


Galaxy kinematics and mass calibration in massive SZE-selected galaxy clusters to $z = 1.3$

R. Capasso^{1,2}  ^{1,2}★ A. Saro,^{1,2,3} J. J. Mohr,^{1,2,4} A. Biviano,³ S. Bocquet,^{5,6} V. Strazzullo,¹ S. Grandis,^{1,2} D. E. Applegate,⁵ M. B. Bayliss,⁷ B. A. Benson,^{5,8,9} L. E. Bleem,^{5,6,10} M. Brodwin,¹¹ E. Bulbul,⁷ J. E. Carlstrom,^{5,6,8,10,12} I. Chiu,¹³ J. P. Dietrich,^{1,2} N. Gupta,^{1,2} T. de Haan,^{14,15} J. Hlavacek-Larrondo,^{16,17,18} M. Klein,^{1,4} A. von der Linden,^{16,17,19,20} M. McDonald,⁷ D. Rapetti,^{1,2,21,22} C. L. Reichardt,²³ K. Sharon,²⁴ B. Stalder,²⁵ S. A. Stanford,^{26,27} A. A. Stark,²⁵ C. Stern^{1,2} and A. Zenteno²⁸

Affiliations are listed at the end of the paper

ABSTRACT

The galaxy phase-space distribution in galaxy clusters provides insights into the formation and evolution of cluster galaxies, and it can also be used to measure cluster mass profiles. We present a dynamical study based on ~ 3000 passive, non-emission-line cluster galaxies drawn from 110 galaxy clusters. The galaxy clusters were selected using the Sunyaev–Zel’dovich effect (SZE) in the 2500 deg² South Pole Telescope (SPT)–SZ survey and cover the redshift range $0.2 < z < 1.3$. We model the clusters using the Jeans equation, while adopting NFW mass profiles and a broad range of velocity dispersion anisotropy profiles. The data prefer velocity dispersion anisotropy profiles that are approximately isotropic near the centre and increasingly radial toward the cluster virial radius, and this is true for all redshifts and masses we study. The pseudo-phase-space density profile of the passive galaxies is consistent with expectations for dark matter particles and subhaloes from cosmological N -body simulations. The dynamical mass constraints are in good agreement with external mass estimates of the SPT cluster sample from either weak lensing, velocity dispersions, or X-ray Y_X measurements. However, the dynamical masses are lower (at the 2.2σ level) when compared to the mass calibration favoured when fitting the SPT cluster data to a Λ cold dark matter model with external cosmological priors, including cosmic microwave background anisotropy data from *Planck*. The discrepancy grows with redshift, where in the highest redshift bin the ratio of dynamical to SPT + Planck masses is $\eta = 0.63_{-0.08}^{+0.13} \pm 0.06$ (statistical and systematic), corresponding to a 2.6σ discrepancy.

Key words: galaxies: clusters: general – galaxies: evolution – galaxies: kinematics and dynamics – galaxies: statistics – cosmology: observations – large-scale structure of Universe.

1 INTRODUCTION

In the current paradigm of structure formation, haloes form through the gravitational collapse of overdense regions that are seeded by processes in the early universe. The formation of cold dark matter (CDM)-dominated haloes proceeds through a sequence of mergers and the accretion of surrounding material, leading to the formation of the galaxy groups and clusters we observe. Baryonic processes associated with the intracluster medium (ICM) and the galaxies also play a role, making galaxy clusters important laboratories for

investigations of structure formation and galaxy evolution as well as useful cosmological probes.

Studies of structure formation using cosmological N -body simulations have been used to demonstrate that haloes formed from collisionless CDM have, on average, a universal mass density profile (Navarro, Frenk & White 1996, 1997, hereinafter NFW). This profile is characterized by two parameters: the virial radius r_{200} ,¹ and the scale radius r_s , which is the radius at which the logarithmic

¹ r_{Δ} defines the sphere within which the cluster overdensity with respect to the critical density at the cluster redshift is Δ . Throughout this paper, we consider $\Delta = 200$ and refer to r_{200} simply as the virial radius.

* E-mail: Raffaella.Capasso@physik.uni-muenchen.de

slope of the density profile is -2 . Numerous observational studies have found the mass distributions of clusters to be well described by this model (e.g. Carlberg et al. 1997; van der Marel et al. 2000; Biviano & Girardi 2003; Katgert, Biviano & Mazure 2004; Umetsu et al. 2014).

Another interesting feature is the finding in N -body simulations that the quantity ρ/σ^3 , where ρ is the mass density and σ the velocity dispersion, has a power-law form. This quantity is known as pseudo-phase-space density (PPSD) profile, $Q(r)$, and its power-law form resembles that of the self-similar solution for halo collapse by Bertschinger (1985) and is thought of as a dynamical equivalent of the NFW mass density profile (Taylor & Navarro 2001). Others (Austin et al. 2005; Barnes et al. 2006) have suggested that the PPSD profile results from dynamical collapse processes, and should therefore be a robust feature of approximately virialized haloes that have undergone violent relaxation (Lynden-Bell 1967).

The galaxy population is more difficult to study in simulations, because of the overmerging problem, i.e. the premature destruction of dark matter (DM) haloes in the dense clusters environments in dissipationless N -body simulations (e.g. Moore, Katz & Lake 1996), and the additional baryonic physics that must be included. However, from the observational side the properties of the galaxy population and trends with mass and redshift can be readily measured and interpreted as long as: (1) selection effects are understood and (2) precise cluster mass measurements are available to ensure that the same portion of the virial region is being studied in clusters of all masses and redshifts. By comparing the galaxy properties to the expectations for collisionless particles studied through N -body simulations, one can characterize the impact of possible additional interactions beyond gravity that are playing a role in the formation of the galaxy population.

As an example, the radial distribution of galaxies in clusters is well fitted by an NFW model when clusters are stacked in the space of r/r_Δ (e.g. Lin, Mohr & Stanford 2004; Muzzin et al. 2007; van der Burg et al. 2014, 2015a; Zenteno et al. 2016). In cluster samples, extending to redshift $z \sim 1$, it is clear that the concentration c_Δ , defined as the ratio between r_Δ and r_s , varies dramatically from cluster to cluster, and that when stacked, the c_Δ varies systematically with the prevalence of star formation (Hennig et al. 2017, hereafter H17). The red, passively evolving galaxies have concentrations similar to those expected for the DM on these halo mass scales, while the star forming (SF), and presumably infalling blue galaxies are far less concentrated. The number of luminous cluster galaxies (magnitudes $m < m_* + 2$) within the virial region scales with cluster mass as $N \propto M^\alpha$ where $\alpha \sim 0.85$ (Lin et al. 2004), and this property appears to be unchanged since redshift $z \sim 1$ (H17). The departure from $\alpha = 1$ in this relation is puzzling, given that massive clusters accrete lower mass clusters and groups (Lin & Mohr 2004) and this is presumably evidence for galaxy destruction processes that are more efficient in the most massive haloes or for a mass accretion history that varies with mass on cluster scales (see discussion in Chiu et al. 2016b). There is evidence for an increase in the fraction of cluster galaxies that are dominated by passively evolving stellar populations since $z \sim 1$ (H17), and this observed increase provides constraints on the time-scales over which quenching of star formation occurs in those galaxies that are accreted by clusters (see e.g. McGee et al. 2009).

Understanding the dynamics of galaxy accretion into clusters, from either lower mass clusters and groups or even individual systems from within the surrounding low-density region, can shed additional light on galaxy evolution. A simulation-based study argues that satellite orbits should become marginally more radial at higher

redshifts, especially for systems with a higher host halo mass (Wetzel 2011). Probes of redshift trends in the orbital characteristics of cluster galaxies have already been carried out (Biviano & Poggianti 2009), providing some indication that passive galaxies have systematically different orbits at low and high redshifts. In other studies of high redshift, relatively low-mass systems, evidence has emerged that recently quenched galaxies have a preferred phase-space distribution that is different from that of passive galaxies (Muzzin et al. 2014; Noble et al. 2016).

In this paper, we attempt to build upon these studies by focusing on a dynamical analysis of galaxies within a large ensemble of Sunyaev–Zel’dovich effect (SZE) selected galaxy clusters extending to redshift $z \sim 1.3$. In contrast to these previous dynamical studies, our cluster sample has a well understood selection that does not depend on the galaxy properties, and the sample extends over a broad redshift range, allowing a cleaner examination of redshift trends. Moreover, each cluster has an SZE-based mass estimate with ~ 25 per cent uncertainty (Bocquet et al. 2015), enabling us to estimate virial radii r_Δ with ~ 8 per cent uncertainties and thereby ensuring that we are examining comparable regions of the cluster at all masses and redshifts.

Our goals are to study (1) whether there is evidence that the orbital characteristics of the passive galaxies are changing with redshift or mass in the cluster ensemble, (2) whether there is evidence within the galaxy dynamics for dynamical equilibrium and self-similarity with mass and redshift, and (3) whether the cluster mass constraints from our analysis are consistent with masses obtained through independent calibration in previously published South Pole Telescope (SPT) analyses.

Combining spectroscopic observations obtained at Gemini South, the Very Large Telescope (VLT) and the Magellan telescopes in a sample of 110 SPT-detected galaxy clusters, we construct a large sample of ~ 3000 passive cluster members, spanning the wide redshift range of $0.2 < z < 1.3$. With this data set, we carry out a Jeans analysis (e.g. Binney & Tremaine 1987) that adopts a framework of spherical symmetry and allows for a range of different velocity dispersion anisotropy profiles. Specifically, we use the Modeling Anisotropy and Mass Profiles of Observed Spherical Systems code (Mamon, Biviano & Boué 2013, hereafter MAMPOSS) to explore the range of models consistent with the data, and then we use the results to characterize the velocity dispersion anisotropy profile, to test for evidence of virialization with the PPSD profile and to probe for trends with cluster mass or redshift in both. Exploring a broad range of possible velocity dispersion anisotropy profiles then allows us to extract robust constraints on the cluster virial masses as well. Throughout this paper, we address a number of limitations that have to be taken into account, such as the degeneracy between the mass and the velocity anisotropy profiles (see Section 2), the assumptions of spherical symmetry and dynamical equilibrium, and the presence of foreground/background interloper galaxies projected on to the cluster virial region. Mamon et al. (2013) have tested the accuracy of MAMPOSS by analysing a sample of clusters extracted from numerical simulations, recovering r_{200} estimates with mean bias at ≤ 2.5 per cent and *rms* scatter of 6 per cent for kinematic samples with 500 tracers.

The paper is organized as follows: in Section 2, we give an overview of the theoretical framework. In Section 3, we summarize the data set used for our analysis. The results are presented in Section 4, where we discuss the outcome of our analysis of the velocity dispersion anisotropy profile, the PPSD profiles, the virial mass comparisons, and the impact of disturbed clusters on

our analysis. We present our conclusions in Section 5. Throughout this paper, we adopt a flat Λ CDM cosmology with the Hubble constant $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, and assume the matter density parameter $\Omega_M = 0.3$. The virial quantities are computed at radius r_{200} . All quoted uncertainties are equivalent to Gaussian 1σ confidence regions, unless otherwise stated.

2 THEORETICAL FRAMEWORK

The dynamical analysis implemented in this paper is based on the application of the Jeans equation to spherical systems (Binney & Tremaine 1987). In spherical coordinates, the Jeans equation can be written as

$$\frac{d(v\sigma_r^2)}{dr} + \frac{v}{r} [2\sigma_r^2 - (\sigma_\theta^2 + \sigma_\phi^2)] = -v \frac{d\Phi}{dr}, \quad (1)$$

where v is the number density profile of the tracer galaxy population, Φ is the gravitational potential, and $\sigma_r, \sigma_\theta, \sigma_\phi$ are the components of the velocity dispersion along the three spherical coordinates r, θ, ϕ . It is convenient to write this equation as

$$\frac{GM(r)}{r} = -\sigma_r^2 \left(\frac{d \ln v}{d \ln r} + \frac{d \ln \sigma_r^2}{d \ln r} + 2\beta \right), \quad (2)$$

where $M(r)$ is the enclosed mass within radius r , G is Newton's constant, $\beta \equiv 1 - (\sigma_\theta^2/\sigma_r^2)$ is the velocity dispersion anisotropy that is generically a function of radius, and $\sigma_\theta = \sigma_\phi$.

In principle, it is therefore possible to use equation (2) to estimate the mass distribution $M(r)$ of the system. However, we as external observers can measure only projected quantities, such as the surface density profile of the tracer population $\Sigma(R)$ and the line-of-sight (LOS) velocity distribution (or, alternatively, the LOS velocity dispersion $\sigma_{\text{LOS}}(R)$). These two observed functions are not sufficient to derive a unique mass model – at least within the context of the typical observational uncertainties where full knowledge of the LOS velocity distribution is lacking (Merritt 1987). This degeneracy between the mass and the velocity anisotropy profiles can be addressed in several ways.

2.1 Dynamical analysis with MAMPOSSt

A first method consists of assuming that, at a given projected radius, the LOS velocity distribution can be described by a Gaussian (Strigari et al. 2008; Wolf et al. 2010). However, because the true distribution may deviate from Gaussianity – including the case of anisotropic systems (Merritt 1987) – the constraints from this approach on the anisotropy are weak (Walker et al. 2009). A step forward, therefore, consists of analysing the kurtosis of the LOS velocity distribution. This was found to be a powerful tool to break the mass–anisotropy degeneracy (Merritt 1987; Gerhard 1993; van der Marel & Franx 1993; Zabludoff, Franx & Geller 1993; Łokas 2002; Łokas & Mamon 2003). Much effort has been put into further constraining the anisotropy taking into account more of the information contained in the projected phase-space velocity distribution as a function of projected radius, considering, for example, the full set of even moments (Kronawitter et al. 2000; Wojtak et al. 2008, 2009).

In this work, we fit the whole projected phase-space velocity distribution using the MAMPOSSt code (Mamon et al. 2013). We use this code to determine the mass and anisotropy profiles of a cluster in parametrized form by performing a likelihood exploration to the distribution of the cluster galaxies in projected phase space, constraining the parameters describing these two profiles. This method

is based on the assumption of spherical symmetry, and adopts a Gaussian as the shape of the distribution of the 3D velocities without demanding Gaussian LOS velocity distributions. We emphasize that it does not assume that light traces mass, allowing the scale radius r_s of the total mass distribution to differ from that of the galaxy distribution. For more details on the code, we refer the reader to Mamon et al. (2013).

MAMPOSSt requires parametrized models for the number density, the mass, and the velocity anisotropy profiles, without any limitation on the choice of these models. We will address the issue of the number density profile in Section 3.3.3.

The current estimates of the systematic error on MAMPOSSt-derived dynamical masses are ≈ 10 per cent, where this number comes from an analysis of clusters extracted from numerical simulations (see Mamon et al. 2013). In their study, they find that for the dynamical tracers defined to lie within a sphere of r_{100} that the estimate of the virial radius r_{200} is biased at ≤ 3.3 per cent (see table 2; Mamon et al. 2013). Thus, as a prior on the virial mass M_{200} bias, we adopt a Gaussian with $\sigma = 10$ per cent centred at no bias.

2.2 Mass and anisotropy profiles

For the mass profile in our analyses, we consider five models, namely the Navarro, Frenk, and White profile (NFW; Navarro, Frenk & White 1996), the Einasto profile (Einasto 1965), the Burkert profile (Burkert 1995), the Hernquist profile (Hernquist 1990), and the Softened Isothermal Sphere (SIS; Geller, Diaferio & Kurtz 1999). All these models have been applied to galaxy clusters in previous works (e.g. Mohr et al. 1996; Rines et al. 2003; Katgert et al. 2004; Rines & Diaferio 2006; Biviano et al. 2006). As we show later in Section 4, our data cannot distinguish among these different mass profiles, and so in the analyses described below we adopt the NFW model

$$\rho(r) = \rho_0 \left(\frac{r}{r_s} \right)^{-1} \left(1 + \frac{r}{r_s} \right)^{-2} \quad (3)$$

where ρ_0 is the central density and r_s is the scale radius where the logarithmic derivative of the density profile reaches -2 .

