.062$ (68% CL) and conclude that $\sigma_{\mathrm{DM}}/m = 0.36^{+0.46}_{-0.45}\,\mathrm{cm}^2/\mathrm{g}$ (68% CL, two-tailed). Second, we tried identifying the position of the 'Brightest Group Galaxy' (BGG), since its formal error is small, and it has proved optimal in studies of isolated groups (*38*). In merging systems however, the brightest nearby galaxy is frequently unassociated with the infalling group (*39*). Accounting for our observed 1.7 ± 0.9 arcsecond offset to any brighter galaxy within 25 arcseconds of X-ray emission (the search region that will be used to identify gravitational lensing signals), again yields a consistent constraint on σ_{DM}/m, but with much larger final error.

6

Position of dark matter, measured via weak gravitational lensing

We measured the ellipticities of galaxies in HST images using the *RRG* method (*40*). This corrects galaxies' Gaussian-weighted moments for convolution with the Point Spread Function (PSF), to measure the shear γ_1 (γ_2) corresponding to elongations along (at 45 degrees to) the x axis. This method has been empirically calibrated on simulated HST imaging in which the true shear is known (*41*), applying a multiplicative correction of $\langle m \rangle = -3.0 \times 10^{-3}$ and a additive bias of $\langle c \rangle = -2.1 \times 10^{-4}$.

HST's PSF varies across the field of view and, because thermal variations change the telescope's focus, at different epochs. Modelling the net PSF in our stacked images therefore required a flexible procedure. We first identified stars in the deep, stacked image using their locus in size–magnitude space. We then measured the ellipticity of each star in individual exposures. By comparing these to *TinyTim* (*42*) models of the HST PSF (created by raytracing through the telescope at different focus positions but at the appropriate wavelengths for the band), we determined the focus position for each exposure. We then interpolated (second and fourth shape moments of) the *TinyTim* PSF model to the position of the galaxies, rotating into the reference frame of the *MultiDrizzle* mosaic. We then summed the PSF moments from each exposure in which a galaxy was observed. Figure S4 shows an example of the final PSF model for one cluster.

We measured the shear of all galaxies that appear in 3 or more exposures, with a combined signal-to-noise in the stacked image > 4.4 and size > 0.1 arcseconds. These cuts (*41*) remove noisy measurements at the edges of the field or in the gaps between detectors. We also masked out galaxies that lie near bright stars or large galaxies, whose shapes appear biased. Figure S5 compares shear catalogues for a single cluster, derived from independent analyses of data in the F814W and F606W bands. There is the expected level of scatter between the two measurements – but, most importantly, there is no detectable bias.

We reconstructed the distribution of mass in the clusters using the parametric model-fitting algorithm *Lenstool* (*43*). Using Bayesian likelihood minimization, *Lenstool* simultaneously fits multiple mass haloes to an observed shear field, with the position and shape of each halo described by the NFW (*44*) density profile. This is an efficient technique to record a unique position for each halo, marginalizing over nuisance parameters that include mass and morphology, that are not of direct interest to our study. Assuming this density profile does not bias measurements of the position of halos within current statistical limits (*24*). *Lenstool* requires positional priors to be defined in which it searches for the lensing signal. Except in a few well-studied systems (where we use the extra information), we obtained an initial lensing model using one prior search radius centered on each gas position and large enough to incorporate any nearby groups of galaxies. Following this scheme, we used an automated procedure to identify and associate the mutually closest galaxy, gas and lensing signals into systems of three mass components. In all systems, we then modeled the lensing system a final time, adopting priors centred on the galaxy position (we redid this step when trying different position estimators for the galaxies). Henceforth, we could center the coordinate system for each combined system of galaxies, gas and dark matter on the galaxies, to avoid prior bias in the Bayesian fits.

