Mass proxy quality of massive halo properties in the ILLUSTRISTNG and FLAMINGO simulations: I. Hot gas

Aljamal, E.; Evrard, A.E.; Farahi, A.; Pillepich, A.; Nelson, D.; Schaye, J.; ... ; Braspenning, J.

Citation

Note: To cite this publication please use the final published version (if applicable).

Mass proxy quality of massive halo properties in the ILLUSTRISTNG and FLAMINGO simulations: I. Hot gas

Eddie Aljamal [1]★ August E. Evrard [1], Arya Farahi [2,3], Annalisa Pillepich [4], Dylan Nelson [5], Joop Schaye [6], Matthieu Schaller [6,7] and Joey Braspenning[4]

[1]*Department of Physics and Leinweber Center for Theoretical Physics, University of Michigan, Ann Arbor, MI 48109, USA*
[2]*Departments of Statistics and Data Sciences, University of Texas at Austin, Austin, TX 78712, USA*
[3]*The NSF-Simons AI Institute for Cosmic Origins, University of Texas at Austin, Austin, TX 78712, USA*
[4]*Max-Planck-Insitut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[5]*Universität Heidelberg, Zentrum für Astronomie, ITA, Albert-Ueberle-Str. 2, D-69120 Heidelberg, Germany*
[6]*Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, the Netherlands*
[7]*Lorentz Institute for Theoretical Physics, Leiden University, PO box 9506, NL-2300 RA Leiden, the Netherlands*

ABSTRACT

We examine scale and redshift dependence of mass–property relations (MPRs) for five hot gas properties of two large group- bservationsand cluster-scale halo samples realized by the IllustrisTNG, TNG-Cluster, and FLAMINGO cosmological hydrodynamical simulations. For intrinsic properties of (i) hot gas mass ($M_{\rm gas}$), (ii) spectroscopic-like temperature ($T_{\rm sl}$), (iii) soft-band X-ray luminosity ($L_{\rm X}$), and (iv) X-ray ($Y_{\rm X}$), and (v) Sunyaev–Zel'dovich ($Y_{\rm SZ}$) thermal energies, we use MPR parameters to infer mass proxy quality (MPQ) – the implied scatter in total halo mass conditioned on a property – for haloes with $M_{\rm 500c} \geq 10^{13}\,{\rm M_\odot}$ at redshifts, $z \in \{0, 0.5, 1, 2\}$. We find: (1) in general, scaling relation slopes and covariance display moderate to strong dependence on halo mass, with redshift dependence secondary; (2) for haloes with $M_{\rm 500c} > 10^{14}\,{\rm M_\odot}$, scalings of $M_{\rm gas}$ and $Y_{\rm SZ}$ simplify toward self-similar slope and constant intrinsic scatter (5 and 10 per cent, respectively) nearly independent of scale, making both measures ideal for cluster finding and characterization to $z = 2$; (3) halo mass-conditioned likelihoods of hot gas mass and thermal energy closely follow a lognormal form; and (4) despite normalization differences up to 0.4 dex between the two simulations, higher order scaling features such as slopes and property covariance show much better agreement. Slopes show appreciable redshift dependence at the group scale, while redshift dependence of the scatter is exhibited by low-mass FLAMINGO haloes only; (5) property correlations are largely consistent between the simulations, with values that mainly agree with existing empirical measurements. We close with a literature survey placing our MPR slopes and intrinsic scatter estimates into community context.

Key words: galaxies: clusters: general – galaxies: clusters: intracluster medium – galaxies: evolution – galaxies: formation – galaxies: groups: general – galaxies: groups: individual: ILLUSTRISTNG, TNG-CLUSTER, and FLAMINGO.