For the velocity anisotropy profile, we consider the following five models that have been used in previous MAMPOSSt analyses:

C a radially constant anisotropy model,

$$\beta_C(r) = \theta_\beta; \quad (4)$$

T the Tiert anisotropy profile (Tiert et al. 2007),

$$\beta_T(r) = \theta_\beta \frac{r}{r + r_s}, \quad (5)$$

which is isotropic at the centre and characterized by the anisotropy value θ_β at large radii. The transition radius r_s is the scale radius of the NFW density profile;

O a model with anisotropy of opposite sign at the centre and at large radii,

$$\beta_O(r) = \theta_\beta \frac{r - r_s}{r + r_s}; \quad (6)$$

ML the Mamon & Łokas (2005) profile,

$$\beta_{ML}(r) = 0.5 \frac{r}{r + \theta_\beta}; \quad (7)$$

OM the Osipkov–Merritt anisotropy profile (Osipkov 1979; Merritt 1985),

$$\beta_{\text{OM}}(r) = \frac{r^2}{r^2 + \theta_\beta^2}. \quad (8)$$

Summarizing, we run MAMPOSSt with three free parameters: the virial radius r_{200} , the scale radius r_s of the mass distribution, and a velocity anisotropy parameter θ_β . This parameter represents the usual $\beta = 1 - (\sigma_\theta^2/\sigma_r^2)$ for the first three models (C, T, and O), while for the M \mathbb{L} and OM models it defines a characteristic radius $\theta_\beta = r_\beta$. The maximum likelihood solutions are obtained using the NEWUOA software (Powell 2006) and are shown in Section 4.

3 CLUSTER DATA

In this section, we present the cluster sample, spectroscopic data, and the method for constructing composite clusters.

3.1 Cluster sample

The cluster sample analysed in this study consists of galaxy clusters detected with the SPT, a 10-m telescope located within 1 km of the geographical South Pole, observing in three mm-wavebands centred at 95, 150, and 220 GHz (see Carlstrom et al. 2011). The SPT–SZ survey, imaging 2500 deg² of the southern sky, has produced data that are used to select galaxy clusters via their thermal SZE signature in the 95 and 150 GHz maps. This SZE signature arises through the inverse Compton scattering of the cosmic microwave background (CMB) photons and the hot ICM (Sunyaev & Zel’dovich 1972). We refer the reader to Schaffer et al. (2011) for details on the survey strategy and data processing. All of the clusters studied in this work have a high detection significance ($\xi > 4.8$; Song et al. 2012; Bleem et al. 2015) and have spectroscopic redshifts.

A major advantage of selecting clusters with the SZE rather than other cluster observables lies in the fact that the surface brightness of the SZE signature is independent of the cluster redshift, which together with the expected temperature and density evolution of the cluster ICM at fixed virial mass and the changing solid angle of clusters with redshift lead an SZE signal-to-noise-selected sample to be approximately mass-limited (Haiman, Mohr & Holder 2001; Holder, Haiman & Mohr 2001).

Because we plan to carry out a dynamical analysis, the cluster sample we analyse includes only those SPT systems with spectroscopic follow-up. This spectroscopic subsample is not a signal-to-noise-selected sample, but rather is the largest sample of SPT-selected clusters we could assemble for the analysis. We plot the sample of 110 clusters in the space of redshift versus mass in Fig. 1, together with the full distribution of the 2500 deg² SPT–SZ sample with the latest available redshifts (Bayliss et al. 2016; Khullar et al. 2018; Strazzullo et al. 2018, Bocquet et al., in preparation). Note that in comparison to the SPT–SZ cosmology sample (de Haan et al. 2016), the median redshift and mass for our sample is 0.56 and $7.26 \times 10^{14} M_\odot$ as compared to 0.55 and $6.08 \times 10^{14} M_\odot$. Approximately 6 per cent of the sample lies at $z > 1$, as compared to 9 per cent of the full SPT–SZ $\xi > 5$ sample (with updated redshifts from Bocquet et al., in preparation). Thus, in comparison to the complete SPT–SZ sample, the subset of those clusters we study here have somewhat higher masses and are somewhat under-represented at high redshift. Full information on the cluster sample is provided in Table 1. From left to right,

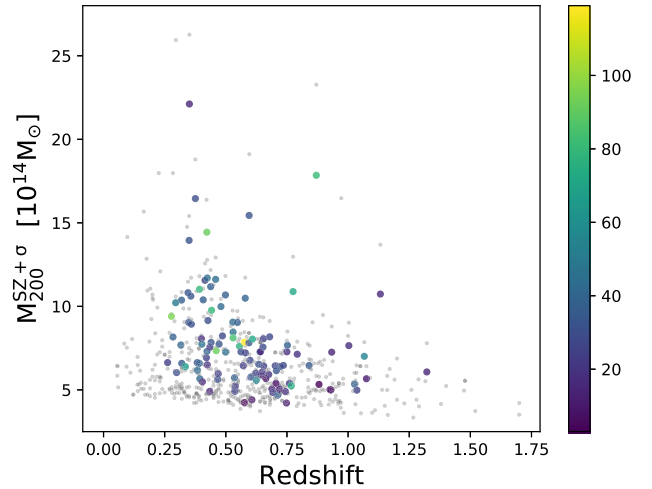


Figure 1. Masses versus redshifts of the cluster sample. The coloured dots are coded according to the number of member galaxies in each cluster (see colour bar on right). Overplotted in grey is the full distribution of the 2500 deg² SPT–SZ sample.

the columns correspond to the SPT cluster designation, the total number of passive galaxy redshifts available, the number of passive galaxy member redshifts, the cluster redshift, the SZE-based mass $M_{200}^{SZ+\sigma}$, and the corresponding virial radius $R_{200}^{SZ+\sigma}$ (described in Section 3.3.1).

3.2 Spectroscopic sample

We use spectroscopic follow-up including data taken using the Gemini Multi Object Spectrograph (Hook et al. 2004) on Gemini South, the Focal Reducer and low dispersion Spectrograph (Appenzeller et al. 1998) on VLT Antu, the Inamori Magellan Areal Camera and Spectrograph (Dressler et al. 2006) on Magellan Baade, and the Low Dispersion Survey Spectrograph (Allington-Smith et al. 1994) on Magellan Clay.

We combine the data sets presented in Ruel et al. (2014) and Bayliss et al. (2016) with the one described in Sifón et al. (2013), obtaining a sample of 4695 redshifts. In what follows, we split the spectroscopic galaxy sample according to the spectral features into two subsamples of those with emission lines (EL) and those without emission lines (nEL), adopting the determinations presented in the aforementioned analyses.

Due to the observational strategy of these surveys, which targeted red-sequence (RS) galaxies, the number of nEL galaxies (3834) greatly exceeds the number of EL galaxies. Because EL and nEL galaxies are characterized by different spatial and kinematic properties (see e.g. Mohr et al. 1996; Biviano et al. 1997; Dressler et al. 1999; Biviano et al. 2002, 2016; Bayliss et al. 2017; Hennig et al. 2017), in the analysis that follows we focus only on the passive nEL population, for which we have sufficient statistics to properly carry out the dynamical analysis. In Fig. 2, we show the normalized distribution of galaxies and clusters as a function of redshift.

3.3 Construction of composite clusters

To enable a precise determination of the cluster masses and the velocity anisotropy profiles, we cannot rely on the few spectroscopic members in the individual clusters of our sample. For this reason,

Table 1. Cluster spectroscopic sample: we present the cluster name, total number of passive galaxies in the spectroscopic sample N_{tot} , members used in our analysis after interloper rejection N_{mem} , cluster redshift z_{clus} , and cluster mass $M_{200}^{\text{SZ}+\sigma}$ and corresponding virial radius $R_{200}^{\text{SZ}+\sigma}$ from Bocquet et al. (2015).

Cluster	N_{tot}	N_{mem}	z_{clus}	$M_{200}^{\text{SZ}+\sigma}$ ($10^{14}M_{\odot}$)	$R_{200}^{\text{SZ}+\sigma}$ (Mpc)
SPT-CL J0000–5748	27	24	0.702	6.44	1.37
SPT-CL J0013–4906	37	36	0.408	10.39	1.81
SPT-CL J0014–4952	38	27	0.752	7.67	1.42
SPT-CL J0033–6326	28	14	0.599	6.77	1.45
SPT-CL J0037–5047	37	17	1.026	5.34	1.13
SPT-CL J0040–4407	33	33	0.350	13.95	2.04
SPT-CL J0102–4603	35	16	0.841	6.46	1.30
SPT-CL J0102–4915	81	80	0.870	17.84	1.80
SPT-CL J0106–5943	35	26	0.348	9.04	1.76
SPT-CL J0118–5156	8	7	0.705	5.40	1.29
SPT-CL J0123–4821	26	18	0.655	6.43	1.40
SPT-CL J0142–5032	28	23	0.679	8.17	1.50
SPT-CL J0200–4852	45	34	0.499	7.26	1.55
SPT-CL J0205–5829	15	8	1.322	6.07	1.06
SPT-CL J0205–6432	19	12	0.744	4.90	1.23
SPT-CL J0212–4657	26	20	0.654	8.06	1.51
SPT-CL J0232–5257	77	61	0.556	7.61	1.54
SPT-CL J0233–5819	9	9	0.664	5.68	1.33
SPT-CL J0234–5831	24	21	0.415	11.56	1.87
SPT-CL J0235–5121	96	82	0.278	9.41	1.84
SPT-CL J0236–4938	66	63	0.334	6.39	1.58
SPT-CL J0240–5946	19	17	0.400	8.06	1.67
SPT-CL J0243–4833	39	37	0.498	10.68	1.76
SPT-CL J0243–5930	32	25	0.634	6.53	1.42
SPT-CL J0252–4824	27	22	0.421	6.92	1.57
SPT-CL J0254–5857	37	32	0.438	11.18	1.83
SPT-CL J0257–5732	15	14	0.434	4.91	1.39
SPT-CL J0304–4401	45	35	0.458	11.62	1.84
SPT-CL J0304–4921	79	72	0.392	11.02	1.85
SPT-CL J0307–6225	26	17	0.580	7.21	1.49
SPT-CL J0310–4647	33	28	0.707	6.15	1.35
SPT-CL J0317–5935	25	18	0.469	5.97	1.47
SPT-CL J0324–6236	19	9	0.750	7.28	1.40
SPT-CL J0330–5228	80	71	0.442	9.77	1.75
SPT-CL J0334–4659	29	25	0.486	8.23	1.62
SPT-CL J0346–5439	85	79	0.530	8.11	1.59
SPT-CL J0348–4515	31	24	0.359	8.93	1.75
SPT-CL J0352–5647	22	16	0.649	6.07	1.37
SPT-CL J0356–5337	26	5	1.036	4.98	1.10
SPT-CL J0403–5719	31	24	0.467	5.62	1.44
SPT-CL J0406–4805	28	26	0.736	6.46	1.35
SPT-CL J0411–4819	45	42	0.424	11.70	1.87
SPT-CL J0417–4748	40	30	0.579	10.49	1.69
SPT-CL J0426–5455	15	11	0.642	7.28	1.46
SPT-CL J0433–5630	24	18	0.692	4.84	1.25
SPT-CL J0438–5419	92	87	0.422	14.44	2.00
SPT-CL J0449–4901	20	16	0.792	7.13	1.37
SPT-CL J0456–5116	31	20	0.562	7.27	1.51
SPT-CL J0509–5342	93	88	0.461	7.34	1.58
SPT-CL J0511–5154	18	14	0.645	5.90	1.36
SPT-CL J0516–5430	51	47	0.295	10.22	1.88
SPT-CL J0521–5104	21	21	0.675	5.92	1.35
SPT-CL J0528–5300	75	63	0.768	5.24	1.25
SPT-CL J0533–5005	4	4	0.881	5.33	1.20
SPT-CL J0534–5937	3	3	0.576	4.24	1.25
SPT-CL J0540–5744	24	17	0.760	5.38	1.26
SPT-CL J0542–4100	36	29	0.640	7.27	1.46
SPT-CL J0546–5345	54	49	1.066	7.01	1.22
SPT-CL J0549–6205	31	26	0.375	16.45	2.13

Table 1 – *continued*

Cluster	N_{tot}	N_{mem}	z_{clus}	$M_{200}^{\text{SZ}+\sigma}$ ($10^{14}M_{\odot}$)	$R_{200}^{\text{SZ}+\sigma}$ (Mpc)
SPT-CL J0551–5709	39	30	0.423	7.30	1.60
SPT-CL J0555–6406	34	27	0.345	10.82	1.88
SPT-CL J0559–5249	72	67	0.609	8.04	1.53
SPT-CL J0655–5234	33	30	0.472	7.74	1.60
SPT-CL J2017–6258	38	35	0.535	5.73	1.41
SPT-CL J2020–6314	27	18	0.537	4.91	1.34
SPT-CL J2022–6323	34	29	0.383	6.63	1.57
SPT-CL J2026–4513	15	11	0.688	5.05	1.27
SPT-CL J2030–5638	42	36	0.394	5.67	1.48
SPT-CL J2032–5627	39	32	0.284	8.16	1.75
SPT-CL J2035–5251	42	29	0.529	9.05	1.65
SPT-CL J2040–5725	7	4	0.930	5.00	1.15
SPT-CL J2043–5035	33	21	0.723	6.40	1.36
SPT-CL J2056–5459	13	11	0.718	5.06	1.26
SPT-CL J2058–5608	10	6	0.607	4.40	1.25
SPT-CL J2100–4548	37	18	0.712	4.68	1.23
SPT-CL J2106–5844	16	16	1.132	10.74	1.37
SPT-CL J2115–4659	31	28	0.299	6.03	1.57
SPT-CL J2118–5055	55	33	0.624	5.57	1.35
SPT-CL J2124–6124	23	21	0.435	7.56	1.61
SPT-CL J2130–6458	41	40	0.316	7.69	1.69
SPT-CL J2135–5726	31	30	0.427	9.15	1.72
SPT-CL J2136–4704	19	19	0.424	6.50	1.53
SPT-CL J2136–6307	6	6	0.926	5.02	1.15
SPT-CL J2138–6008	39	32	0.319	10.37	1.87
SPT-CL J2140–5727	17	11	0.404	5.49	1.46
SPT-CL J2145–5644	43	35	0.480	9.99	1.73
SPT-CL J2146–4633	15	7	0.933	7.26	1.30
SPT-CL J2146–4846	26	25	0.623	6.00	1.38
SPT-CL J2146–5736	34	23	0.602	5.77	1.38
SPT-CL J2148–6116	24	24	0.571	6.45	1.45
SPT-CL J2155–6048	23	19	0.539	5.23	1.36
SPT-CL J2159–6244	38	36	0.391	6.59	1.56
SPT-CL J2222–4834	29	25	0.652	7.56	1.48
SPT-CL J2232–5959	34	26	0.595	7.82	1.53
SPT-CL J2233–5339	33	28	0.440	7.84	1.62
SPT-CL J2248–4431	15	14	0.351	22.11	2.37
SPT-CL J2300–5331	25	21	0.262	6.63	1.64
SPT-CL J2301–5546	12	8	0.748	4.21	1.17
SPT-CL J2306–6505	46	42	0.530	8.46	1.61
SPT-CL J2325–4111	33	27	0.357	10.61	1.86
SPT-CL J2331–5051	119	108	0.576	7.85	1.54
SPT-CL J2332–5358	47	45	0.402	7.74	1.64
SPT-CL J2335–4544	37	33	0.547	9.04	1.63
SPT-CL J2337–5942	28	19	0.775	10.88	1.59
SPT-CL J2341–5119	18	13	1.003	7.65	1.29
SPT-CL J2342–5411	12	7	1.075	5.67	1.14
SPT-CL J2344–4243	33	25	0.595	15.44	1.91
SPT-CL J2351–5452	42	30	0.384	6.18	1.53
SPT-CL J2355–5055	36	33	0.320	6.59	1.61
SPT-CL J2359–5009	44	37	0.775	5.17	1.24

we either create composite clusters with much more dynamical information or, as for the results presented in Section 4.4.2, fit to a common model across the cluster ensemble by combining the likelihoods associated with each individual cluster. The composite cluster approach has been adopted in previous analyses (Carlberg et al. 1997; van der Marel et al. 2000; Katgert et al. 2004; Biviano & Poggianti 2009), and it is supported by cosmological simulations that predict cosmological haloes can be characterized by a universal mass density profile with a concentration that de-

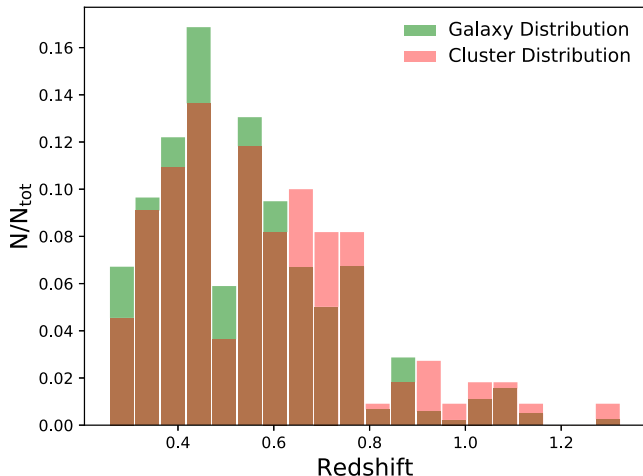


Figure 2. Redshift distribution of member galaxies (green) and clusters (red) in our sample, normalized to the total number of objects in each case. The two distributions are similar, but a clear trend to have fewer galaxies per cluster at high redshift is visible.