Lenstool samples the posterior surface in two ways. To obtain the best fitting position, we iterated to the best-fit solution with a converging MCMC step size, using ten simultaneous sampling chains to avoid local maxima. To sample the entire posterior surface (whose width quantifies uncertainty on model parameters), we then reran the algorithm with a fixed step size. The 1σ error on position was on average 11.4 arcseconds (roughly 60 kpc at redshift z=0.4). As a sanity check we compare our measured centroids to those systems included in previous studies. Our statistical uncertainty is sometimes larger because we
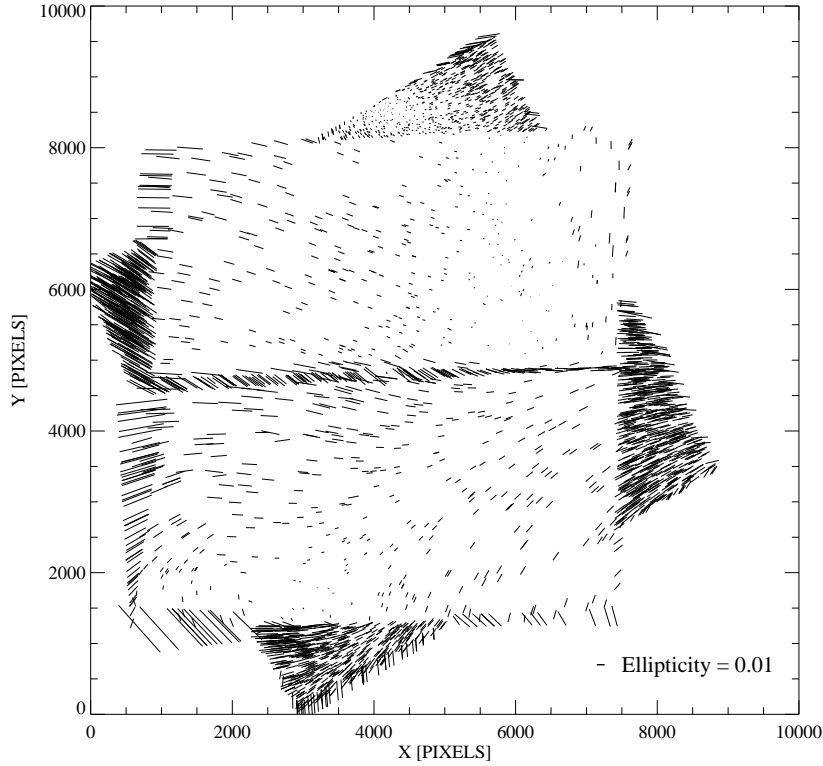
7

Figure S4: An example model of the Point Spread Function (PSF) of the Hubble Space Telescope/Advanced Camera for Surveys (HST/ACS). Each tick mark represents the ellipticity of the PSF at that particular position in the HST field. Its orientation shows the PSF's major axis and its length shows the ellipticity; a dot would indicate a circular PSF. The PSF tends to be highly elliptical near the edge of the field and more circular in the centre. Tick marks are plotted at the position of every "detected" source. The mosaic pattern of dithered exposures can be seen: noisier regions with fewer exposures contain more spurious sources, which are removed during analysis (but are shown here for clarity). The example shown is for observations of galaxy cluster MACSJ0416.
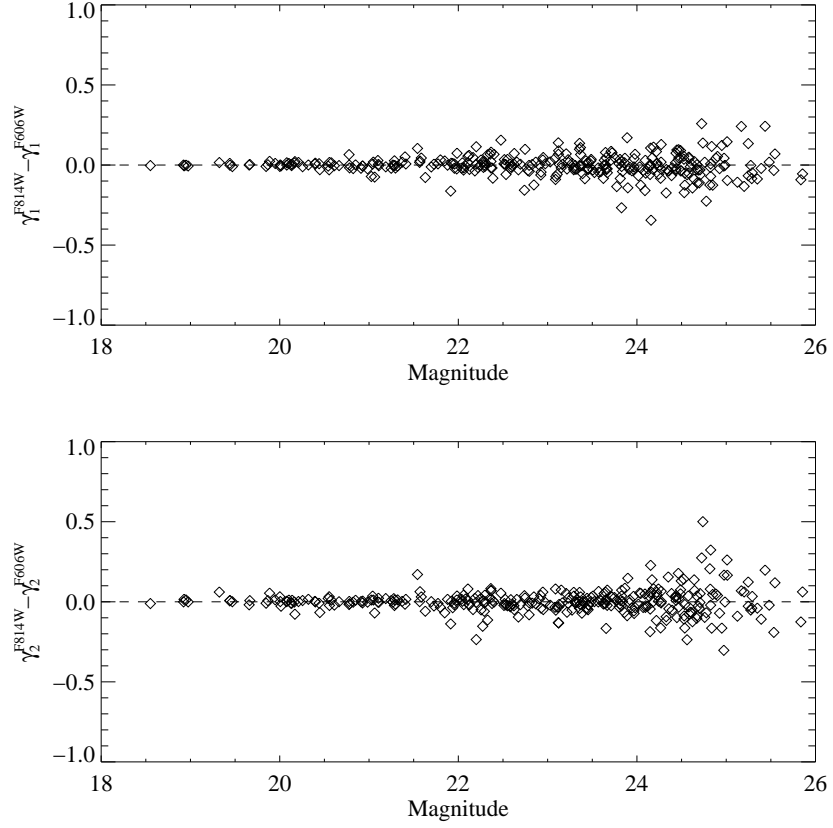
8

Figure S5: A comparison of the gravitational lensing shears measured independently behind a single cluster, in two different HST filters. The top (bottom) panel shows the difference between γ_1 (γ_2) for each galaxy, which traces to elongations along (at 45 degrees to) the x axis. We find scatter as expected due to observational noise, but no systematic bias.

use only weak gravitational lensing, but we find no evidence for any bias. For example, our measured positions in the 'bullet cluster' lie within one standard deviation of those reported in (*31*).