1 INTRODUCTION

Dark matter (DM) haloes are the nurseries wherein galaxies first form (e.g. J. P. Ostriker, P. J. E. Peebles & A. Yahil 1974; S. D. M. White & M. J. Rees 1978; J. Silk & G. A. Mamon 2012), then cluster into larger structures such as groups and clusters of galaxies, whose abundance across cosmic time informs studies of cosmological parameters (e.g. S. D. M. White, G. Efstathiou & C. S. Frenk 1993; G. M. Voit 2005; S. W. Allen, A. E. Evrard & A. B. Mantz 2011; A. V. Kravtsov & S. Borgani 2012; D. H. Weinberg et al. 2013; C. E. Norton, F. C. Adams & A. E. Evrard 2024). At the population level, haloes exhibit a rich variety of bulk observable properties with magnitudes driven primarily by scale (total system mass), and secondarily by astrophysical processes integrated over their hierarchical merger histories, large-scale environmental effects, and other factors. At any redshift, the measurable characteristics of stellar and hot gas contents of haloes will, in the mean, scale in some manner with total mass, and variations in formation history and astrophysical evolution will drive intrinsic covariance in such properties for haloes of fixed total mass.

The differential comoving space density of haloes as a function of their total mass – the halo mass function (HMF) – is a sensitive probe of structure growth and cosmology that has been studied with cluster samples selected by the optical–infrared (IR) galaxy content of massive haloes (M. D. Gladders et al. 2007; E. Rozo et al. 2010; E. S. Rykoff et al. 2014; A. H. Gonzalez et al. 2019; M. H. Abdullah, A. Klypin & G. Wilson 2020; T. M. C. Abbott et al. 2020; M. Aguena et al. 2021; H. Miyatake et al. 2022; Z. L. Wen & J. L. Han 2022; M. Maturi et al. 2023) as well as the thermal Sunyaev–Zel'dovich (tSZ) effect on the cosmic microwave background (CMB, N. Sehgal et al.

★ E-mail: ealjamal@umich.edu

2011; T. de Haan et al. 2016; Planck Collaboration XXIV 2016; S. Bocquet et al. 2019; S. Bocquet et al. 2024) and the extended X-ray emission (H. Böhringer et al. 2007; A. Vikhlinin et al. 2009; A. Mantz et al. 2010a; N. Mehrtens et al. 2012; A. B. Mantz et al. 2015; M. Pierre et al. 2016; F. Pacaud et al. 2018; J. Ider Chitham et al. 2020; I. N. Chiu et al. 2023) that arises from their hot intracluster plasma. Since the true masses of DM haloes in the sky are not directly measurable, empirical determination of the HMF requires observable proxies, defined by scaling relations (e.g. S. Giodini et al. 2013), to statistically map true halo masses to observable features of clusters used by the above surveys.

Considerable effort has therefore been directed to understanding scaling relations for properties such as brightest central galaxy magnitude, count of galaxies above a certain size threshold, hot gas mass, temperature, X-ray luminosity, and integrated electron pressure that manifests the thermal SZ effect. Results from a number of such studies are presented below (Section 6). Cluster sample sizes with high-quality observations are typically dozens to hundreds, and sample selection effects must be treated carefully to ensure unbiased results (A. B. Mantz 2019; S. L. Mulroy et al. 2019). Large volume simulations can generate 10 000 or more massive haloes (though many studies are smaller), which enables precise scaling relation analysis for samples complete above some minimum total mass. The upper mass scale is set by the simulation's volume, and to avoid excessive computational expense many studies selectively perform full physics treatment within specific subvolumes enclosing only the most massive haloes (e.g. D. J. Barnes et al. 2017a; D. Nelson et al. 2024).

We employ the term *mass–property relation* (MPR) to refer to the subset of scaling relations that relate true mass, defined with a spherical overdensity condition, to intrinsic properties of haloes measured within the same spherical region. MPRs for massive haloes have been studied using cosmological simulations beginning with the first generation of available codes (e.g. A. E. Evrard, C. A. Metzler & J. F. Navarro 1996; G. L. Bryan & M. L. Norman 1998). Modern simulations that include complex feedback from stars and active galactic nuclei (AGNs) show that non-gravitational processes drive systematic shifts in scaling relations away from simple self-similar expectations founded on purely gravitational evolution (S. Bhattacharya, T. Di Matteo & A. Kosowsky 2008; E. Puchwein, D. Sijacki & V. Springel 2008; D. Fabjan et al. 2010; I. G. McCarthy et al. 2010).