Table 2. Characteristics of the composite clusters in redshift: columns show the bin number, redshift range, mean redshift, number of included clusters N_{clus} , number of total galaxies N_{tot} , number of member galaxies N_{mem} , and mean SPT-based mass $\langle M_{200}^{\text{SZ}+\sigma} \rangle$.

Bin	Redshift range	$\langle z \rangle$	N_{clus}	N_{tot}	N_{mem}	$\langle M_{200}^{\text{SZ}+\sigma} \rangle$ ($10^{14} M_{\odot}$)
1	0.26–0.38	0.33	18	712	593	9.23
2	0.39–0.44	0.42	19	744	644	9.50
3	0.46–0.56	0.51	17	758	615	7.98
4	0.56–0.71	0.62	30	904	675	7.31
5	0.71–1.32	0.86	26	716	459	7.60
-	0.26–1.32	0.55	110	3834	2966	8.28

pendently on the halo mass (Navarro et al. 1997). We return to the method using a combination of individual cluster likelihoods in Section 4.4.2.

We create composite clusters by combining cluster subsamples within different mass and redshift ranges in such a way that each subsample includes an adequate number of members. Considering that tests done using MAMPOSSt on cosmological simulations indicate that with 500 tracers the code is able to recover the mass and anisotropy parameters with a suitably small uncertainty (Mamon et al. 2013), we create similarly sized subsamples. With this approach, we can construct five composite clusters selected either as redshift or mass subsamples.

Table 2 lists the characteristics of the composite clusters created within redshift bins. From left to right, the columns correspond to the bin number, the redshift range of the clusters combined to create the composite system in that bin, the mean redshift of these clusters, the number of clusters, the total number of galaxy redshifts, the number of cluster members with redshifts, and the average SPT-based mass of the individual clusters in the sample. In the following subsections, we describe in detail how these composite clusters are created and how interlopers are removed from the spectroscopic samples.

3.3.1 Rescaling observables with SZE-based mass estimates

To create composite clusters, we choose the brightest cluster galaxy (BCG) as the cluster centre. We determine these positions using results from previous and ongoing work (Song et al. 2012; Sifón et al. 2013; Stalder et al., in preparation), allowing us to calculate a projected separation R for each galaxy from the cluster centre. We extract the rest-frame LOS velocity v_{rf} from the galaxy redshift z and equivalent velocity $v(z)$ as $v_{\text{rf}} \equiv (v(z) - v(z_c))/(1 + z_c)$, where z_c is the cluster redshift.

Because our clusters span a wide range of mass and redshift, we must scale the galaxy-projected cluster distances R and LOS velocities to the values of a fiducial cluster mass using an estimate of the individual cluster virial radius and velocity. For convenience, in the composite cluster, we use the mean mass $\langle M_{200} \rangle$ and the mean redshift $\langle z \rangle$ of the ensemble of clusters in that bin to calculate the associated $\langle R_{200} \rangle$ and $\langle v_{200} \rangle$ that we adopt as fiducial values for the bin. Thus, the rescaled observables for each galaxy i within a cluster j in a particular bin are $R_{i,j} = R_i \langle R_{200} \rangle / R_{200,j}$ and $v_{i,j} = v_{\text{rf},i} \langle v_{200} \rangle / v_{200,j}$, where $R_{200,j}$ and $v_{200,j}$ are the virial radius and circular velocity, respectively, of cluster j . The circular velocity is defined using the virial condition $v_{200}^2 = GM_{200}/R_{200}$.

To perform the normalization, we need precise estimates of the cluster mass. Given that the SPT sample is SZE selected, we adopt the SZE observable ξ , which is the detection signal to noise and is correlated to the underlying cluster virial mass (Anderson et al. 2011; Benson et al. 2013), as the source of our cluster mass estimates. SPT masses derived from ξ have a statistical uncertainty that depends on the intrinsic scatter in the ξ –mass relation and on the observational uncertainty in the signal-to-noise ratio of the detection; together, these lead characteristically to ~ 20 per cent statistical uncertainty in the cluster virial mass. In addition, there are systematic uncertainties remaining from the calibration of the mass–observable relation that are currently at the ~ 15 per cent level. As discussed further below, the systematic mass-scale uncertainties at this level have no impact on our analysis, but the cluster-to-cluster statistical mass uncertainty of ~ 20 per cent introduces a corresponding uncertainty in R_{200} and v_{200} of ~ 7 per cent.

The mass–observable relation for the SPT sample can be calibrated in different ways, and in the current analysis, we use two different approaches. The first approach uses direct cluster mass measurements from weak lensing, velocity dispersions, X-ray measurements to calibrate the SZE–mass observable relation. The second approach fits the SPT cluster distribution in signal-to-noise ξ and redshift z , and the mass–observable relation to a flat Λ CDM model with external cosmological priors from other data sets. These constraints from, e.g. CMB anisotropy, baryon acoustic oscillations, etc., effectively constrain the mass–observable relation so that the cluster mass function implied by SPT cluster distribution in ξ and z is consistent with that expected given the external cosmological priors. As an example (see Bocquet et al. 2015), the inclusion of external cosmological constraints from Planck CMB in the SPT mass–observable calibration leads to a shift of cluster masses to ~ 25 per cent higher values in comparison to the case where the mass calibration is undertaken with direct cluster mass measurements from velocity dispersions.

Given the need to have an initial mass estimate to enable the stacking of the clusters for a dynamical analysis, we carry out the analysis with initial SPT-based masses calibrated in two different ways. The first uses the SPT cluster counts together with direct cluster

mass information from velocity dispersions. Here, we refer to these masses as $M_{200}^{SZ+\sigma}$. The second uses the SPT cluster counts together with external cosmological priors from *Planck* CMB anisotropy and distance measurements (Baryon Acoustic Oscillations and supernovae), specifically we use mass calibration inferred from the SPTCL +Planck +WP+BAO + SNIa data set in Bocquet et al. (2015). We refer to these masses as $M_{200}^{SZ+Planck}$. In both cases, the SPT mass–observable relation is calibrated using M_{500} (Bocquet et al. 2015), and so we transform from M_{500} by assuming an NFW model with a concentration c_{200} sampled from cosmological N -body simulations (Duffy et al. 2008). The r_{200} and v_{200} are then easily obtained, depending only on the cluster redshift and the adopted cosmology.

As discussed in Section 4.4.1, the dynamical mass measurements of the composite clusters have only a weak dependence on the mass scale of the velocities and radii used to build the composite clusters; specifically, the ~ 25 per cent shift between the masses $M_{200}^{SZ+\sigma}$ and $M_{200}^{SZ+Planck}$ used to build the two sets of composite clusters has no significant impact on the final dynamical masses.

Because we are carrying out a Jeans analysis, which is based on the assumption of dynamical equilibrium, we restrict our analysis to the cluster virial region ($R \leq r_{200}$). Moreover, we exclude the very central cluster region ($R \leq 50$ kpc). In fact, the composite clusters are centred on the BCG, which we exclude from the dynamical analysis. We note also that in the composite clusters the characteristic asphericity of individual clusters is averaged down, leading to a combined system that is approximately spherical in agreement with the dynamical model we are employing.

3.3.2 Interloper rejection

One benefit of constructing composite clusters is that we can more easily identify and reject some interloper galaxies, i.e. galaxies that are projected inside the cluster virial region, but do not actually lie inside it. We do so by using the ‘CLEAN’ method (Mamon et al. 2013), which is based on the identification of the cluster members on the basis of their projected phase-space location (R, v). The LOS velocity dispersion σ_{LOS} of the composite cluster is used to estimate the cluster mass using a scaling relation calibrated using numerical simulations (e.g. Saro et al. 2013), and an NFW mass profile with concentration sampled from the theoretical mass–concentration relation of Macciò, Dutton & van den Bosch (2008). Thereafter, assuming the M Σ velocity anisotropy profile model, and given the $M(r)$ of the cluster, an LOS velocity dispersion profile $\sigma_{LOS}(R)$ is calculated and used to iteratively reject galaxies with $|v_{rfl}| > 2.7\sigma_{LOS}$ at any clustercentric distance (see Mamon, Biviano & Murante 2010; Mamon et al. 2013). As an example, we present in Fig. 3, the location of galaxies in projected phase space with the identification of cluster member galaxies for the full sample.

We note that even after cleaning the dynamical sample, it is still contaminated to some degree by interlopers. One reason for this is that galaxies near the cluster turnaround radius (i.e. well outside the virial radius) will have small LOS velocities. Therefore, if those galaxies are projected on to the cluster virial region they simply cannot be separated from the galaxies that actually lie within the cluster virial radius. Analysis of cosmological N -body simulations shows that when passive galaxies are selected this contamination is characteristically ~ 20 per cent (Saro et al. 2013) for SPT mass-scale clusters and represents a systematic, whose effects on the dynamical

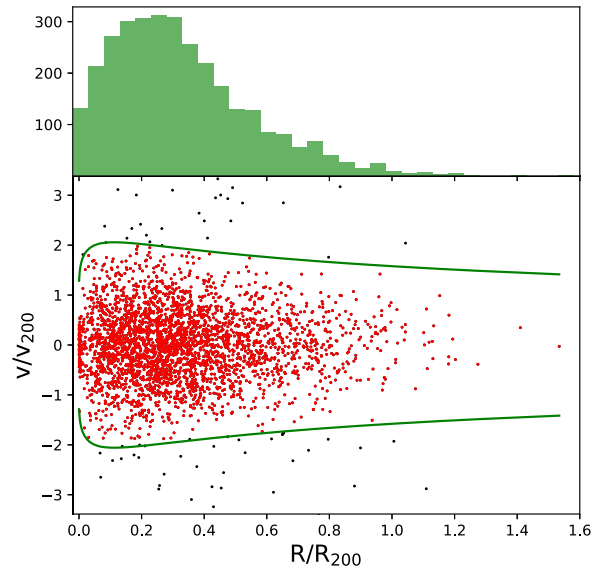


Figure 3. The projected phase-space diagram (below) in the rescaled coordinates ($R/R_{200}, v/v_{200}$) for the full sample. Green lines represent the radially dependent $2.7\sigma_{LOS}$ cut used to reject interlopers (indicated by black dots). The radial distribution of the member galaxies with measured redshifts is shown in the upper panel.

analysis require further exploration (see also discussion in Old et al. 2015).

3.3.3 Galaxy number density profiles

We cannot simply adopt the spectroscopic sample to measure the number density profiles of the galaxy population, because the structure of the spectrographs and our observing strategy of visiting each cluster typically with only two masks will generally result in a radially dependent incompleteness. This incompleteness, if not accounted for, would affect the determination of the cluster-projected number density profile $\Sigma(R)$, related to the 3D number density profile $\nu(r)$ through the Abel inversion (Binney & Tremaine 1987). As only the logarithmic derivative of $\nu(r)$ enters the Jeans equation (see equation 2), the absolute normalization of the galaxy number density profile has no impact on our analysis. However, a radially dependent incompleteness in the velocity sample would impact our analysis.

Thus, for our analysis, we rely on a study of the galaxy populations in 74 SZE-selected clusters from the SPT survey that were imaged as part of the Dark Energy Survey (DES) Science Verification phase (H17). This complete sky-area-selected subsample of the SPT–SZ cluster sample has a redshift and mass distribution that are consistent with our sample. The number density profile of the RS population of galaxies is found to be well fitted by an NFW model out to radii of $4r_{200}$, with a concentration of cluster galaxies $c_{gal} = 5.37^{+0.27}_{-0.24}$. The non-RS population is much less centrally concentrated, but in our analysis, we are focused only on the nEL galaxy populations. No statistically significant redshift or mass trends in the galaxy concentration were identified. Therefore, we adopt the number density profile described by an NFW profile with the above-mentioned value of c_{gal} and a scale radius $r_{s,gal} = R_{200}/c_{gal}$. Implicit in this approach is the assumption that the dynamical properties of our spectroscopic sample are consistent with the

dynamical properties of the RS galaxy population used to measure the radial profiles.

We test this assumption of a fixed c_{gal} by including in the likelihood an additional parameter $r_{s, \text{gal}}$, given by the ratio R_{200}/c_{gal} , where c_{gal} is given by H17

$$c_{\text{gal}} = A \left(\frac{M_{200}}{M_{\text{piv}}} \right)^B \left(\frac{1+z}{1+z_{\text{piv}}} \right)^C, \quad (9)$$

where, for RS galaxies, $A = 5.47 \pm 0.53$, $B = -0.01 \pm 0.10$, and $C = 0.15 \pm 0.30$. We choose mass and redshift pivot points $M_{\text{piv}} = 5 \times 10^{14}$ and $z_{\text{piv}} = 0.53$, respectively, corresponding to the median mass and redshift of our sample. With this additional parameter, fitting for A , B , and C with Gaussian priors corresponding to the uncertainties listed above, we adopt the Markov Chain Monte Carlo method for sampling from the probability distribution, utilizing the EMCEE code (Foreman-Mackey et al. 2013). Our analysis provides values of the scale and virial radii consistent within 1σ of the measured values reported in H17.

4 RESULTS

In this section, we present the results of our dynamical analysis. In the first subsection, we explore the mass profiles and present several crucial pieces of information that are needed to understand the results in the following sections. The following subsections present our measurements of the anisotropy profiles, the PPSD profiles, and the dynamical mass constraints. The final subsection examines the impact on our analysis of ongoing mergers.

4.1 Testing mass and anisotropy profiles

As a first step, we use MAMPOSSt to analyse the composite cluster constructed using the full dynamical data set on all mass and anisotropy models described in Section 2.1. Our goal is to determine which mass models are appropriate for this study. For each mass model, we explore all the different anisotropy models.

4.1.1 Comparison of different mass profiles

To quantitatively differentiate among the mass models (NFW, Einasto, Burkert, Hernquist, and SIS), we compare the likelihood of the data being consistent with the model, employing the so-called Bayes factor. This factor is the marginalized likelihood of the model (see Hoeting et al. 1999, and references therein). It is computed by averaging the likelihood in a specific model M_j over the available prior range $P(\theta_j | M_j)$, reading

$$\mathcal{L}(D | M_j) = \int \mathcal{L}(D | \theta_j, M_j) P(\theta_j | M_j) d\theta_j, \quad (10)$$

where $\mathcal{L}(D | \theta_j, M_j)$ is the likelihood of the data D given the model parameters θ_j .

Considering J models M_1, \dots, M_J , we define the Bayes factor $B_{j, \text{max}}$ of each model j by normalizing by the most probable model, yielding

$$B_{j, \text{max}} = \frac{\mathcal{L}(D | M_j)}{\mathcal{L}(D | M_{\text{max}})}, \quad (11)$$

where M_{max} indicates the model with the highest marginalized likelihood. The averaged posterior distribution on the parameters common to all models is then simply given by the weighted average

Table 3. Likelihood ratio (or Bayes factor; see equation 11) for the NFW, Burkert, Einasto, Hernquist, and SIS mass profiles is presented for each anisotropy profile analysed using the composite cluster constructed using the full dynamical sample. As noted in the text, we cannot discard any of the mass models aside from the SIS.

$\beta(r)$ model	Mass model				
	NFW	Burkert	Einasto	Hernquist	SIS
C	0.52	0.19	0.47	1.00	9.04×10^{-4}
ME	0.68	0.07	1.00	0.89	7.46×10^{-4}
OM	0.45	0.18	0.46	1.00	4.44×10^{-6}
O	0.90	0.03	1.00	0.43	
T	0.75	0.06	0.84	1.00	

of the posterior distributions of each model, with the Bayes factor as weight. Models can also be rejected using their Bayes factors. According to the Jeffreys scale (Jeffreys 1998), M_j is considered decisively rejectable if $B_{j, \text{max}} < 0.01$.

Table 3 contains the measured values. One can see that for each anisotropy model there is one preferred model (Bayes factor 1.0), but that the likelihood ratios for all but one of the mass profile models are on the order of 1. Note that for the SIS profile, we can only consider three of the five anisotropy models, because the T and O models need the value of the scale radius in the density profile, and that is not uniquely defined for the SIS. Indeed, with the exception of the SIS model, we cannot reject any of the mass models we consider here.