Positional offsets between components

When assigning different mass components to one another, for almost all the clusters, we used an automated matching algorithm to associate the nearest clumps of dark matter, gas and stars. This was made robust by performing the matching in both directions (e.g., dark matter to stars, and stars to dark matter). In a few cases where detailed analyses of individual systems were available in the literature (for example, using strong lensing, X-ray shocks, optical spectroscopy or imaging additional bands, which were outside the scope of our work), we inserted that prior information by hand during association. This was most useful in systems A520 and A2744. As a further test, we carry out a jackknife test to ensure that the association does not effect the overall constraints, and moreover, no single cluster dominates the result. We find no evidence for such an effect, and derive consistent error bars of $\Delta\sigma_{\mathrm{DM,JK}}/m = \pm 0.5 \mathrm{cm}^2/\mathrm{g}$, further supporting the error bars quoted in our final result.

We drew an offset vector δ_{SG} in angle between the observed position of the gas and galaxies, which we took to define the system's direction of motion. We then measured the position of the total mass along that vector and (in a right handed coordinate system) perpendicular to it, defining offset vectors δ_{SI}, δ_{GI}, and δ_{DI} from the intersection point I of these vectors.

Gravitational lensing measures the position of total mass, rather than that of just dark matter. We corrected the measured offsets δ_{SI} and δ_{GI} for the contribution from the next most massive component. To calibrate this correction, we analysed mock lensing data from a dominant mass component (with an NFW (*44*) profile) plus a less massive component at some offset δ. The corrections were always small but, for a subdominant component with the same profile, normalised to contain a fraction f of the total mass, we found that the lensing position is pulled by an amount $f\delta e^{-0.01\delta/r_s}$, and we corrected for that. If we do not calibrate for the extra pull of gas on the lensing peak we infer an upper limit of $\sigma_{\mathrm{DM}}/m < 0.54 \mathrm{cm}^2/\mathrm{g}$ (68% CL, one-tailed).

To test the hypothesis that dark matter does not exist, we required a model of the δ_{GI} data expected if this were true. To generate that model, we assumed that the true positions of the X-ray and lensing signals coincided, but that the observed positions were offset by a random amount determined by the appropriate level of noise in each (see above). We calculated the 2D offset, then projected this onto the direction to the stars, which is also selected at random. We could have slightly increased the model δ_{GI} offset to account for the mass in stars (the increase must be positive because the vector δ_{SG} is defined from the galaxies to the gas). However, it is better to instead decrease the observed δ_{GI} offset. The two approaches are equivalent in principle, but the latter allowed information to be added to our analysis because the absolute value of δ_{SG} was known in each system. When comparing the model and observed δ_{GI} offsets via a Kolmogorov-Smirnov test (in which we computed critical values using a Taylor series), we also used the errors on σ_{GI} determined for each system individually.

When measuring the interaction cross-section of dark matter, we converted offset measurements in arcseconds to physical units of kpc (using a standard cosmological model, which assumes dark matter exists). This enabled a more detailed comparison of the offsets between different systems. The (noisily determined) error estimates of offsets in a few systems were anomalously low, and likely smaller than the uncertainty in our knowledge of the merging configuration. To more robustly quantify the total uncertainty of offsets (which should include observational noise plus the possibility of component

misidentification and merging irregularities), we empirically exploited the control test δ_{DI}, which has an rms variation between systems $\sigma_{DI} = 60\,\text{kpc}$. This value is consistent with most of the individually measured errors, but more robust. We therefore adopted it globally as the error on every measurement of δ_{DI} and δ_{SI}, rescaling to a value in arcseconds at the redshift of each system. Errors in δ_{SG} must be smaller than this, because they do not involve observational noise in the lensing position. However, they also include the possibility of component misidentification, which is best estimated through this global approach. We therefore adopted the conservative approach of also assigning this value as the error on every measurement of δ_{SG}. Thus we set $\sigma_{SG} = \sigma_{SI} = \sigma_{DI} = 60\,\text{kpc}$. To combine our measurements of $\beta = \delta_{SI}/\delta_{SG}$ and $\beta_{\perp} = \delta_{DI}/\delta_{SG}$ from individual systems, we multiplied their posterior probabilities (approximated as a normal distribution even though it is a Cauchy distribution, but with a width determined by propagating errors on the individual offsets).