In particular, MPR slopes for hot gas mass tend to be steeper than self-similar because the efficiency of converting gas into stars decreases as halo mass increases (A. V. Kravtsov et al. 2006; D. Nagai 2006; D. Nagai et al. 2007; V. Biffi et al. 2014; S. Planelles et al. 2014; N. Truong et al. 2018). The action of AGN feedback also steepens slopes by preferentially expelling gas from the shallow potential wells of lower mass systems (S. Planelles et al. 2014; D. J. Barnes et al. 2017a, b; A. M. C. Le Brun et al. 2017; A. Farahi et al. 2018; N. Truong et al. 2018; N. A. Henden, E. Puchwein & D. Sijacki 2019; A.-R. Pop et al. 2022). Despite consensus on steeper than self-similar slopes, there remains disagreement about the magnitude of the slope and the level of intrinsic scatter around the mean. This may partially be due to disparate mass ranges used in each analysis, the choice of which affects MPR parameters when significant mass-dependence exists (A. M. C. Le Brun et al. 2017; A. Farahi et al. 2018; A.-R. Pop et al. 2022).

When population samples are large, thousands or more, and span a wide dynamic range in halo mass, single power-law (SPL) MPRs are likely to be inadequate. To tease out scale-dependent features in scaling relations, Kernel Localized Linear Regression (KLLR; A.

Farahi, D. Anbajagane & A. E. Evrard 2022) is a powerful approach. The method introduces a single parameter to regression analysis, the width of the kernel applied to the primary scale variable. In the limit of infinite sample size, this scale shrinks toward zero and differential forms for conditional probabilities are recovered (see Appendix A). The method was first applied by A. Farahi et al. (2018) to hot gas and stellar masses in large halo samples from the BAHAMAS[1] (I. G. McCarthy et al. 2010) and MACSIS[2] (D. J. Barnes et al. 2017a) simulations. Non-monotonic behaviours in the slope and scatter of both properties is found, with scatter in gas mass reaching a low of ~ 5 per cent for the largest haloes.

The work we present here is an extension of the aforementioned studies to incorporate a broader set of intrinsic halo properties derived from two large-volume simulation campaigns employing different galaxy formation physics treatments with independent modelling approaches: ILLUSTRISTNG (R. Weinberger et al. 2017; A. Pillepich et al. 2018a) and FLAMINGO (R. Kugel et al. 2023; J. Schaye et al. 2023). We report MPRs for five hot gas properties of simulated samples containing many thousands of haloes above $10^{13}\,M_{\odot}$ at redshift zero. Subsequently, we report stellar property MPRs (Aljamal et al., in preparation, Paper II).

The local intrinsic scatter and slope of a property's scaling with true mass determines its value as a proxy for the latter. Steeper slopes map a fixed property range onto a narrower range of mass, and the limit of zero intrinsic scatter maps a property uniquely to mass. We introduce the term *mass proxy quality* (MPQ) to refer to the *precision* with which true halo mass can be estimated from the value of an intrinsic property. The standard deviation of the conditional likelihood, $\Pr(M|S, z)$, of true halo mass, M, conditioned on a chosen property, S, at redshift, z is a natural measure of that property's quality as a mass proxy. We employ instead shape parameters from the complementary MPR likelihood, $\Pr(S|M, z)$, to define the MPQ. When likelihoods are lognormal, the population model of A. E. Evrard et al. (2014, hereafter E14) expresses the second moment of $\Pr(M|S, z)$ in terms of MPR and HMF parameters. We test the performance of this model in Section 3.3 below.