We note that our choice of parameter priors for these analyses does not affect the calculation of the Bayes factor, because these priors are set to be flat with allowed ranges that extend beyond the preferred range of each parameter.

Taking guidance from both theoretical expectation and observational results, we select the NFW model as a good description of the mass profile. As mentioned in Introduction, in fact, cosmological simulations produce DM haloes with mass profiles well described by an NFW profile. This result is in good agreement with a variety of observational analyses using both dynamics and weak lensing (Carlberg et al. 1997; van der Marel et al. 2000; Biviano & Girardi 2003; Katgert et al. 2004; Umetsu et al. 2011), even though some results have preferred different models (Merritt et al. 2006; Navarro et al. 2010; Dutton & Macciò 2014; van der Burg et al. 2015b; Sereno & Etori 2017).

4.1.2 Bayesian model averaging with anisotropy models

In contrast to the mass profiles, the literature does not provide us with strong predictions for the radial form of the velocity anisotropy or β profile. In Table 4, we list – for each of the composite clusters as defined in Table 2 – the results of the MAMPOSSt analysis, with the anisotropy parameter θ_β being β_C for the C model, β_∞ for the T and O models, and r_β for the ME and OM models. The errors on each of the parameters listed in the table are obtained by a marginalization procedure, i.e. by integrating the probabilities $p(r_{200}, r_s, \beta)$ provided by MAMPOSSt over the remaining two free parameters. In addition to the anisotropy parameters, Table 4 contains the dynamical constraints on the composite cluster virial radius R_{200}^{dyn} and the Bayesian weight described above.

Because we cannot strongly reject any of the models, we combine the results from the different anisotropy profiles by performing

Table 4. Parameter constraints from the MAMPOSSt analysis of the composite clusters defined in Table 2. Columns represent the velocity anisotropy model $\beta(r)$, the virial radius r_{200} , the scale radius r_s , the anisotropy parameter θ_β , and the Bayes factor from equation (11). For each composite cluster, we also report the value of r_v , which is the scale radius of the galaxy number density profile, obtained from the ratio between the mean value of the $R_{200}^{SZ+\sigma}$ in each redshift bin, and the fixed c_{gal} value (see Section 3.3.3). Note that for the anisotropy models C, T, and O the anisotropy parameter does not have units, while for the M \mathbb{L} and OM the values are evaluated in Mpc.

$\beta(r)$ model	r_{200} (Mpc)	r_s (Mpc)	θ_β	Bayes factor
$\langle z \rangle = 0.33 ; r_v = 0.33$ Mpc				
C	$1.84^{+0.12}_{-0.07}$	$0.34^{+0.54}_{-0.12}$	$0.28^{+0.91}_{-0.24}$	0.39
M \mathbb{L}	$1.79^{+0.14}_{-0.05}$	$0.31^{+0.25}_{-0.08}$	$0.37^{+3.99}_{-0.37}$	0.94
OM	$1.68^{+0.17}_{-0.07}$	$0.19^{+0.13}_{-0.05}$	$1.97^{+5.19}_{-0.77}$	1.00
O	$1.87^{+0.17}_{-0.14}$	$0.24^{+0.15}_{-0.06}$	$-0.56^{+2.38}_{-0.43}$	0.15
T	$1.81^{+0.09}_{-0.09}$	$0.27^{+0.19}_{-0.08}$	$0.19^{+2.08}_{-0.38}$	0.28
$\langle z \rangle = 0.42 ; r_v = 0.32$ Mpc				
C	$1.85^{+0.10}_{-0.07}$	$0.44^{+0.53}_{-0.14}$	$0.20^{+0.90}_{-0.27}$	0.09
M \mathbb{L}	$1.79^{+0.12}_{-0.06}$	$0.40^{+0.17}_{-0.11}$	$0.86^{+8.79}_{-0.56}$	1.00
OM	$1.62^{+0.17}_{-0.05}$	$0.27^{+0.15}_{-0.06}$	$1.19^{+2.97}_{-0.37}$	0.25
O	$1.78^{+0.15}_{-0.10}$	$0.35^{+0.12}_{-0.07}$	$0.39^{+1.66}_{-0.41}$	0.12
T	$1.81^{+0.09}_{-0.09}$	$0.42^{+0.22}_{-0.10}$	$0.39^{+1.59}_{-0.32}$	0.18
$\langle z \rangle = 0.51 ; r_v = 0.29$ Mpc				
C	$1.69^{+0.08}_{-0.05}$	$0.32^{+0.23}_{-0.09}$	$0.19^{+0.77}_{-0.26}$	0.02
M \mathbb{L}	$1.66^{+0.08}_{-0.05}$	$0.33^{+0.13}_{-0.10}$	$0.53^{+5.13}_{-0.30}$	0.32
OM	$1.51^{+0.07}_{-0.04}$	$0.33^{+0.19}_{-0.12}$	$0.76^{+1.35}_{-0.20}$	0.13
O	$1.56^{+0.06}_{-0.05}$	$0.31^{+0.07}_{-0.05}$	$0.93^{+0.21}_{-0.05}$	1.00
T	$1.64^{+0.06}_{-0.06}$	$0.40^{+0.17}_{-0.09}$	$0.68^{+0.88}_{-0.18}$	0.21
$\langle z \rangle = 0.62 ; r_v = 0.27$ Mpc				
C	$1.53^{+0.06}_{-0.05}$	$0.35^{+0.31}_{-0.11}$	$0.39^{+0.64}_{-0.20}$	0.27
M \mathbb{L}	$1.52^{+0.05}_{-0.05}$	$0.39^{+0.13}_{-0.13}$	$0.06^{+0.50}_{-0.05}$	0.15
OM	$1.43^{+0.09}_{-0.04}$	$0.18^{+0.09}_{-0.04}$	$2.40^{+5.74}_{-0.81}$	1.00
O	$1.51^{+0.06}_{-0.06}$	$0.23^{+0.09}_{-0.05}$	$-0.02^{+1.77}_{-0.35}$	0.09
T	$1.51^{+0.05}_{-0.06}$	$0.27^{+0.15}_{-0.07}$	$0.34^{+1.46}_{-0.30}$	0.22
$\langle z \rangle = 0.86 ; r_v = 0.25$ Mpc				
C	$1.30^{+0.06}_{-0.05}$	$0.26^{+0.31}_{-0.09}$	$0.23^{+1.04}_{-0.31}$	0.12
M \mathbb{L}	$1.28^{+0.07}_{-0.04}$	$0.25^{+0.12}_{-0.08}$	$0.58^{+7.20}_{-0.38}$	1.00
OM	$1.19^{+0.09}_{-0.03}$	$0.19^{+0.11}_{-0.05}$	$1.08^{+3.07}_{-0.44}$	0.32
O	$1.25^{+0.06}_{-0.05}$	$0.25^{+0.08}_{-0.05}$	$0.68^{+1.01}_{-0.24}$	0.33
T	$1.28^{+0.05}_{-0.06}$	$0.30^{+0.17}_{-0.08}$	$0.55^{+1.36}_{-0.26}$	0.36

a Bayesian model averaging, weighting every model by its Bayes factor and combining statistics from the different β models. This approach was first proposed by Vázquez et al. (2012a), and has subsequently been used in multiple analyses aside from our own (Vázquez et al. 2012b; Aslanyan et al. 2014; Hee et al. 2016; Planck Collaboration et al. 2016). For a proof that the weighted and co-added probability distribution functions constitute a proper probability distribution function, we direct the reader to examine

Table 5. Parameter constraints for the composite cluster mass profiles. Columns represent the redshift range, virial radius R_{200}^{dyn} , scale radius r_s , dynamical mass M_{200}^{dyn} , and concentration. These results are obtained using Bayesian model averaging of the different anisotropy models.

Bin	Redshift range	R_{200}^{dyn} (Mpc)	r_s (Mpc)	M_{200}^{dyn} ($10^{14} M_\odot$)	c_{200}
1	0.26–0.38	$1.81^{+0.11}_{-0.10}$	$0.26^{+0.16}_{-0.10}$	$9.44^{+1.70}_{-1.65}$	$5.4^{+2.6}_{-2.1}$
2	0.39–0.44	$1.82^{+0.10}_{-0.09}$	$0.36^{+0.17}_{-0.11}$	$10.57^{+1.93}_{-1.55}$	$4.1^{+1.5}_{-1.3}$
3	0.46–0.56	$1.56^{+0.11}_{-0.06}$	$0.31^{+0.15}_{-0.13}$	$7.42^{+1.58}_{-0.92}$	$4.4^{+1.7}_{-1.8}$
4	0.56–0.71	$1.51^{+0.05}_{-0.06}$	$0.25^{+0.11}_{-0.08}$	$7.31^{+1.13}_{-0.62}$	$5.2^{+1.9}_{-1.9}$
5	0.71–1.32	$1.28^{+0.06}_{-0.06}$	$0.24^{+0.11}_{-0.10}$	$6.20^{+0.85}_{-0.88}$	$4.2^{+2.0}_{-1.6}$
-	0.26–1.32	$1.62^{+0.03}_{-0.05}$	$0.29^{+0.06}_{-0.07}$	$8.71^{+0.52}_{-0.80}$	$5.1^{+1.1}_{-1.0}$

section 8.2, equation (69) of the last reference. The results of this analysis are shown in Table 5 where we present the virial radius R_{200}^{dyn} , the NFW scale radius r_s , the virial mass M_{200}^{dyn} , and the concentration c_{200} . The 1σ parameter uncertainties are computed through a marginalization procedure, as before.

4.1.3 Impact of mass priors on composite clusters

As mentioned in Section 3.3.1, when creating the composite clusters, we perform a rescaling of the observables (the galaxy-projected cluster distances R and rest-frame velocities v_{rf}) using two different initial mass estimates, $M_{200}^{SZ+\sigma}$ and $M_{200}^{SZ+\text{Planck}}$ (available for each individual cluster). The dynamical mass constraints that result from the composite clusters created using these two different scalings do not differ significantly, demonstrating the stability of our analysis to underlying systematic mass uncertainties in the initial mass estimates used to construct the composite clusters. We return to this point in Section 4.4.1 below, where we show the dynamical masses obtained in the two cases (see Table 6).

Moreover, the masses derived through the full dynamical analysis are in good agreement with the mean SPT-based masses $M_{200}^{SZ+\sigma}$ listed in Table 2. Regarding the precision of the constraints, it is clear that a composite cluster with ~ 600 cluster members allows one to determine the dynamical mass with what is effectively a ~ 15 per cent uncertainty (~ 8 per cent for the full data set, using ~ 3000 tracers). In contrast, the NFW scale radius and the corresponding concentration are only weakly constrained.

4.1.4 Goodness of fit of dynamical models

We have also examined whether the best-fitting models are an adequate description of the data. To do this, we have created 1000 mock galaxy samples of similar size to the full galaxy sample by sampling the likelihood distribution in LOS velocity and projected distance produced using the best-fitting model to the full galaxy sample. The best-fitting model has an M \mathbb{L} anisotropy profile, and the Bayes factors of the other four models for the full galaxy sample are similar to those seen in Table 4 for the redshift subsamples of the data, indicating that all five anisotropy models fit similarly well. We then analyse each of these 1000 mock samples in the same way as the real data, examining the mean $\log(\text{likelihood})$ per galaxy for the real data and the mocks. The likelihood of the real data is somewhat lower than the typical likelihood of the mocks. However, 6.1 per cent of the mock samples have even lower likelihood than the real data. Thus, the dynamical models we are fitting are indeed an adequate description of the data.

Table 6. Comparisons of dynamical masses from composite clusters calculated using different initial masses. From left to right, the columns contain the redshift range of the cluster sample, the derived dynamical mass given an initial SPT plus velocity dispersion mass, the derived dynamical mass given an SPT + Planck initial mass, and the ratios of these dynamical to initial masses in each case. Finally, we report the constraints on η and η' as described in Section 4.4.2.

Bin	Redshift range	$M_{200}^{\text{dyn}} M_{200}^{\text{SZ}+\sigma}$ ($10^{14} M_{\odot}$)	$M_{200}^{\text{dyn}} M_{200}^{\text{SZ}+\text{Planck}}$ ($10^{14} M_{\odot}$)	$\frac{M_{200}^{\text{dyn}} M_{200}^{\text{SZ}+\sigma}}{\langle M_{200}^{\text{SZ}+\sigma} \rangle}$	$\frac{M_{200}^{\text{dyn}} M_{200}^{\text{SZ}+\text{Planck}}}{\langle M_{200}^{\text{SZ}+\text{Planck}} \rangle}$	η	η'
1	0.26–0.38	$9.44^{+1.70}_{-1.65}$	$9.29^{+1.96}_{-1.45}$	$1.02^{+0.18}_{-0.18}$	$0.86^{+0.15}_{-0.15}$	$0.77^{+0.28}_{-0.13}$	$1.21^{+0.17}_{-0.36}$
2	0.39–0.44	$10.57^{+1.93}_{-1.55}$	$10.41^{+1.88}_{-1.19}$	$1.11^{+0.20}_{-0.16}$	$0.91^{+0.17}_{-0.13}$	$0.92^{+0.17}_{-0.20}$	$1.24^{+0.21}_{-0.27}$
3	0.46–0.56	$7.42^{+1.58}_{-0.92}$	$7.14^{+1.49}_{-0.88}$	$0.93^{+0.20}_{-0.11}$	$0.74^{+0.16}_{-0.09}$	$0.74^{+0.15}_{-0.09}$	$0.97^{+0.23}_{-0.12}$
4	0.56–0.71	$7.31^{+1.13}_{-0.62}$	$7.36^{+1.01}_{-0.78}$	$1.00^{+0.15}_{-0.08}$	$0.73^{+0.11}_{-0.06}$	$0.79^{+0.10}_{-0.10}$	$1.10^{+0.13}_{-0.13}$
5	0.71–1.32	$6.20^{+0.85}_{-0.88}$	$5.95^{+0.83}_{-0.75}$	$0.82^{+0.11}_{-0.12}$	$0.55^{+0.07}_{-0.08}$	$0.63^{+0.13}_{-0.08}$	$0.90^{+0.18}_{-0.11}$
-	0.26–1.32	$8.71^{+0.52}_{-0.80}$	$8.50^{+0.59}_{-0.67}$	$1.05^{+0.06}_{-0.10}$	$0.81^{+0.05}_{-0.07}$	$0.81^{+0.06}_{-0.06}$	$1.14^{+0.06}_{-0.07}$

4.2 Velocity dispersion anisotropy profiles

In the following, we first present the results of our examination of mass and redshift trends (Section 4.2.1) and then discuss implications for a particular mass accretion model (Section 4.2.2) and then finally compare with previous studies of the velocity dispersion anisotropy profile (Section 4.2.3).

4.2.1 Constraints on redshift and mass trends

Fig. 4 contains the measured anisotropy $\beta(r)$ profiles and their 1σ confidence regions for the five composite clusters in different redshift ranges together with the results from the full sample (lower, rightmost panel). These profiles are obtained by using the posterior distribution in the anisotropy parameter θ_{β} extracted from each of the five anisotropy models. Specifically, for each of the models, a large number of θ_{β} values are drawn, consistent with the posterior. Each value corresponds to an anisotropy profile. The number of θ_{β} values drawn for each model is weighted according to the Bayes factor. The sum of all these anisotropy profiles provides a measure of the probability distribution in the anisotropy profile value at each radius. The red line represents the median value of this distribution, while the shaded region is defined by the 16th and 84th percentiles of the distribution (1σ confidence region). As noted in the previous section, the mass profile model in all cases is an NFW with concentration and mass free to vary.

Our analysis indicates that the orbits of passive, red galaxies are nearly isotropic close to the cluster centre, and become increasingly radial going towards larger radii, reaching a radial anisotropy $\beta \simeq 0.15$ – 0.6 at $R/r_{200} \simeq 1$. There is no clear evidence for a redshift trend in the anisotropy profile of the passive galaxy population out to $z \approx 1$. We have carried out a similar analysis of five composite clusters built from the same cluster sample divided into mass bins rather than redshift bins, and we find no evidence for trends with mass, either. We show this result in Fig. 5.

For this reason, we analyse also the full sample, providing our best available constraints. The orbital anisotropy varies from values consistent with zero in the cluster core to a value 0.4 ± 0.15 at the virial radius. For reference, anisotropy values of 0.4 correspond to tangential components of the velocity dispersion ellipsoid having amplitudes that are only 60 per cent as large as the radial component. For the full sample we show (blue dashed line) also the anisotropy profile recovered using the best-fitting NFW parameters and using the Jeans equation to solve for the velocity dispersion anisotropy

profile (Binney & Mamon 1982; Solanes & Salvador-Sole 1990; Biviano et al. 2013). This result is in good agreement with the solution recovered using the Bayesian model averaging over the five adopted anisotropy profiles.