Interpreting positional offsets as an interaction cross-section

Similarly to previous studies of the cross-section of dark matter (*19, 23*), we interpreted observations of offset dark matter in terms of its optical depth for interactions. However, we have developed a more sophisticated model (*28*) intended to take into account the 3D and time-varying trajectories of infalling halos. First, calculating the dimensionless ratio $\beta = \delta_{SI}/\delta_{SG}$ removes dependence on the angle of the collision with respect to the line of sight. Furthermore, a set of analytic assumptions suggests that β is a physically meaningful quantity that should be the same for every system. The main assumption of quasi-steady state equilibrium is reasonable for the detectable systems in our sample, but caution would be needed to interpret dark matter substructure that had passed directly through the cluster core (and had its gas stripped) or substructure on a radial orbit caught at the brief moment of turnaround (this is a negligible fraction in our mock data). The model also incorporates the results of simulations (*11*) in which dark sector interactions that are frequent but exchange little momentum (e.g. via a light mediator particle that produces a long-ranged force and anisotropic scattering) produce a drag force and separate dark matter from the stars. On the other hand, simulations of 'billiard ball' interactions that are rare but exchange a lot of momentum (e.g. via a massive mediator that produces a short-ranged force and isotropic scattering) tend to scatter dark matter away from a system and produce mass loss (*11, 18, 19*). However, we note that the ref. (*18*) also reports an unexpected small separation between galaxies and dark matter after billiard ball scattering. In this paper, we explicitly follow the prescription in (*11*).

According to our model of dark matter dynamics (see equation 33 of ref. (*28*)), the offset of dark matter from galaxies, calibrated against the offset of gas, is

$$\langle \beta \rangle = B \left(1 - \exp\left(\frac{-(\sigma_{DM} - \sigma_{gal})}{\sigma^{\star}} \right) \right). \tag{S1}$$

Since the gaps between galaxies are vast compared to their size, they interact very rarely, so we assumed that $\sigma_{gal} \approx 0$. If this assumption were wrong, or in the presence of observational noise, our analysis can therefore produce negative values of σ_{DM}/m. Our quoted errors include observational errors, pair-assignment errors, and model parameter errors.

The value of $\sigma^{\star}/m$ depends upon the geometrical properties of the dark matter halo, but is proportional to its mass and inversely proportional to its cross-sectional area. For our set of merging systems, we conservatively adopted a $\sigma^{\star}/m = 6.5 \pm 3\,\text{cm}^2/\text{g}$ by assuming the system masses are

$\log(M_{200}/M_\odot) = 14 \pm 1$, with NFW density profiles and concentration varying with mass as observed in numerical simulations (*45*). By assuming a conservative range in halo masses we propagate a much larger error in $\sigma^\star/m$ than one would expect if we were to measure the true values. We then analytically marginalized over $\sigma^\star/m$, propagating the uncertainty through to our final constraints. The top panel of Figure S6 shows the values of $\sigma^\star/m$ instead assuming different, fixed system masses; the bottom panel shows the effect on $\sigma_{\rm DM}/m$. The inferred estimate of $\sigma_{\rm DM}/m$ is broadly insensitive to $\sigma^\star/m$, varying from $\sigma_{\rm DM}/m = -0.23 \pm 0.60\,{\rm cm}^2/{\rm g}$ for an assumed halo of $M_{200} = 10^{13}M_\odot$ to $\sigma_{\rm DM}/m = -0.1 \pm 0.28\,{\rm cm}^2/{\rm g}$ for a halo of $M_{200} = 10^{15}M_\odot$.

The relative behaviour of gas and dark matter was compared through a ratio in the prefactor

$$B = \frac{C_{\rm DM}A_{\rm DM}M_{\rm gas}\rho_{\rm DM}}{C_{\rm gas}A_{\rm gas}M_{\rm DM}\rho_{\rm gas}}, \tag{S2}$$

where C, A and M are the drag coefficient, size and mass of the merging halo, and ρ is the density of material through which it is moving. We assumed a conservative lower limit of $B \gtrsim 1$, leading to a conservative upper limit on our constraints on $\sigma_{\rm DM}/m$.