Note that this definition of proxy quality ignores the issue of *accuracy* when inverting a scaling relation. For example, previous work used a simulation-based prior on the mean gas mass fraction in clusters to derive total masses from hot gas mass measurements (A. Mantz et al. 2010b). If this prior value for gas fraction were systematically biased, so too would be the implied mass values. Indeed, MPR normalizations differ between the two simulations we study, as discussed in Section 4.2 below.

The MPQ concept is not new. Using mock X-ray images of 16 simulated clusters spanning roughly a decade in final mass. A. V. Kravtsov et al. (2006) and D. Nagai et al. (2007) highlighted gas thermal energy as superior to either gas mass or temperature in terms of MPQ. Even earlier, A. E. Evrard et al. (1996) noted that masses derived from the virial scaling of X-ray temperature alone produced mass estimates with smaller scatter compared to a hydrostatic mass treatment, because the latter included extra variance induced by the outer gas profile slope. More recently, deep neural networks trained on mock observations of hydrodynamical simulations have shown promise in improving the accuracy and reducing the scatter in cluster mass estimation using X-ray (S. B. Green et al. 2019; M. Ntampaka et al. 2019; Z. Yan et al. 2020; M. Ho et al. 2023; S. Krippendorf

---

[1] BAryons and HAloes of MAssive Systems.

[2] MAssive ClusterS and Intercluster Structures.

et al. 2024), microwave (J. D. Cohn & N. Battaglia 2020; D. de Andres et al. 2022; D. Wadekar et al. 2023a, b), and optical data (M. Ntampaka et al. 2016; M. Ho et al. 2019, 2022; D. Kodi Ramanah et al. 2020; D. Kodi Ramanah, R. Wojtak & N. Arendse 2021).

Our use of the term MPR intentionally avoids confusion with *mass–observable relations* (MORs) that aim to study directly observable features of clusters, such as the number of red galaxies defined to lie in some clustercentric sky region and photometric colour window. Such measures require integration along the line of sight and potentially observer-specific treatments. The hot gas properties we study here are certainly 'observable', but extracting unbiased measurements of intrinsic values[3] from actual observations entails a variety of data analysis methods that depend on the telescope/instrument combination, observing conditions, and other factors that lie beyond the scope of this work.

This paper is organized as follows. In Section 2, we define the gas properties used in this paper, review self-similar MPR scalings and relevant aspects of the E14 model, and describe the simulation samples used in this work. We introduce our results for the mass and redshift dependence of the MPQs of five hot gas properties in Section 3, while Section 4 presents the MPR parameters (slope, scatter, and normalization) that underlie the MPQs at redshifts of 0, 0.5, 1, and 2. In Section 5, we discuss halo mass-conditioned correlations of the five gas properties. In Section 4.3, we examine the MPR-likelihood shapes. We then compare our results to the literature in Section 6, followed by a brief summary in Section 7.

In terms of primary scale variable, we employ a mass (M_{500c}) and radius (R_{500c}) convention defined by an interior spherical density contrast of 500 times the critical density, $\rho_{crit}(z)$.

2 METHODS AND SIMULATIONS

In this section, we introduce the halo properties being used (Section 2.1), review the self-similar scaling model (Section 2.2), present the framework for mass variance estimation (Section 2.3), and finish with a description of the simulated samples (Section 2.4). Readers familiar with the basics may choose to review Tables 1–3 and examine our MPQ definition in Section 2.3 before moving to the next section.

2.1 Halo properties

As stated above, all properties are intrinsic to a halo, meaning they involve only material within a spherical region defined by R_{500c}.

2.1.1 True halo mass

We adopt the common spherical overdensity true mass convention of M_Δ, for which the summed mass of all components within a sphere of radius R_Δ encloses a mean density of Δ times the critical density, $\rho_{crit}(z) = 3H(z)^2/(8\pi G)$, thus

$$M_\Delta = \frac{4\pi}{3}\Delta\rho_{crit}(z)R_\Delta^3. \tag{1}$$

We use a value $\Delta = 500$, a scale commonly used by the X-ray and SZ community. For the statistics we study, results using $\Delta = 200$ are qualitatively and quantitatively similar.