4.2.2 Comparison with two-phase accretion model

The behaviour of the anisotropy profile is consistent with the theoretical model discussed in Lapi & Cavaliere (2009), according to which the growth of structure proceeds in two phases: an early, fast accretion phase during which the cluster undergoes major merging events, and a second slower accretion phase involving minor mergers and smooth accretion (see e.g. White 1986; Zhao et al. 2003; Diemand, Kuhlen & Madau 2007). Lynden-Bell (1967) discusses how a dynamical system rapidly relaxes from a chaotic initial state to a quasi-equilibrium. The first stage of fast accretion provokes rapid changes in the cluster gravitational potential, inducing the collisionless components of the cluster to undergo violent relaxation, resulting in orbits that are more isotropic. Galaxies accreted by the cluster during the fast accretion phase would then be expected to exhibit approximately isotropic orbits, while galaxies accreted during the second phase would maintain their preferentially radial orbitals over longer time-scales. Given that as the cluster accretes its mass and virial radius also grow, a two-phase scenario like this would tend to lead to anisotropy profiles that are isotropic in the core and become more radial at larger radius.

However, one might also expect to see a time or redshift variation of the anisotropy profile, with typical galaxy orbits in high-redshift clusters showing less of a tendency for radial orbits near the virial radius. The fact that our analysis shows no strong redshift trend in the anisotropy profiles is an indication that, in massive galaxy clusters, the passive galaxy population orbits are not changing significantly with cosmic time since $z \approx 1$. This suggests that the merging and relaxation processes responsible for the anisotropy profiles are underway at all cosmic epochs probed here. Other indicators of cluster merging have shown similar results; namely, the fraction of systems with disturbed X-ray morphologies (typically measured using centroid variations or ellipticities; Mohr, Fabricant & Geller 1993) does not change significantly with redshift in samples of homogeneously SZE-selected cluster samples (Nurgaliev et al. 2017; McDonald et al. 2017). This may suggest that the fast accretion phase of such a two-phase model could be very short and happening primarily at redshifts above those probed by our sample. Clearly, a detailed examination of cosmological N -body simulations with sufficient

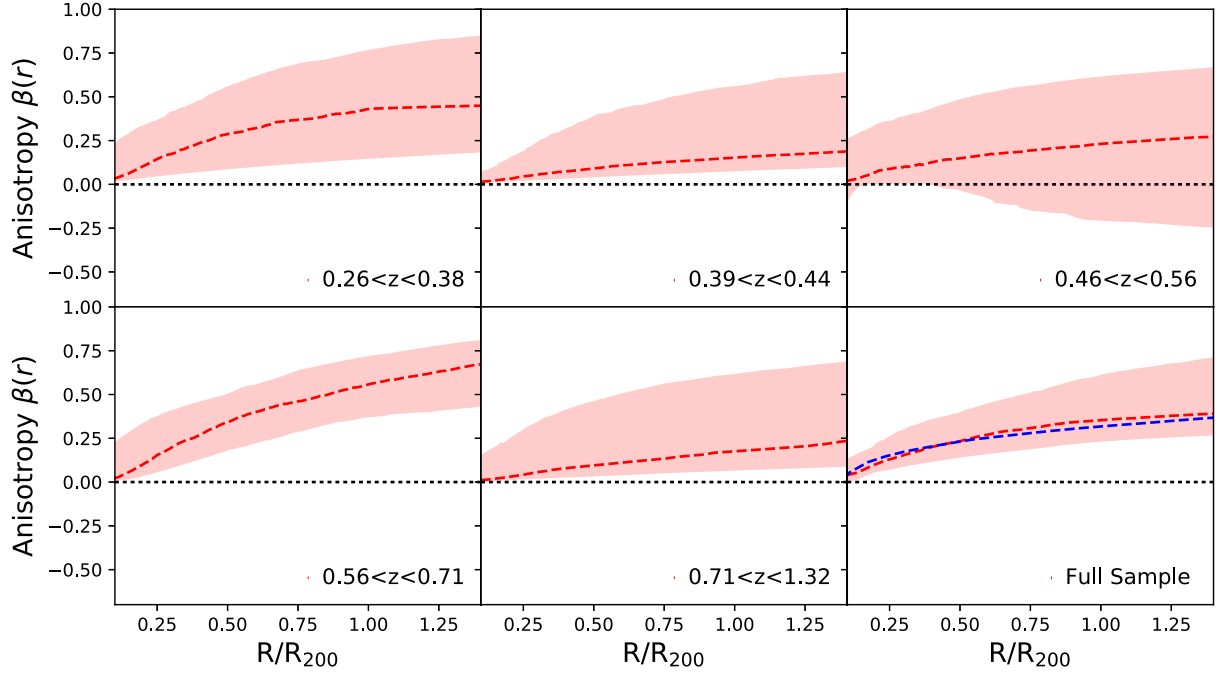


Figure 4. Velocity anisotropy profile $\beta(r)$ for each redshift bin. The red dashed line represents the profile obtained implementing a Bayesian model averaging, with the pink shaded region indicating the 1σ confidence region around this solution. There is no clear evidence of a redshift trend. The blue line in the lower right panel (full sample) shows the result obtained when adopting the best-fitting NFW model and the Jeans equation inversion to solve for the anisotropy profile. This result is in good agreement with the model averaging result. Our analysis shows that passive galaxies preferentially move on nearly isotropic orbits close to the cluster centre, and on increasingly radial orbits as one moves to the virial radius.

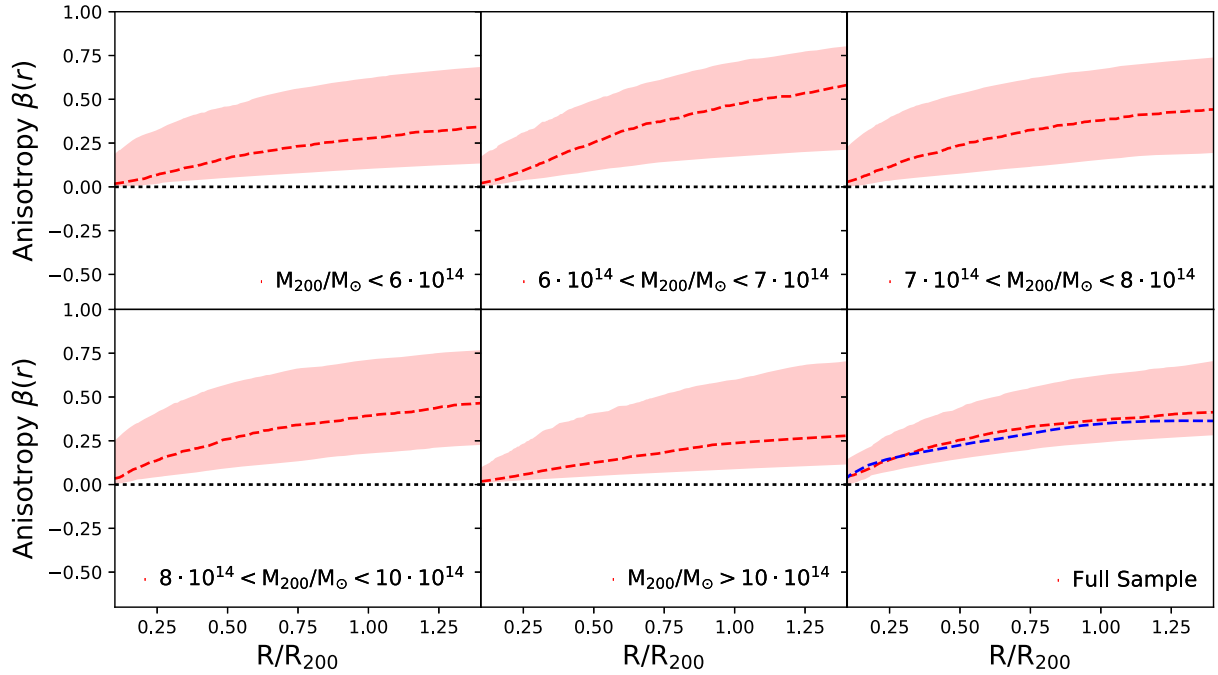


Figure 5. Velocity anisotropy profile $\beta(r)$ for each mass bin. As in Fig. 4, the anisotropy-redshifted dashed line represents the profile obtained implementing a Bayesian model averaging, with the pink shaded region indicating the 1σ confidence region around this solution. The blue line in the lower right panel (full sample) shows the result obtained when adopting the best-fitting NFW model and the Jeans equation inversion to solve for the anisotropy profile. There is no clear evidence of a mass trend.

volume to contain rare, massive clusters and sufficient mass resolution to ensure the survival of galaxy-scale subhaloes after accretion into the cluster is warranted. Moreover, a dynamical analysis like the one we have carried out that focuses on systems at redshift $z > 1$ would enable a more sensitive probe for time variation in the growth of structure.

4.2.3 Comparison with previous results

Previous studies of the velocity dispersion anisotropy conducted on passive cluster members at intermediate to high redshifts present hints of an anisotropy profile that is nearly isotropic close to the cluster centre, and increasingly radial at larger radii. Biviano & Poggianti (2009), stacking 19 clusters between redshift ≈ 0.4 – 0.8 with a mean mass of $\approx 3 \times 10^{14} M_{\odot}$, suggest radially anisotropic orbits. Biviano et al. (2016) reach the same conclusion when analysing a stacked sample of 10 clusters at $0.87 < z < 1.34$. Studying a single cluster, Annunziatella et al. (2016) show that the same trend is found for galaxies characterized by a stellar mass $M_{\star} > 10^{10} M_{\odot}$, while lower mass galaxies move on more tangential orbits, avoiding small pericentres, presumably because those that cross the cluster centre are more likely to have been tidally destroyed. These results are consistent with ours. Our result, obtained through the analysis of a large sample of passive galaxies within a homogeneously selected sample of massive clusters over a wide redshift range and with low scatter mass estimates, allows us to cleanly probe for redshift and mass trends in the velocity dispersion anisotropy profile.

Some published analyses carried out at lower redshifts ($z \lesssim 0.1$) than our sample show similar results. Wojtak & Łokas (2010) analysed a sample of 41 nearby relaxed clusters, finding that galaxy orbits are isotropic at the cluster centres and more radial at the cluster virial radius. A similar result is obtained by Lemze et al. (2009) and Aguerra et al. (2017). However, other analyses show that the orbits of passive galaxies at these redshifts are more isotropic at all radii (Biviano & Katgert 2004; Katgert et al. 2004; Biviano & Poggianti 2009; Munari, Biviano & Mamon 2014), hinting at a possible change in galaxy orbits over time due to processes such as violent relaxation, dynamical friction, and radial orbital instability (Bellovary et al. 2008). At present, results from numerical simulations predict a range of behaviour (Wetzel 2011; Iannuzzi & Dolag 2012; Munari et al. 2013), so further study is definitely needed. Extending our own observational analysis towards lower redshifts could also help clarify this picture.

4.3 Pseudo phase-space density profiles

The determination of the anisotropy profile $\beta(r)$ allows us to investigate the behaviour of the PPSD profile $Q(r)$ introduced in Introduction. According to numerical simulations of virialized haloes (Taylor & Navarro 2001; Dehnen & McLaughlin 2005; Lapi & Cavaliere 2009), there is a scaling between the density ρ and the velocity dispersion, best appreciated by considering the quantity $Q(r) = \rho/\sigma^3$. The profile of this quantity is found to follow a universal power law of fixed slope, $\propto r^{-1.875}$. Remarkably, this is the same power law predicted by the similarity solution of Bertschinger (1985) for secondary infall and accretion on to an initially overdense perturbation in an Einstein-de Sitter universe. In that work, the authors found that the relaxation process is self-similar, meaning that each new shell falling in is virialized and adds a constant contribution to the resulting power-law density profile.

Because the velocity dispersion profile is constrained by the galaxies, we derive the PPSD profile using the number density

profile of galaxies ρ_{gal} instead of the mass density profile. For each stacked cluster, we fix the halo occupation number N_{200} of red galaxies to that found by H17 to find the central density ρ_0 of the NFW ρ_{gal} profile, such that

$$\rho_0 = \frac{N_{200}}{4\pi \int_0^{R_{200}} \frac{r^2}{r/r_s(1+r/r_s)^2} dr}. \quad (12)$$

We present the results obtained investigating both the total PPSD profile $Q_{\text{gal}}(r) \equiv \rho_{\text{gal}}/\sigma^3$ and the radial PPSD profile $Q_{\text{gal},r}(r) \equiv \rho_{\text{gal}}/\sigma_r^3$, where ρ_{gal} is obtained as described above and $\sigma_r(r)$ is recovered using the following equation (van der Marel 1994; Mamon & Łokas 2005; Mamon et al. 2013)

$$\sigma_r^2(r) = \frac{1}{v(r)} \int_r^{\infty} \exp \left[2 \int_r^s \beta(t) \frac{dt}{t} \right] v(s) \frac{GM(s)}{s^2} ds, \quad (13)$$

evaluated over an adequate grid of r , and $\sigma(r)$ is given by

$$\sigma(r) = \sqrt{3 - 2 \beta(r)} \sigma_r(r). \quad (14)$$

Fig. 6 shows the ratio between the derived galaxy PPSD profiles and the fixed-slope best-fitting relations $Q(r) \propto r^{-1.84}$ and $Q_r(r) \propto r^{-1.92}$, where the slopes are those obtained by Dehnen & McLaughlin (2005) by studying DM particles in numerically simulated haloes, while the normalizations are fitted to the data of the full sample. A similar value for the slope of $Q(r)$ is found by Faltenbacher et al. (2007), a work based on numerical simulations and focused on the gas and DM entropy profiles in galaxy clusters. The similarity between radial profiles of the passive galaxy PPSD and simulated PPSD profiles is an indication that the passive galaxy radial distribution and kinematics are similar to those of the DM particles in those simulations.

We emphasize here that it is not possible for us to predict the PPSD profile for the underlying DM in our sample, because in general, the blue or EL galaxy population exhibits different kinematical properties than those of the passive, red or nEL galaxies. One might expect the full DM kinematics to exhibit a mix of the properties of the different galaxy populations. We note that because the measured concentration of the passive population (H17) is similar to the derived concentration of the mass profiles from our analysis, that if we replace the galaxy density profile with our derived mass density profile, the PPSD profiles have a very similar radial behaviour.

In Fig. 6, we also show (solid black lines) the expected amplitude if there is self-similar behaviour of passive galaxy profiles and kinematics. That is, we have $Q(r) \equiv \rho/\sigma^3 \propto E(z)^2/[M(r)E(z)] = E(z)/M(r)$, where $H(z) = H_0 E(z)$ is the Hubble parameter. The passive galaxy PPSD amplitudes are in reasonably good agreement with the amplitudes expected under self-similarity, providing some additional evidence that this population is approximately self-similar.

In the previous subsection, we mentioned that the anisotropy profiles could be the result of violent relaxation. This process, driven by gravity alone, would also tend to create a scale-invariant phase-space density. Relaxation into dynamical equilibrium would then lead the slope of the PPSD profile to approach a critical value, resulting in the particular form of the density profile for DM particles within simulated haloes. An anisotropy profile isotropic in the inner regions and increasingly radial at larger radii gives the PPSD profile slope observed in numerical simulations. The agreement then suggests that the passive galaxies we analyse here have reached a similar level of dynamical equilibrium as the DM particles in the simulations, and that this is true for all redshifts up to $z \approx 1$.