The first requirement to have ensured a conservative treatment is that the infalling substructure's gas envelope is smaller than its dark matter envelope, $A_{\rm gas} < A_{\rm DM}$. This is generically true of isolated structures in numerical simulations and, as gas is stripped during the collision, it will become smaller still. The geometric size of a gas halo also depends upon its temperature – and hot gas may be more easily stripped than cold gas. To test whether that has an statistically significant effect, we measured the X-ray temperature of each observed infalling system, and separately analyzed the hotter and cooler halves of our sample. As shown in Figure S7, the results for each half remain consistent, with error bars larger by approximately $\sqrt{2}$. For the hotter sample ($T > 8\,{\rm keV}$), we found $\sigma_{\rm DM}/m = -0.10 \pm 0.58\,{\rm cm}^2/{\rm g}$ and for the cooler sample ($T < 8\,{\rm keV}$) we found $\sigma_{\rm DM}/m = -0.50 \pm 0.64\,{\rm cm}^2/{\rm g}$. Although there was marginal evidence that hot gas is more easily stripped than cold gas, which could be investigated with a much larger sample, our conclusions remain unaffected within current statistical precision.

The second requirement to have ensured a conservative treatment is that the gas fraction in the medium through which the bullet is traveling, $f_{\rm gas} \equiv \rho_{\rm gas}/\rho_{\rm DM}$ is less than that of the infalling structure $M_{\rm gas}/M_{\rm DM}$. We assumed that, overall, infalling structure contains the universal fraction $\Omega_{\rm B}/\Omega_{\rm D} = 0.184$ (*33*), and we measured $f_{\rm gas}$ in mock data realised from cosmological simulations of structure formation (*46*). The mean $f_{\rm gas}$ over all simulations (the solid line in Figure S8) is lower than the universal fraction, and is indeed constant (within 10%) at the radii of observable substructures (points in Figure S8). These conclusions from simulations are consistent with deep X-ray observations of galaxy clusters, e.g. (*47*).
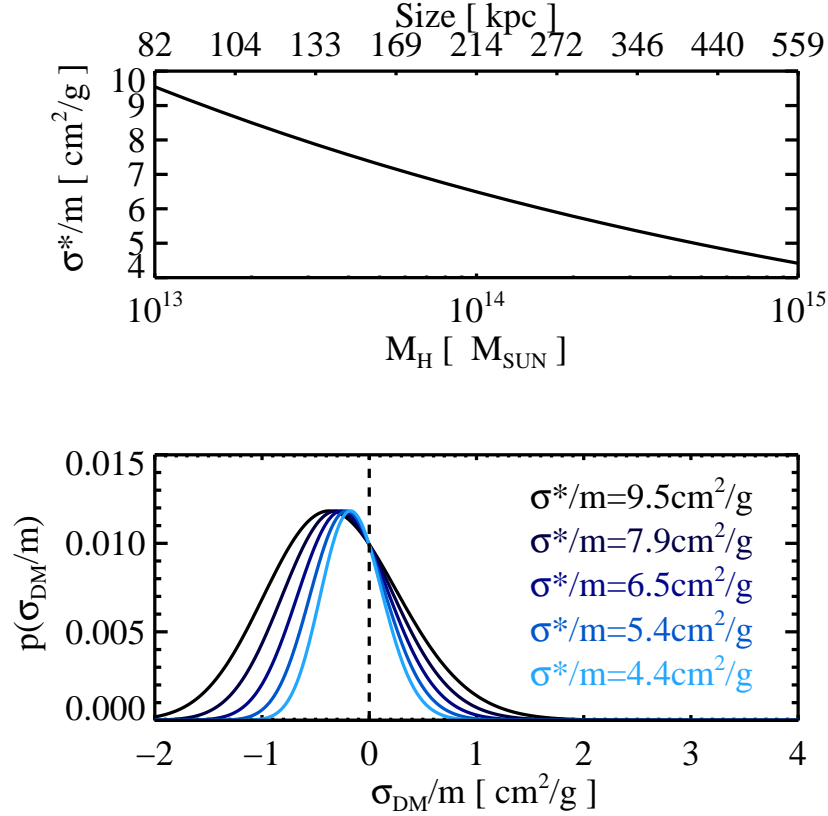