It is difficult to compare our results directly to previous studies, because most of those studies have tended to focus on $Q(r)$ and $Q_r(r)$

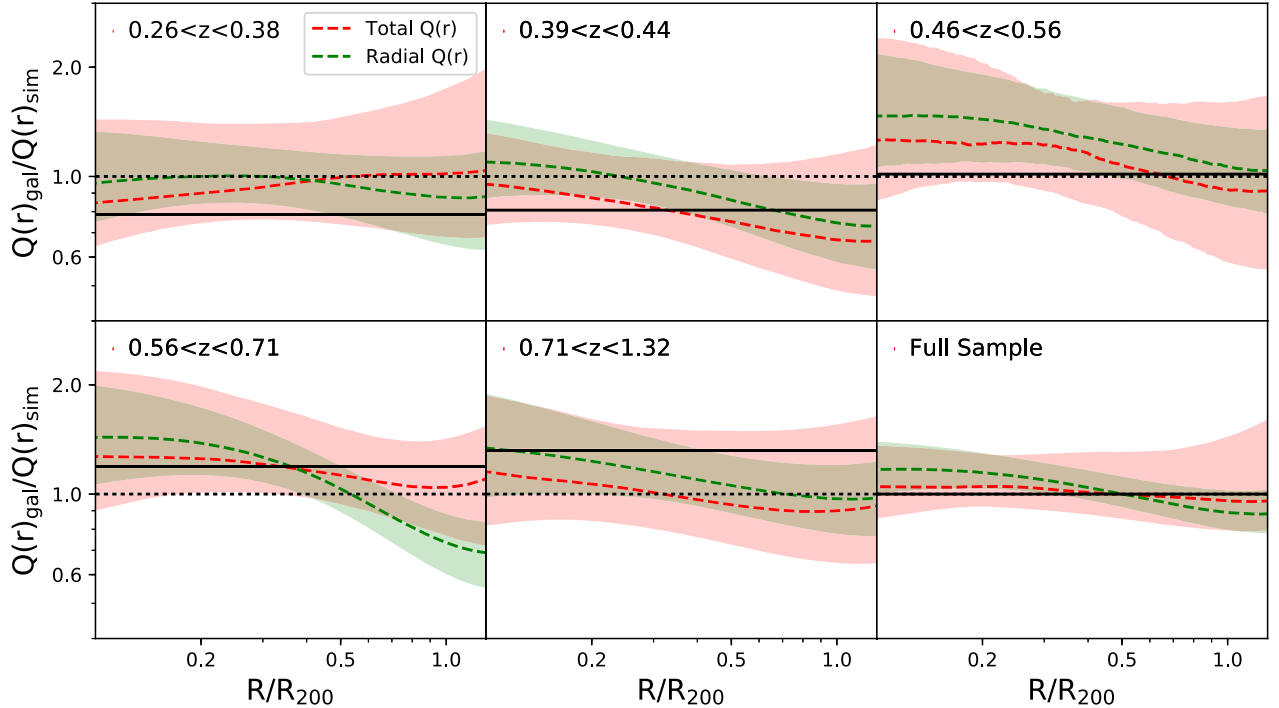


Figure 6. The ratios of the passive galaxy PPSD profiles $Q_{\text{gal}}(r)$ to that of DM particles in simulations $Q_{\text{sim}}(r)$ (Dehnen & McLaughlin 2005) are shown in green for the total PPSD $\rho_{\text{gal}}/\sigma^3$ and in red for the radial PPSD $\rho_{\text{gal}}/\sigma_r^3$ for five redshift bins and the full sample. Shading corresponds to the 1σ confidence region. The normalization of each $Q(r)$ profile (both of the galaxies and of the simulations) is fixed to that of the $Q_{\text{gal}}(r)$ of the full sample, which is determined by fitting to the data. Thus, the deviations from $Q_{\text{gal}}(r)/Q_{\text{sim}}(r) = 1$ of the full sample are only driven by the deviations of the $Q_{\text{gal}}(r)$ slope from the $Q_{\text{sim}}(r)$ one, but in the other samples, they are also driven by real normalization differences. The solid black line marks the expected amplitude in each redshift bin under the assumption of cluster self-similar redshift evolution. Both passive galaxy PPSD profiles are in reasonably good agreement with N -body simulations, and the data are roughly consistent with self-similar scaling with mass and redshift.

using the inferred total matter density from their dynamical analyses rather than the galaxy density profile. Such an approach produces a mixed PPSD profile that contains both total matter and galaxy properties, because the dispersion profiles used are necessarily those of the galaxies rather than the DM. This complicates the interpretation, because a mismatch between the observed and simulated PPSD profiles could be because the total matter density profile does not match the simulated DM profile or it could mean that the velocity dispersion profile of the DM and a particular galaxy population do not match. In general, given that different galaxy populations tend to exhibit different kinematic properties (i.e. radial distributions and velocity dispersion and anisotropy profiles), one would not expect the mixed PPSD profiles defined in this way to agree for the different populations.

Indeed, Biviano et al. (2013) measured $Q(r)$ using the anisotropy profile of the galaxies and the NFW mass density profile inferred from their analysis, separately for SF and passive galaxies in a single cluster. They found that their observed PPSD profile agreed with simulations only for the passive galaxy population. They argue that passive members have undergone violent relaxation and have reached dynamical equilibrium, while SF members have not yet reached equilibrium. But another possible interpretation would be that the Jeans analyses of both populations are equally accurate, but that the passive population exhibits a velocity dispersion and anisotropy profile more similar to that of the simulated DM. A more recent study came to similar conclusions (Munari, Biviano & Mamon 2015).

On the other hand, a recent study of the nearby cluster Abell 85 (Aguerri et al. 2017) shows that $Q_r(r)$ follows the theoretical power-law form independent of the galaxy colour or luminosity, concluding that all the different families of galaxies under study reached a virialized state. Given the discussion above, this agreement in the mixed PPSD profile derived from different galaxy populations also indicates that the different populations must have similar kinematics (i.e. velocity dispersion and anisotropy profiles). In their study, they emphasize that the anisotropy profiles of the blue and red galaxies are different, which would make the agreement in $Q_r(r)$ surprising. However, fig. 3 in their paper suggests that the anisotropy profiles have similar character (isotropic in the centre, more radial at larger radius) and present evidence for inconsistency that is weak ($\leq 2\sigma$).

4.4 Dynamical mass constraints

We use the dynamical information to study the masses of these clusters in two different ways. In Section 4.4.1, we analyse the composite clusters and the consistency of the dynamical masses when using different initial masses to scale the galaxy observables. In Section 4.4.2, we examine the differences between the dynamical masses and the SPT + Planck masses $M_{200}^{\text{SZ+Planck}}$ using a different approach where the scaling values for each cluster r_{200} and v_{200} are altered self-consistently in each iteration of the Markov Chain to reflect the SPT + Planck masses scaled by a free parameter η , defined in equation (15).

4.4.1 Mass constraints on the composite clusters

Constructing the composite clusters (described fully in Section 3.3) requires a scaling of our galaxy observables, v_{rf} and R , by estimates of the virial radius and velocity r_{200} and v_{200} respectively, and therefore requires an initial mass estimate. As described above, this is a potential problem, because there is currently a ~ 25 per cent shift between the SPT cluster masses derived using cluster counts and velocity dispersion information, and the masses derived using the cluster counts and external cosmological priors from *Planck* (Bocquet et al. 2015; de Haan et al. 2016). Thus, we examine the dynamical mass constraints in each case: (1) those derived using initial masses derived from the cluster counts and velocity dispersion measurements $P(M_{200}^{\text{dyn}} | M_{200}^{\text{SZ}+\sigma})$ and (2) those derived using initial masses from the cluster counts and external cosmological constraints from *Planck* $P(M_{200}^{\text{dyn}} | M_{200}^{\text{SZ}+\text{Planck}})$.

In Fig. 7, we display the marginalized distribution of the dynamical masses obtained with our analysis where the $M_{200}^{\text{SZ}+\sigma}$ masses were adopted for the initial scaling. The green regions mark the 1σ confidence regions, and the red lines represent the mean initial masses derived from the cluster counts and velocity dispersion measurements $\langle M_{200}^{\text{SZ}+\sigma} \rangle$, where the masses from each cluster are weighted by the number of galaxy velocities available for that cluster. There is good agreement between the dynamical and the initial masses in all redshift bins and also for the full sample (lower, rightmost panel). The second column of Table 6 contains the measurement results and uncertainties for each subset. Characteristic dynamical mass uncertainties are at the ~ 15 per cent level for individual subsamples and at the ~ 8 per cent level for the full sample. These are quite encouraging mass constraints, given that they are marginalized over the velocity dispersion anisotropy profile uncertainties.

We find that by fixing the concentration and radial anisotropy parameters to their best-fitting values when fitting for mass in a composite cluster with 600 tracers, the resulting mass uncertainty is not significantly impacted. This suggests that our uncertainties in the individual redshift bins are not dominated by the freedom in mass and anisotropy profiles. A similar test on the composite cluster built from the full dynamical sample leads to a ~ 5 per cent mass uncertainty, which is comparable to what we find when using a single anisotropy model, before performing the Bayesian model averaging. This is an interesting result when taken together with the discussion of velocity dispersion based mass estimates in Sifón et al. (2016), where the scaling presented suggests that with samples of 600 dynamical tracers mass estimates should be closer to ~ 7 per cent accurate rather than the 15 per cent we recover. Further examination of the assumptions built into the dynamical mass measurements using velocity dispersions and full Jean analysis modelling is warranted and is planned for a future analysis.

As discussed in Section 2.1, current estimates from studies of clusters in numerical simulations indicate there are remaining systematic uncertainties associated with MAMPOSSt analysis at the 10 per cent level.

To test the stability of the recovered dynamical masses to the initial input masses used for scaling, we perform the same analysis using the SPT + Planck cluster masses. The third column of Table 6 shows these results. These dynamical masses with the different initial masses are quite close to the values derived with the other set of initial masses. This shows that there is no strong dependence of the dynamical mass on the initial mass. This is because any change in the masses used for rescaling the cluster observables during stacking will impact, on average, the individual cluster masses and the final mean mass in the bin in a similar manner. The overall scale

of the dynamical data in projected radius and LOS velocity remains approximately invariant.

As columns 4 and 5 of Table 6 make clear, the dynamical masses, while being in good agreement with the cluster counts plus velocity dispersion masses $M_{200}^{\text{SZ}+\sigma}$, exhibit some discordance with the cluster counts plus external cosmological constraint masses $M_{200}^{\text{SZ}+\text{Planck}}$. While the three lowest redshift bins show no significant disagreement, the upper two redshift bins show masses that are only 73 per cent and 55 per cent as large as the SPT+Planck masses (offsets that are statistically significant at the 2.5σ and 6.5σ levels, respectively). The full sample has a dynamical mass that is only 80 per cent of the SPT + Planck masses, a difference that is significant at the 3.8σ level (statistical only). The direction and scale of this mass shift is similar to that highlighted already in Bocquet et al. (2015). However, with our analysis, we are able to show that this discrepancy seems to grow with redshift.

4.4.2 Comparison with SZE-based masses

To examine this discrepancy more carefully, we use the dynamical analysis to examine the masses of these clusters and compare them to the masses derived separately from the SPT cluster counts in combination with external cosmological constraints from the *Planck* CMB anisotropy. Rather than using the composite clusters, we analyse individual clusters, combining the likelihoods from each cluster and exploring constraints on an overall mass scaling parameter η , that is defined as

$$\eta = \frac{M_{200}^{\text{dyn}}}{M_{200}^{\text{SZ}+\text{Planck}}}. \quad (15)$$

We do this by running MAMPOSSt for each individual cluster in our sample. We calculate the posterior distribution of η by using a multimodal nested sampling algorithm, namely MultiNest (Feroz & Hobson 2008; Feroz, Hobson & Bridges 2009; Feroz et al. 2013), which provides us with the evidence for each model, and allows us to perform a Bayesian model averaging over different subsets of clusters.

Fig. 8 contains a plot of the posterior distributions of η from our analysis within each redshift bin and for the full sample. Results are largely consistent with the results from the composite clusters. Column 6 of Table 6 contains the best-fitting η values and associated uncertainties. The preferred value for the full sample is $\eta = 0.81 \pm 0.06 \pm 0.08$. The constraint is followed by a statistical uncertainty and then a systematic uncertainty.

As already discussed in Section 2.1, studies of dynamical tracers drawn from clusters in structure formation simulations indicate that MAMPOSSt-derived dynamical masses has systematic uncertainties of ≈ 10 per cent (see Mamon et al. 2013). This number comes from an analysis of tracers lying within a sphere of radius r_{100} that are then used to estimate the virial radius r_{200} . For the systematic uncertainty presented above, we have therefore adopted as a Gaussian with $\sigma = 10$ per cent centres at no bias.

If one combines the statistical and systematic uncertainty in quadrature, the implication would be a difference at the 1.9σ level. As mentioned already, this tendency for the dynamical masses to be lower than those masses derived from the cluster counts in combination with external cosmological constraints is consistent with the tendencies seen previously (Bocquet et al. 2015) using simply dispersions and the cluster mass function in (see also Rines et al. 2016; Sifón et al. 2016). More recent weak-lensing analyses also

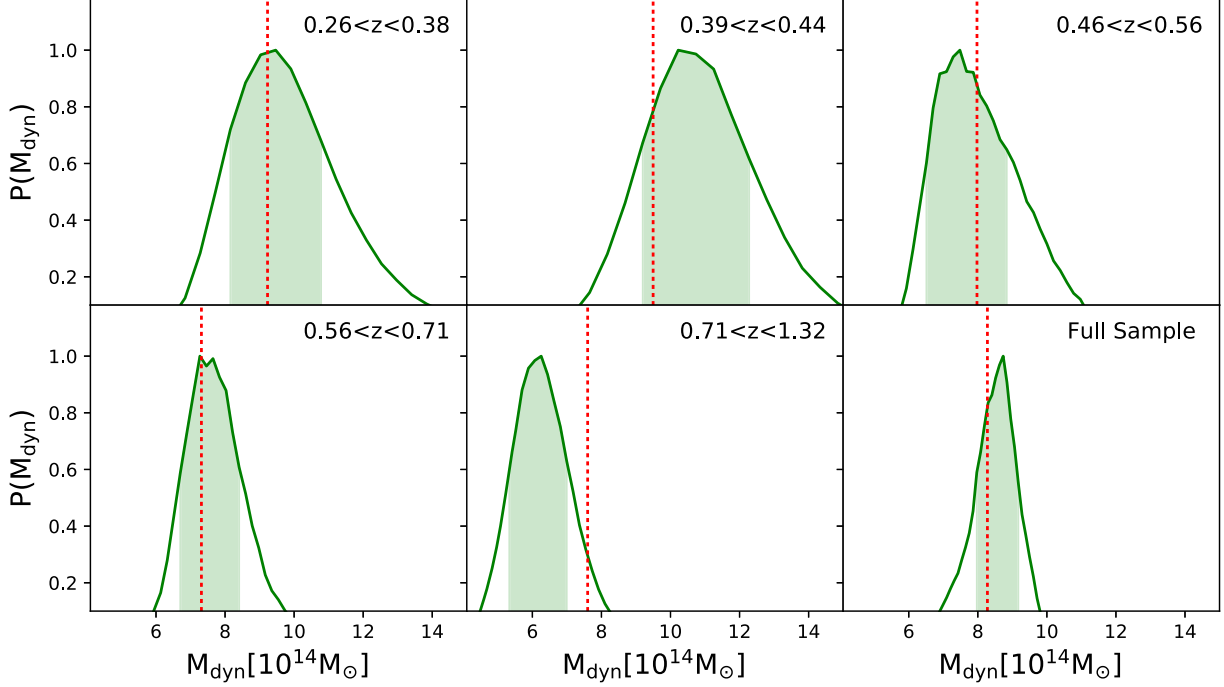


Figure 7. Marginalized distribution of the dynamical masses. Each panel corresponds to a different redshift range, and the final panel shows the results of the analysis of the full sample. In green, we highlight the 1σ confidence region. The red line represents the mean SPT + σ mass $\langle M_{200}^{SZ+\sigma} \rangle$ for the clusters in the bin, weighted by the number of member galaxies in the individual clusters. There is a good agreement between the dynamical masses and the originally inferred SPT masses in all cases (see Table 6).

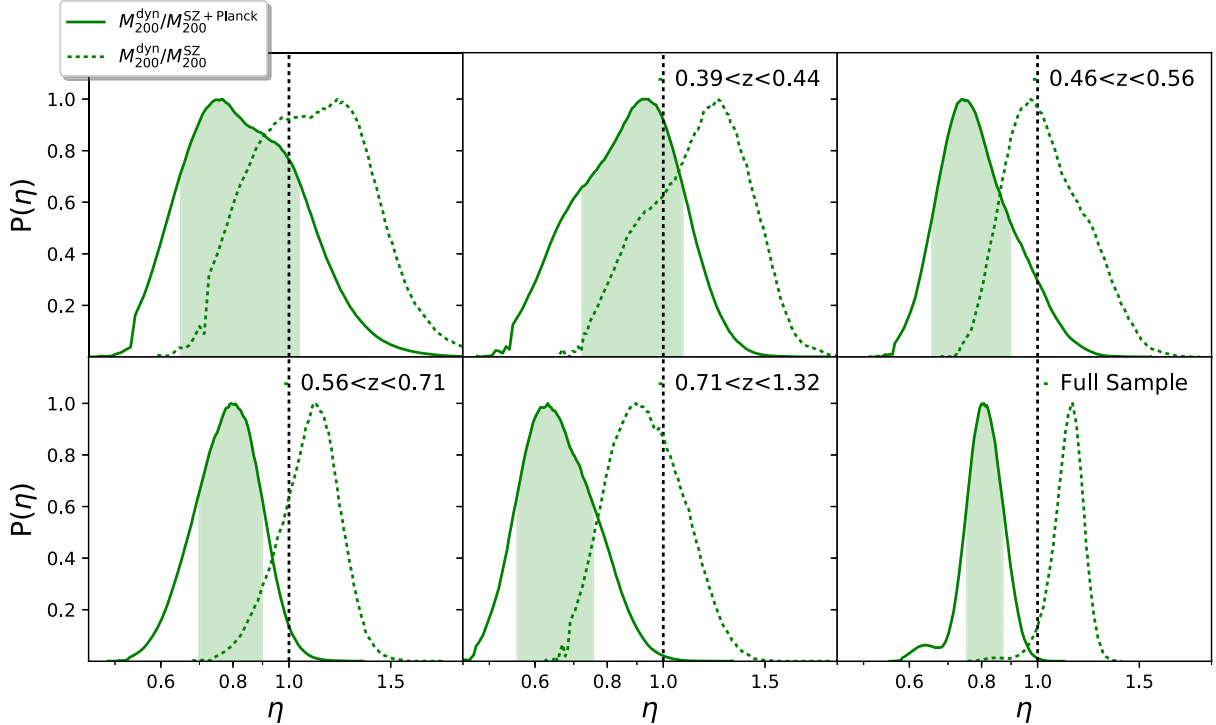


Figure 8. Posterior distribution of $\eta = M_{200}^{\text{dyn}} / M_{200}^{\text{SZ+Planck}}$ (solid green lines) arising from a dynamical analysis of each cluster subsample. The green region shows the 1σ region, while the vertical dotted black line marks the value $\eta = 1$. These distributions show that in the two high-redshift bins and for the full sample there is disagreement between the dynamical masses and the SPT + Planck calibrated cluster masses. For the full sample, the discrepancy is 1.9σ significant when including estimates of the systematic uncertainties. In the highest redshift bin, the discrepancy is 2.6σ . The dashed green lines show the estimated posterior for $\eta' = M_{200}^{\text{dyn}} / M_{200}^{\text{SZ}}$, where M_{200}^{SZ} represents the de Haan et al. (2016) masses calibrated using the SPT mass function and Y_X measurements for many of the systems. In contrast to the SPT + Planck masses, these masses are in good agreement (1σ offset) with the dynamical masses.

support the lower mass scale of SPT clusters (Dietrich et al. 2017; Stern et al. 2018).

To emphasize, Fig. 8 also shows (dotted line) the distribution of

$$\eta' = \frac{M_{200}^{\text{dyn}}}{M_{200}^{\text{SZ}}}, \quad (16)$$

where M_{200}^{SZ} are masses calibrated from a cosmological analysis carried out in de Haan et al. (2016) using the X-ray mass proxy Y_x and the abundance of clusters as a function of redshift, Y_x and the SZE mass proxy ξ , without the inclusion of the external *Planck* cosmological constraints. These results indicate that the dynamical masses are in good agreement with the M_{200}^{SZ} masses at all redshifts and for the full sample. The η' distribution for the full sample prefers a value of $\eta' = 1.14 \pm 0.07 \pm 0.11$, indicating no disagreement.

To summarize, our dynamical mass measurements, which are derived using only dynamical information and no information from the mass function or cluster counts, are in good agreement with masses derived using information from the cluster counts together with additional information from either velocity dispersions or from X-ray Y_x measurements that have been externally calibrated. However, our mass measurements exhibit moderate disagreement with those masses obtained similarly but when also adopting external cosmological priors from *Planck* CMB anisotropy. Progress in testing these two mass scales would require better control of the systematic uncertainties in the dynamical masses (Mamon et al. 2013). The agreement between the dynamical and the *Planck*-based masses is best at low redshift, with the dynamical masses preferring ever smaller η with increasing redshift. In the highest redshift bin ($0.71 \leq z \leq 1.32$), we measure $\eta = 0.63_{-0.08}^{+0.13} \pm 0.06$, discrepancy at the 3σ level.

4.5 Impact of disturbed clusters

One assumption in applying the Jeans equation to analyse our sample is that the galaxies we are analysing in approximate dynamical equilibrium. Thus, it would seem important to remove the obviously disturbed clusters – those undergoing or having undergone recent, major mergers. As discussed in Mohr et al. (1993), the asymmetry and isophotal ellipticity of the X-ray surface brightness distribution provide information about the merger state of galaxy clusters, which is generally superior to the constraints possible from the galaxy distribution (Geller & Beers 1982) due to higher signal-to-noise ratio. However, neither of these measures are sensitive to mergers along the LOS, where the galaxy velocity distribution typically provides more information (Dressler & Shectman 1988).

Merger signatures from X-ray cluster surface brightness distributions were extracted uniformly from a large X-ray flux-limited sample of nearby clusters, indicating that over half of them exhibit statistically significant centroid variations and that the bulk of them are elliptical (Mohr et al. 1995). That study provided clear evidence that even massive clusters at low redshift are still undergoing continued accretion of subclusters. This observation is in agreement with expectations from structure formation within our standard Λ CDM model. We can use a similar approach to identify major mergers (not along the LOS) in our cluster sample. Nurgaliev et al. (2017) quantified the X-ray morphology of a subsample of 90 SZE-selected galaxy clusters using a measure of the photon asymmetry A_{phot} (Nurgaliev et al. 2013) closely related to the centroid variation (Mohr et al. 1993, 1995) and power ratios (Buote & Tsai 1995, 1996).

Table 7. Sensitivity of dynamical mass measurements to the dynamical state of clusters. We compare the masses for those clusters exhibiting large X-ray surface brightness asymmetries (unrelaxed) and those exhibiting small asymmetries (relaxed). Columns list the subsample, the mean redshift, concentration, dynamical mass given initial scaling using SPT dispersion-based masses, and ratio of the dynamical mass to the SZE dispersion-based mean mass for the subsample.

Subsample	$\langle z \rangle$	c_{200}	$M_{200}^{\text{dyn}} M_{200}^{\text{SZ}+\sigma}$ [$10^{14} M_{\odot}$]	$\frac{M_{200}^{\text{dyn}} M_{200}^{\text{SZ}+\sigma}}{\langle M_{200}^{\text{SZ}+\sigma} \rangle}$
Relaxed	0.57	$6.4_{-2.4}^{+1.4}$	$7.0_{-0.9}^{+0.8}$	$0.8_{-0.1}^{+0.1}$
Unrelaxed	0.6	$4.5_{-1.1}^{+1.2}$	$9.5_{-1.2}^{+1.0}$	$1.0_{-0.1}^{+0.1}$

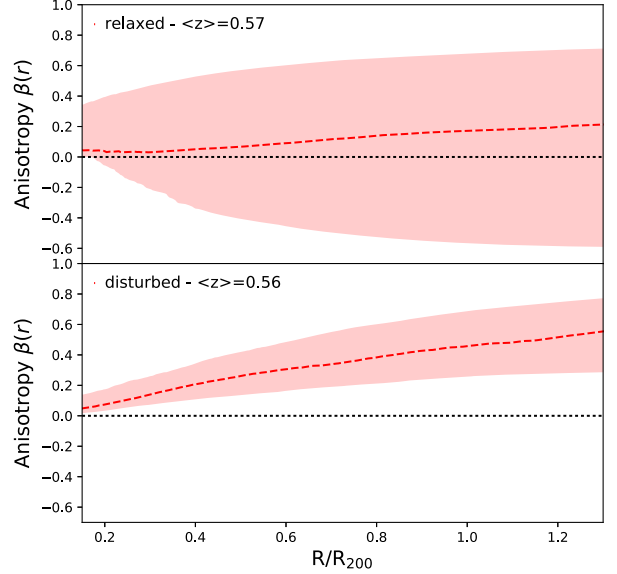


Figure 9. Velocity anisotropy profile $\beta(r)$ recovered when separately analysing relaxed and unrelaxed clusters. For the latter, the anisotropy profile indicates that the galaxies in the clusters are moving on nearly isotropic orbits near the centre and increasingly radial orbits out toward the virial radius R_{200} . The recovered anisotropy profile for the relaxed clusters is consistent with these results, but exhibits much larger uncertainties that allow for a much broader range of galaxy orbits.

To test the dependence of our results on the dynamical state of the clusters, we adopt these A_{phot} values as a measure of departures from equilibrium and separately analyse relaxed and disturbed clusters. We take here relaxed clusters to be those with $A_{\text{phot}} \leq 0.2$. Out of our 110 cluster sample, 68 (2152 spectra) have measured A_{phot} and 39 of these (1258 galaxies) are classified as relaxed. In Table 7, we show the results for the mass M_{200}^{dyn} and concentration c_{200} having performed the analysis on the stacked relaxed/disturbed populations. We find that the $M_{200}^{\text{SZ}+\sigma}$ and the dynamical masses are in agreement at $\approx 1.5\sigma$ level for the relaxed sample and at the 1σ level for the disturbed sample (statistical only).

Fig. 9 shows the recovered anisotropy profiles for the two subsamples. The anisotropy profile recovered for the disturbed clusters indicates that the galaxies are moving on roughly isotropic orbits in the cluster core and increasingly radial orbits at large distance from the cluster core. This is consistent with the behaviour seen in the total sample and most subsamples. On the other hand, we find that the anisotropy profile of relaxed clusters, while consistent with this behaviour, exhibits much larger uncertainties that allow also for very

different behaviour, including simple isotropic orbits at all radii. A comparison shows that in the case of the relaxed sample one of the anisotropy models, namely the $\beta_O(r)$ one, which has anisotropy of opposite sign in the centre and at large radii, has the highest Bayes factor. In the case of this sample, the preference is for radial orbits near the centre and somewhat tangential orbits at larger radius. This is responsible for the extension of the uncertainties to anisotropies of ~ -0.4 at $R/R_{200} \sim 1$. We note that this kind of behaviour is also shown in the upper right panel of Fig. 4, which is the redshift subsample that also prefers the β_O model (see Table 4).

The differences in the character of these results is intriguing and deserves further study with larger dynamical samples in cluster ensembles that have associated measurements of substructure. We note, however, that the results presented in this section on the velocity anisotropy profiles, the PPSD profiles, and the cluster halo masses have been compared to studies of clusters formed in numerical structure formation simulations (e.g. Navarro et al. 1996; Dehnen & McLaughlin 2005; Faltenbacher et al. 2007; Mamon et al. 2013) in which cluster substructure is generic. Observationally, substructure has been established as generic to the real cluster population at low redshift for a timespan approaching four decades (Geller & Beers 1982). With a uniform selection of cluster subsamples defined in similar ways in simulations and the real world, it should be possible in future analyses to sharpen studies like ours to measure differences in galaxy orbital characteristics associated with X-ray substructure, presence of cool cores, and so on.

5 CONCLUSIONS

We present a dynamical analysis of 110 SZE-selected galaxy clusters from the SPT–SZ survey with redshifts between 0.2 and 1.3 that have an associated spectroscopic sample of more than 3000 passive galaxies. We examine subsets of this cluster sample in redshift and mass, each comprising ~ 600 cluster members. These subsets are either combined to form composite clusters or are analysed individually using a Jeans equation-based code called **MAMPOSSt** (Mamon et al. 2013) that allows one to adopt different parametric models for the mass, galaxy, and velocity dispersion anisotropy profiles. In our analysis, we adopt an NFW mass profile, and use the measured concentration of the RS galaxy population from a complete subsample of the SPT SZE-selected cluster sample (H17), and employ five different velocity dispersion anisotropy profiles (see Section 2.2). We perform Bayesian model averaging to combine results from the different dispersion anisotropy models, because none of the five models are excluded by the data.

The velocity dispersion anisotropy profiles show the same radial features at all redshifts: orbits are isotropic near the centre and increasingly radial at larger radii. We also find no variations with cluster mass. The radial variation is broadly consistent with that seen in a recent analysis of near-infrared-selected clusters at $z \sim 1$ (Biviano et al. 2016) and also studies at low redshift ($z \lesssim 0.1$; Lemze et al. 2009; Wojtak & Łokas 2010). These trends of anisotropy with radius resemble those of DM particles in haloes extracted from cosmological numerical simulations (Mamon & Łokas 2005; Mamon et al. 2010, 2013, and references therein). The absence of a redshift trend is inconsistent with the results presented in Biviano & Poggianti (2009), where they report that passive galaxy orbits are becoming more isotropic over time. The absence of a redshift trend in the velocity anisotropy profiles suggests that the process of infall and relaxation for the passive galaxy population is occurring similarly at all redshifts since at least $z \sim 1$.

We measure the PPSD profiles $Q_{\text{gal}}(r)$ and $Q_{\text{gal},r}(r)$, using quantities derived from cluster galaxies. We find good agreement with theoretical predictions from N -body simulations of DM particles (Dehnen & McLaughlin 2005). We examine whether the amplitude of the profile scales as expected with redshift and mass under the assumption of self-similarity, finding that they do. To the extent that the PPSD profile provides constraints on the equilibrium nature of the galaxy dynamics, the good agreement with simulations suggests that galaxies behave approximately as collisionless particles and are as relaxed as the DM particles in haloes forming within cosmological structure formation simulations. Moreover, the lack of evidence for redshift trends in the power-law index of the PPSD profiles suggests again that the passive galaxy population in clusters is dynamically similar at all redshifts and mass ranges probed in our study.

We carry out a consistency check between our dynamical masses M_{200}^{dyn} , which are marginalized over uncertainties in the velocity dispersion anisotropy profiles, with masses calibrated using the SPT cluster counts with and without strong external cosmological priors. We find that our masses are smaller than those derived with strong external cosmological priors $M_{200}^{\text{SZ+Planck}}$ by $\eta = 0.81 \pm 0.06 \pm 0.08$, corresponding to a 1.9σ discrepancy when systematic uncertainties in our dynamical masses are included. Moreover, our analysis shows that the agreement is best at low redshift, while lower values of η are preferred at higher redshift. In the highest redshift bin, the best-fitting mass ratio is $\eta = 0.63_{-0.08}^{+0.13} \pm 0.06$, which corresponds to a disagreement at the 2.6σ level.

In addition, we find good agreement, $\eta' = 1.14 \pm 0.07 \pm 0.11$, between our dynamical masses and those masses extracted from the SPT cluster counts in combination with 82 externally calibrated X-ray Y_X mass estimates (de Haan et al. 2016) when the cosmological parameters are allowed to vary (and also those masses calibrated in combination with 63 velocity dispersions; see Bocquet et al. 2015). Our mass constraints are also consistent with those from related studies of SPT-selected clusters, using both weak-lensing magnification (Chiu et al. 2016a) and tangential shear (Dietrich et al. 2017; Schrabback et al. 2018; Stern et al. 2018).

Using *Chandra* X-ray data, we examine the impact of the dynamical state of the clusters on our dynamical analysis by separately analysing relaxed and unrelaxed clusters. We find dynamical masses to be in good agreement with our combined sample for both the relaxed and unrelaxed clusters. Concerning the anisotropy profiles, we find that, for the disturbed sample, the shape of the orbits resembles the one seen in the total sample and most subsamples. On the other hand, the anisotropy profile of relaxed clusters, while still consistent with this behaviour, exhibits much larger uncertainties that allow also for isotropic orbits at all radii. Further investigation with larger dynamical samples in cluster ensembles is required in order to understand the different behaviour of these objects.

As a next step, our analysis can be extended to cluster samples that include many low-mass systems. One such sample that is being analysed presently has been defined in the project known as SPIDERS (SPectroscopic IDentification of eROSITA Sources, Clerc et al. 2016), an optical spectroscopic survey of X-ray-selected galaxy clusters discovered in *ROSAT* and *XMM–Newton* imaging. Another sample is being built up through spectroscopic observations of optically selected clusters within the DES. Longer term, we expect deep spectroscopic followup of SZE and X-ray-selected clusters to provide ever larger galaxy samples that include both EL and passive galaxies. These samples will allow cluster masses to be constrained in a redshift regime where weak lensing is challenging,

while also enabling studies of the kinematic relationship between cluster EL and passive galaxies out to redshifts well beyond 1.