Figure S6: The sensitivity of measurements of dark matter's self-interaction cross-section to the model parameter $\sigma^\star/m$. This parameter is the characteristic value of cross-section at which an appropriately-sized cloud of standard model particles becomes optically thick. The top panel shows the value of $\sigma^\star/m$ for different various substructure masses, assuming an NFW mass profile and a mass-concentration relation from cosmological simulations (45). The bottom panel demonstrates how a few of those values affect our measurement of the cross-section. The resulting variation is sub-dominant to statistical error in our sample of clusters. We adopted a value of $\sigma^\star/m = 6.5 \pm 3\,\mathrm{cm}^2/\mathrm{g}$, corresponding to dark matter halos of $M = 10^{14\pm1}M_\odot$, and propagated the uncertainty through to our final constraints.
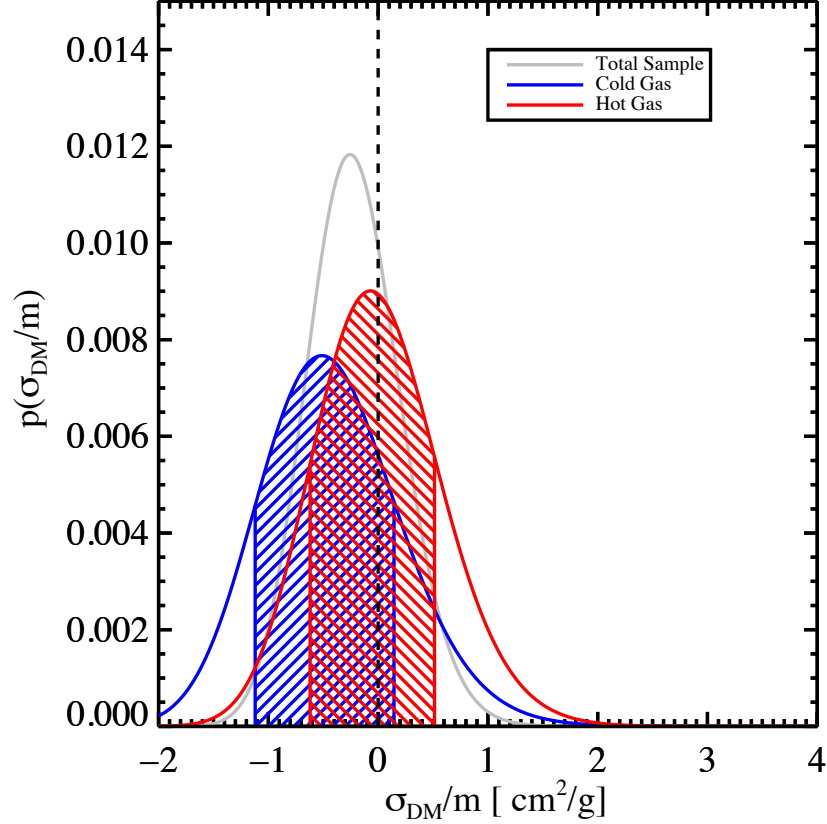
Figure S7: The sensitivity of measurements of dark matter's self-interaction cross-section to the temperature of the gas against which dark matter's trajectory is calibrated. We measured the gas temperature from the X-ray spectra of our 72 systems, and split the sample in two: blue data show substructures with gas temperature $< 8\,\mathrm{keV}$, and red data show substructures with gas temperature $> 8\,\mathrm{keV}$. The constraining power of each sample is approximately $\sqrt{2}$ less than that of the full sample, shown in grey, and no statistically significant difference is measured.
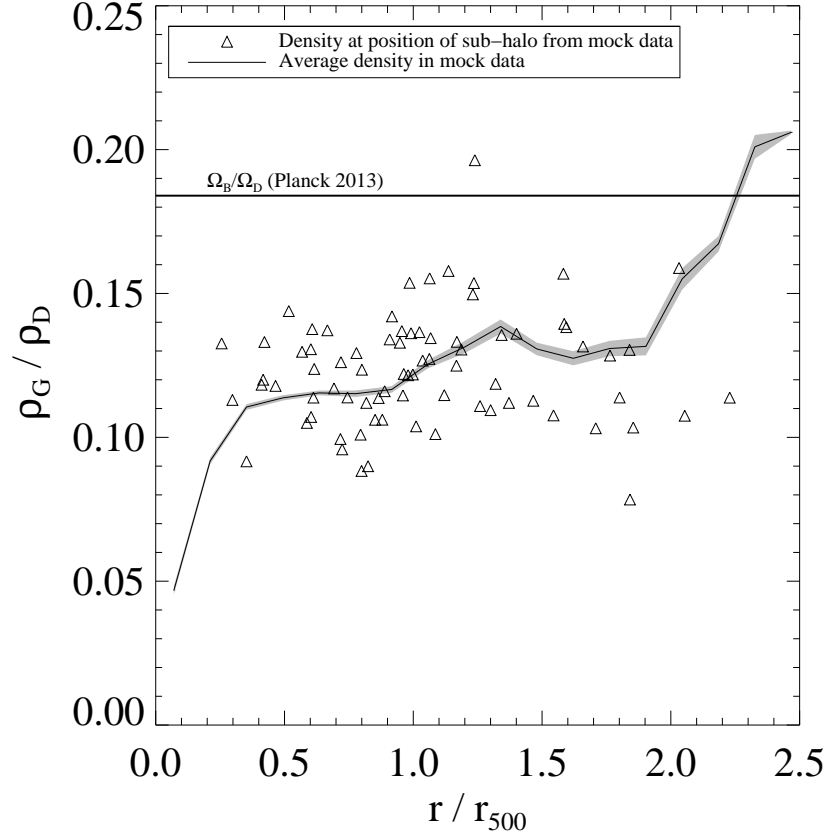
Figure S8: The sensitivity of measurements of dark matter's self-interaction cross-section to the density of gas through which it is moving. The plot shows the gas fraction $f_{\mathrm{gas}} = \rho_{\mathrm{gas}}/\rho_{\mathrm{DM}}$ in simulated galaxy clusters (*46*), as a function of clustercentric radius. The solid line shows the average f_{gas} over 16 clusters, with the 1σ error on the mean given in grey. Triangles show the measured f_{gas} at the radius of substructures observable in mock 2D realisations of the 3D simulations (only the inner $\sim 60\%$ lie inside the HST field of view at the redshifts of the observed systems). Our interpretation of the dark matter and gas trajectories as an interaction cross-section, assumes that these are lower than the universal fraction $\Omega_{\mathrm{B}}/\Omega_{\mathrm{D}} = 0.184$ (*33*).

References and Notes

1. M. Davis, G. Efstathiou, C. S. Frenk, S. D. M. White, The evolution of large-scale structure in a universe dominated by cold dark matter, *ApJ* **292**, 371-394 (1985).

2. J. Dubinski, R. G. Carlberg, The structure of cold dark matter halos, *ApJ* **378**, 496-503 (1991).

3. A. Klypin, A. V. Kravtsov, O. Valenzuela, F. Prada, Where Are the Missing Galactic Satellites?, *ApJ* **522**, 82-92 (1999).

4. B. Moore, *et al.*, Dark Matter Substructure within Galactic Halos, *ApJ* **524**, L19-L22 (1999).

5. M. Boylan-Kolchin, J. S. Bullock, M. Kaplinghat, Too big to fail? The puzzling darkness of massive Milky Way subhaloes, *MNRAS* **415**, L40-L44 (2011).

6. A. Pontzen, F. Governato, Cold dark matter heats up, *Nature* **506**, 171-178 (2014).

7. J. M. Bardeen, J. R. Bond, N. Kaiser, A. S. Szalay, The statistics of peaks of Gaussian random fields, *ApJ* **304**, 15-61 (1986).

8. D. N. Spergel, P. J. Steinhardt, Observational Evidence for Self-Interacting Cold Dark Matter, *Physical Review Letters* **84**, 3760-3763 (2000).

9. M. Rocha, *et al.*, Cosmological simulations with self-interacting dark matter - I. Constant-density cores and substructure, *MNRAS* **430**, 81-104 (2013).

10. J. Zavala, M. Vogelsberger, M. G. Walker, Constraining self-interacting dark matter with the Milky Way's dwarf spheroidals, *MNRAS* **431**, L20-L24 (2013).

11. F. Kahlhoefer, K. Schmidt-Hoberg, M. T. Frandsen, S. Sarkar, Colliding clusters and dark matter self-interactions, *MNRAS* **437**, 2865-2881 (2014).

12. R. Foot, Mirror dark matter: Cosmology, galaxy structure and direct detection, *International Journal of Modern Physics A* **29**, 30013 (2014).

13. K. K. Boddy, J. L. Feng, M. Kaplinghat, T. M. P. Tait, Self-interacting dark matter from a non-Abelian hidden sector, *Phys. Rev. D* **89**, 115017 (2014).

14. Y. Hochberg, E. Kuflik, T. Volansky, J. G. Wacker, The SIMP Miracle, *arXiv:1402.5143* (2014).

15. J. M. Cline, Z. Liu, G. D. Moore, W. Xue, Composite strongly interacting dark matter, *Phys. Rev. D* **90**, 015023 (2014).

16. S. Tulin, H.-B. Yu, K. M. Zurek, Resonant Dark Forces and Small-Scale Structure, *Physical Review Letters* **110**, 111301 (2013).

17. LUX Collaboration, First results from the LUX dark matter experiment at the Sanford Underground Research Facility, *Phys. Rev. Lett.* **112**, 091303 (2013).

18. S. W. Randall, M. Markevitch, D. Clowe, A. H. Gonzalez, M. Bradač, Constraints on the Self-Interaction Cross Section of Dark Matter from Numerical Simulations of the Merging Galaxy Cluster 1E 0657-56, *ApJ* **679**, 1173-1180 (2008).

19. M. Markevitch, *et al.*, Direct Constraints on the Dark Matter Self-Interaction Cross Section from the Merging Galaxy Cluster 1E 0657-56, *ApJ* **606**, 819-824 (2004).

20. D. Eckert, *et al.*, The stripping of a galaxy group diving into the massive cluster A2142, *A&A* **570**, A119 (2014).