ACKNOWLEDGEMENTS

We thank Cristóbal Sifón for providing useful feedback. We acknowledge the support by the DFG Cluster of Excellence ‘Origin and Structure of the Universe’, the Transregio program TR33 ‘The Dark Universe’, and the Ludwig–Maximilians University (LMU). The South Pole Telescope is supported by the National Science Foundation through grant PLR-1248097. Partial support is also provided by the NSF Physics Frontier Center grant PHY-1125897 to the Kavli Institute of Cosmological Physics at the University of Chicago, the Kavli Foundation and the Gordon and Betty Moore Foundation grant GBMF 947. The Melbourne group acknowledges support from the Australian Research Council’s Discovery Projects funding scheme (DP150103208). DR is supported by a NASA Postdoctoral Program Senior Fellowship at NASA’s Ames Research Center, administered by the Universities Space Research Association under contract with NASA. Work at Argonne National Laboratory was supported under U.S. Department of Energy contract DE-AC02-06CH11357. AB acknowledges the hospitality of the LMU, and partial financial support from PRIN-INAF 2014 ‘Glittering kaleidoscopes in the sky: the multifaceted nature and role of Galaxy Clusters’ (PI: Mario Nonino). BB has been supported by the Fermi Research Alliance, LLC under contract no. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics. AS is supported by the ERC-StG ‘ClustersXCosmo’, grant agreement no. 716762.

REFERENCES

Aguerri J. A. L., Agulli I., Diaferio A., Dalla Vecchia C., 2017, *MNRAS*, 468, 364
Allington-Smith J. et al., 1994, *PASP*, 106, 983
Andersson K. et al., 2011, *ApJ*, 738, 48
Annunziatella M. et al., 2016, *A&A*, 585, A160
Appenzeller I. et al., 1998, *The Messenger*, 94, 1
Aslanyan G., Price L. C., Abazajian K. N., Easther R., 2014, *J. Chem. Phys.*, 8, 052
Austin C. G., Williams L. L. R., Barnes E. I., Babul A., Dalcanton J. J., 2005, *ApJ*, 634, 756
Barnes E. I., Williams L. L. R., Babul A., Dalcanton J. J., 2006, *ApJ*, 643, 797
Bayliss M. B. et al., 2016, *ApJS*, 227, 3
Bayliss M. B. et al., 2017, *ApJ*, 837, 88
Bellovary J. M., Dalcanton J. J., Babul A., Quinn T. R., Maas R. W., Austin C. G., Williams L. L. R., Barnes E. I., 2008, *ApJ*, 685, 739
Benson B. A. et al., 2013, *ApJ*, 763, 147
Bertschinger E., 1985, *ApJS*, 58, 39
Binney J., Mamon G. A., 1982, *MNRAS*, 200, 361
Binney J., Tremaine S., 1987, *Galactic Dynamics*. Princeton Univ. Press, Princeton, NJ, p. 747
Biviano A., Girardi M., 2003, *ApJ*, 585, 205
Biviano A., Katgert P., 2004, *A&A*, 424, 779
Biviano A., Poggianti B. M., 2009, *A&A*, 501, 419
Biviano A., Katgert P., Mazure A., Moles M., den Hartog R., Perea J., Focardi P., 1997, *A&A*, 321, 84
Biviano A., Katgert P., Thomas T., Adami C., 2002, *A&A*, 387, 8
Biviano A., Murante G., Borgani S., Diaferio A., Dolag K., Girardi M., 2006, *A&A*, 456, 23
Biviano A. et al., 2013, *A&A*, 558, A1
Biviano A., van der Burg R. F. J., Muzzin A., Sartoris B., Wilson G., Yee H. K. C., 2016, *A&A*, 594, A51
Bleem L. E. et al., 2015, *ApJS*, 216, 27

Bocquet S. et al., 2015, *ApJ*, 799, 214
Buote D. A., Tsai J. C., 1995, *ApJ*, 452, 522
Buote D. A., Tsai J. C., 1996, *ApJ*, 458, 27
Burkert A., 1995, *ApJ*, 447, L25
Carlberg R. G. et al., 1997, *ApJ*, 485, L13
Carlstrom J. E. et al., 2011, *PASP*, 123, 568
Chiu I. et al., 2016a, *MNRAS*, 457, 3050
Chiu I. et al., 2016b, *MNRAS*, 455, 258
Clerc N. et al., 2016, *MNRAS*, 463, 4490
de Haan T. et al., 2016, *ApJ*, 832, 95
Dehnen W., McLaughlin D. E., 2005, *MNRAS*, 363, 1057
Diemand J., Kuhlen M., Madau P., 2007, *ApJ*, 667, 859
Dietrich J. P. et al., 2017, preprint ([arXiv:1711.05344](https://arxiv.org/abs/1711.05344))
Dressler A., Shectman S. A., 1988, *AJ*, 95, 985
Dressler A., Smail I., Poggianti B. M., Butcher H., Couch W. J., Ellis R. S., Oemler A., Jr, 1999, *ApJS*, 122, 51
Dressler A., Hare T., Bigelow B. C., Osip D. J., 2006, *Proc. SPIE*, 6269, 62690F
Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, *MNRAS*, 390, L64
Dutton A. A., Macciò A. V., 2014, *MNRAS*, 441, 3359
Einasto J., 1965, *Tr. Astrofiz. Inst. Alma-Ata*, 5, 87
Faltenbacher A., Hoffman Y., Gottlöber S., Yepes G., 2007, *MNRAS*, 376, 1327
Feroz F., Hobson M. P., 2008, *MNRAS*, 384, 449
Feroz F., Hobson M. P., Bridges M., 2009, *MNRAS*, 398, 1601
Feroz F., Hobson M. P., Cameron E., Pettitt A. N., 2013, preprint ([arXiv:1306.2144](https://arxiv.org/abs/1306.2144))
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
Geller M. J., Beers T. C., 1982, *PASP*, 94, 421
Geller M. J., Diaferio A., Kurtz M. J., 1999, *ApJ*, 517, L23
Gerhard O. E., 1993, *MNRAS*, 265, 213
Haiman Z., Mohr J. J., Holder G. P., 2001, *ApJ*, 553, 545
Hee S., Handley W. J., Hobson M. P., Lasenby A. N., 2016, *MNRAS*, 455, 2461
Hennig C. et al., 2017, *MNRAS*, 467, 4015
Hernquist L., 1990, *ApJ*, 356, 359
Hoeting J. A., Madigan D., Raftery A. E., Volinsky C. T., 1999, *Stat. Sci.*, 14, 382
Holder G., Haiman Z., Mohr J. J., 2001, *ApJ*, 560, L111
Hook I. M., Jørgensen I., Allington-Smith J. R., Davies R. L., Metcalfe N., Murowinski R. G., Crampton D., 2004, *PASP*, 116, 425
Iannuzzi F., Dolag K., 2012, *MNRAS*, 427, 1024
Jeffreys H., 1998, *The Theory of Probability*, Oxford Classic Texts in the Physical Sciences. Oxford Univ. Press, Oxford
Katgert P., Biviano A., Mazure A., 2004, *ApJ*, 600, 657
Khullar G. et al., 2018, preprint ([arXiv:1806.01962](https://arxiv.org/abs/1806.01962))
Kronawitter A., Saglia R. P., Gerhard O., Bender R., 2000, *A&AS*, 144, 53
Lapi A., Cavaliere A., 2009, *ApJ*, 692, 174
Lemze D., Broadhurst T., Rephaeli Y., Barkana R., Umetsu K., 2009, *ApJ*, 701, 1336
Lin Y., Mohr J. J., Stanford S. A., 2004, *ApJ*, 610, 745
Lin Y.-T., Mohr J. J., 2004, *ApJ*, 617, 879
Łokas E. L., 2002, *MNRAS*, 333, 697
Łokas E. L., Mamon G. A., 2003, *MNRAS*, 343, 401
Lynden-Bell D., 1967, *MNRAS*, 136, 101
Macciò A. V., Dutton A. A., van den Bosch F. C., 2008, *MNRAS*, 391, 1940
Mamon G. A., Łokas E. L., 2005, *MNRAS*, 363, 705
Mamon G. A., Biviano A., Murante G., 2010, *A&A*, 520, A30
Mamon G. A., Biviano A., Boué G., 2013, *MNRAS*, 429, 3079
McDonald M. et al., 2017, *ApJ*, 843, 28
McGee S. L., Balogh M. L., Bower R. G., Font A. S., McCarthy I. G., 2009, *MNRAS*, 400, 937
Merritt D., 1985, *MNRAS*, 214, 25P
Merritt D., 1987, *ApJ*, 313, 121
Merritt D., Graham A. W., Moore B., Diemand J., Terzić B., 2006, *AJ*, 132, 2685

- Mohr J. J., Fabricant D. G., Geller M. J., 1993, *ApJ*, 413, 492
- Mohr J. J., Evrard A. E., Fabricant D. G., Geller M. J., 1995, *ApJ*, 447, 8 +
- Mohr J. J., Geller M. J., Fabricant D. G., Wegner G., Thorstensen J., Richstone D. O., 1996, *ApJ*, 470, 724
- Moore B., Katz N., Lake G., 1996, *ApJ*, 457, 455
- Munari E., Biviano A., Borgani S., Murante G., Fabjan D., 2013, *MNRAS*, 430, 2638
- Munari E., Biviano A., Mamon G. A., 2014, *A&A*, 566, A68
- Munari E., Biviano A., Mamon G. A., 2015, *A&A*, 574, C1
- Muzzin A., Yee H. K. C., Hall P. B., Ellingson E., Lin H., 2007, *ApJ*, 659, 1106
- Muzzin A. et al., 2014, *ApJ*, 796, 65
- Navarro J., Frenk C., White S., 1996, *ApJ*, 462, 563
- Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563
- Navarro J., Frenk C., White S., 1997, *ApJ*, 490, 493
- Navarro J. F. et al., 2010, *MNRAS*, 402, 21
- Noble A. G., Webb T. M. A., Yee H. K. C., Muzzin A., Wilson G., van der Burg R. F. J., Balogh M. L., Shupe D. L., 2016, *ApJ*, 816, 48
- Nurgaliev D., McDonald M., Benson B. A., Miller E. D., Stubbs C. W., Vikhlinin A., 2013, *ApJ*, 779, 112
- Nurgaliev D. et al., 2017, *ApJ*, 841, 5
- Old L. et al., 2015, *MNRAS*, 449, 1897
- Osipkov L. P., 1979, *Sov. Astron. Lett.*, 5, 42
- Planck Collaboration et al., 2016, *A&A*, 594, A20
- Powell M. J. D., 2006, in *Large-Scale Nonlinear Optimization*. Springer, Netherlands
- Rines K., Diaferio A., 2006, *AJ*, 132, 1275
- Rines K., Geller M. J., Kurtz M. J., Diaferio A., 2003, *AJ*, 126, 2152
- Rines K. J., Geller M. J., Diaferio A., Hwang H. S., 2016, *ApJ*, 819, 63
- Ruel J. et al., 2014, *ApJ*, 792, 45
- Saro A., Mohr J. J., Bazin G., Dolag K., 2013, *ApJ*, 772, 47
- Schaffer K. K. et al., 2011, *ApJ*, 743, 90
- Schrabback T. et al., 2018, *MNRAS*, 474, 2635
- Sereno M., Ettori S., 2017, *MNRAS*, 468, 3322
- Sifón C. et al., 2013, *ApJ*, 772, 25
- Sifón C. et al., 2016, *MNRAS*, 461, 248
- Solanes J. M., Salvador-Sole E., 1990, *A&A*, 234, 93
- Song J. et al., 2012, *ApJ*, 761, 22
- Stern C. et al., 2018, preprint ([arXiv:1802.04533](https://arxiv.org/abs/1802.04533))
- Strazzullo V. et al., 2018, preprint ([arXiv:1807.09768](https://arxiv.org/abs/1807.09768))
- Strigari L. E., Bullock J. S., Kaplinghat M., Simon J. D., Geha M., Willman B., Walker M. G., 2008, *Nature*, 454, 1096
- Sunyaev R. A., Zel'dovich Y. B., 1972, *Comments Astrophys. Space Phys.*, 4, 173
- Taylor J. E., Navarro J. F., 2001, *ApJ*, 563, 483
- Tiret O., Combes F., Angus G. W., Famaey B., Zhao H. S., 2007, *A&A*, 476, L1
- Umetsu K., Broadhurst T., Zitrin A., Medezinski E., Coe D., Postman M., 2011, *ApJ*, 738, 41
- Umetsu K. et al., 2014, *ApJ*, 795, 163
- van der Burg R. F. J., Muzzin A., Hoekstra H., Wilson G., Lidman C., Yee H. K. C., 2014, *A&A*, 561, A79
- van der Burg R. F. J., Hoekstra H., Muzzin A., Sifón C., Balogh M. L., McGee S. L., 2015a, *A&A*, 577, A19
- van der Burg R. F. J., Hoekstra H., Muzzin A., Sifón C., Balogh M. L., McGee S. L., 2015b, *A&A*, 577, A19
- van der Marel R. P., 1994, *MNRAS*, 270, 271
- van der Marel R. P., Franx M., 1993, *ApJ*, 407, 525
- van der Marel R. P., Magorrian J., Carlberg R. G., Yee H. K. C., Ellingson E., 2000, *AJ*, 119, 2038
- Vázquez J. A., Bridges M., Hobson M. P., Lasenby A. N., 2012a, *J. Chem. Phys.*, 6, 006
- Vázquez J. A., Bridges M., Hobson M. P., Lasenby A. N., 2012b, *J. Chem. Phys.*, 9, 020
- Walker M. G., Mateo M., Olszewski E. W., Peñarrubia J., Wyn Evans N., Gilmore G., 2009, *ApJ*, 704, 1274
- Wetzel A. R., 2011, *MNRAS*, 412, 49
- White S. D. M., 1986, in Kolb E. W., Turner M. S., Lindley D., Olive K., Seckel D., eds, *The Evolution of Large-Scale Structure*, p. 228
- Wojtak R., Łokas E. L., 2010, *MNRAS*, 408, 2442
- Wojtak R., Łokas E. L., Mamon G. A., Gottlöber S., Klypin A., Hoffman Y., 2008, *MNRAS*, 388, 815
- Wojtak R., Łokas E. L., Mamon G. A., Gottlöber S., 2009, *MNRAS*, 399, 812
- Wolf J., Martínez G. D., Bullock J. S., Kaplinghat M., Geha M., Muñoz R., Simon J. D., Avedo F. F., 2010, *MNRAS*, 406, 1220
- Zabludoff A. I., Franx M., Geller M. J., 1993, *ApJ*, 419, 47
- Zenteno A. et al., 2016, *MNRAS*, 462, 830
- Zhao D. H., Mo H. J., Jing Y. P., Börner G., 2003, *MNRAS*, 339, 12

¹*Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, D-81679 Munich, Germany*

²*Excellence Cluster Universe, Boltzmannstr. 2, D-85748 Garching, Germany*

³*INAF – Osservatorio Astronomico di Trieste via G.B. Tiepolo 11, I-34143 Trieste, Italy*

⁴*Max Planck Institute for Extraterrestrial Physics, Giessenbachstr., D-85748 Garching, Germany*

⁵*Kavli Institute for Cosmological Physics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA*

⁶*Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439, USA*

⁷*Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA*

⁸*Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA*

⁹*Fermi National Accelerator Laboratory, Batavia, IL 60510-0500, USA*

¹⁰*Department of Physics, University of Chicago, Chicago, IL 60637, USA*

¹¹*Department of Physics and Astronomy, University of Missouri, 5110 Rockhill Road, Kansas City, MO 64110, USA*

¹²*Enrico Fermi Institute, University of Chicago, Chicago, IL 60637, USA*

¹³*Institute of Astronomy and Astrophysics, Academia Sinica, Taipei 10617, Taiwan*

¹⁴*Department of Physics, University of California, Berkeley, CA 94720, USA*

¹⁵*Department of Physics, McGill University, Montreal, Quebec H3A 2T8, Canada*

¹⁶*Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, 452 Lomita Mall, Stanford, CA 94305-4085, USA*

¹⁷*Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA*

¹⁸*Department of Physics, Université de Montréal, Montreal, Quebec H3T 1J4, Canada*

¹⁹*Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794, USA*

²⁰*Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark*

²¹*Center for Astrophysics and Space Astronomy, Department of Astrophysical and Planetary Sciences, University of Colorado, Boulder, CO 80309, USA*

²²*NASA Ames Research Center, Moffett Field, CA 94035, USA*

²³*School of Physics, University of Melbourne, Parkville, VIC 3010, Australia*

²⁴*Department of Astronomy, University of Michigan, 1085 S. University Ave, Ann Arbor, MI 48109, USA*

²⁵*Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA*

²⁶*Department of Physics, University of California, Davis, CA 95616, USA*

²⁷*Institute of Geophysics and Planetary Physics, Lawrence Livermore National Laboratory, Livermore, CA 94551, USA*

²⁸*Cerro Tololo Inter-American Observatory, Casilla 603, La Serena, Chile*

This paper has been typeset from a \LaTeX file prepared by the author.