21. M. Bartelmann, P. Schneider, Weak gravitational lensing, *Phys. Rep.* **340**, 291-472 (2001).

22. D. Clowe, A. Gonzalez, M. Markevitch, Weak-Lensing Mass Reconstruction of the Interacting Cluster 1E 0657-558: Direct Evidence for the Existence of Dark Matter, *ApJ* **604**, 596-603 (2004).

23. L. L. R. Williams, P. Saha, Light/mass offsets in the lensing cluster Abell 3827: evidence for collisional dark matter?, *MNRAS* **415**, 448-460 (2011).

24. D. Harvey, *et al.*, Dark matter astrometry: accuracy of subhalo positions for the measurement of self-interaction cross-sections, *MNRAS* **433**, 1517-1528 (2013).

25. F. Gastaldello, *et al.*, Dark matter-baryons separation at the lowest mass scale: the Bullet Group, *MNRAS* **442**, L76-L80 (2014).

26. R. Massey, T. Kitching, D. Nagai, Cluster bulleticity, *MNRAS* **413**, 1709-1716 (2011).

27. J. G. Fernández-Trincado, J. E. Forero-Romero, G. Foex, T. Verdugo, V. Motta, The Abundance of Bullet Groups in ΛCDM, *ApJ* **787**, L34 (2014).

28. D. Harvey, *et al.*, On the cross-section of dark matter using substructure infall into galaxy clusters, *MNRAS* **441**, 404-416 (2014).

29. http://archive.stsci.edu/hst/ .

30. http://cxc.harvard.edu/cda/ .

31. D. Clowe, *et al.*, A Direct Empirical Proof of the Existence of Dark Matter, *ApJ* **648**, L109-L113 (2006).

32. S. Giodini, *et al.*, Stellar and Total Baryon Mass Fractions in Groups and Clusters Since Redshift 1, *ApJ* **703**, 982-993 (2009).

33. Planck Collaboration, Planck 2013 results. XVI. Cosmological parameters, *A&A* **571**, A16 (2014).

34. E. Bertin, S. Arnouts, SExtractor: Software for source extraction., *A&AS* **117**, 393-404 (1996).

35. R. Massey, *et al.*, An improved model of charge transfer inefficiency and correction algorithm for the Hubble Space Telescope, *MNRAS* **439**, 887-907 (2014).

36. A. M. Koekemoer, A. S. Fruchter, R. N. Hook, W. Hack, *HST Calibration Workshop : Hubble after the Installation of the ACS and the NICMOS Cooling System*, S. Arribas, A. Koekemoer, B. Whitmore, eds. (2003), p. 337.

37. J. D. Rhodes, *et al.*, The Stability of the Point-Spread Function of the Advanced Camera for Surveys on the Hubble Space Telescope and Implications for Weak Gravitational Lensing, *ApJS* **172**, 203-218 (2007).

38. M. R. George, *et al.*, Galaxies in X-Ray Groups. II. A Weak Lensing Study of Halo Centering, *ApJ* **757**, 2 (2012).

39. H. Martel, F. Robichaud, P. Barai, Major Cluster Mergers and the Location of the Brightest Cluster Galaxy, *ApJ* **786**, 79 (2014).

40. J. Rhodes, A. Refregier, E. J. Groth, Weak Lensing Measurements: A Revisited Method and Application toHubble Space Telescope Images, *ApJ* **536**, 79-100 (2000).

41. A. Leauthaud, *et al.*, Weak Gravitational Lensing with COSMOS: Galaxy Selection and Shape Measurements, *ApJS* **172**, 219-238 (2007).

42. J. E. Krist, R. N. Hook, F. Stoehr (2011), vol. 8127 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*.

43. E. Jullo, *et al.*, A Bayesian approach to strong lensing modelling of galaxy clusters, *New Journal of Physics* **9**, 447 (2007).

44. J. F. Navarro, C. S. Frenk, S. D. M. White, A Universal Density Profile from Hierarchical Clustering, *ApJ* **490**, 493 (1997).

45. A. V. Macciò, A. A. Dutton, F. C. van den Bosch, Concentration, spin and shape of dark matter haloes as a function of the cosmological model: WMAP1, WMAP3 and WMAP5 results, *MNRAS* **391**, 1940-1954 (2008).

46. D. Nagai, A. Vikhlinin, A. V. Kravtsov, Testing X-Ray Measurements of Galaxy Clusters with Cosmological Simulations, *ApJ* **655**, 98-108 (2007).

47. A. B. Mantz, *et al.*, Cosmology and astrophysics from relaxed galaxy clusters - II. Cosmological constraints, *MNRAS* **440**, 2077-2098 (2014).