---

[3]The term *intrinsic* here refers to spherically integrated values given in Table 1.

2.1.2 Hot gas mass

We define the hot gas mass, M_{gas}, to be the sum of all gas elements having temperature $\geq 10^5$ K within R_{500c}. This component represents the majority of baryons at cluster scales at $z = 0$, though we will show in Paper II that stars are an important component for low-mass groups. Low temperature $T < 10^5$ K gas constitutes approximately 4 per cent of the total gas phase within R_{500c} at the group scale, declining to 0.1 per cent for high-mass clusters in TNG at $z = 0$. These fractions increase with redshift at fixed halo mass (E. Rohr et al. 2024; M. Staffehl et al. 2025). For FLAMINGO, these fractions are 6 per cent and 1 per cent, respectively.

2.1.3 Hot gas temperature

While the thermal structure of the intracluster medium (ICM) is not isothermal (D. Chatzigiannakis et al. 2025), there are several common aggregate measures for gas temperature in use (E. Lee et al. 2022; S. T. Kay et al. 2024). We primarily employ a core-excised spectroscopic-like temperature, T_{sl}, defined by a density weighting

$$T_{sl} = \frac{\int n_e^2 T^{1-\alpha} dV}{\int n_e^2 T^{-\alpha} dV} \approx \frac{\sum_i n_{e,i} m_i T_i^{1-\alpha}}{\sum_i n_{e,i} m_i T_i^{-\alpha}}. \tag{2}$$

where the sum is over all gas cells within $0.15R_{500c} < r < R_{500c}$ and with temperature $k_B T_i \geq 0.1$ keV. In the above, $\alpha = 3/4$ and m_i and $n_{e,i}$ are the gas mass and electron density of the fluid element. We employ a temperature threshold of 0.1 keV to better align the comparison with observations from X-ray telescopes, such as *XMM* and *Chandra*, which are typically insensitive to photons below ~ 0.1 keV (M. C. Weisskopf et al. 2000; F. Jansen et al. 2001).

Secondarily, we employ a mass-weighted average temperature,

$$T_{mw} = \frac{\sum_i m_i T_i}{\sum_i m_i}, \tag{3}$$

where the sum is over the gas elements of mass m_i and temperature T_i that contribute to M_{gas}.

The X-ray spectra of hot gas in clusters and groups of galaxies are usually fit with a single or multitemperature thermal model (P. Mazzotta et al. 2004b; A. Vikhlinin 2006) to arrive at a spectroscopic temperature, T_{spec}. Previous work on simulated clusters with mass $M_{500c} > 10^{14}\,h^{-1}\,M_\odot$ found that the spectroscopic-like temperature reproduces T_{spec} to within a few percent (P. Mazzotta et al. 2004a; E. Rasia et al. 2005, 2014). We therefore consider T_{sl} as the simulation analogue of the observable X-ray temperature T_{spec}. Caution must be exercised in comparing T_{sl} to T_{spec} for clusters with $M_{500c} < 10^{14}\,h^{-1}\,M_\odot$ or $T_{spec} < 3$ keV. To extend the temperature range of a comparison between T_{sl} and T_{spec}, a more complicated weighting scheme for T_{sl} is required (A. Vikhlinin 2006).

We note that the density weighting aspect of T_{sl} makes it sensitive to recently heated material near a supermassive black hole (SMBH). The hot gas properties reported by J. Braspenning et al. (2024) for FLAMINGO haloes omit contributions from such recently heated particles. Specifically, for hot gas properties in FLAMINGO, we exclude cells which have been heated in the past 15 Myr and that have temperature in the range $10^{0.1}\Delta T_{AGN} \leq T_i \leq 10^{0.3}\Delta T_{AGN}$ where $\Delta T_{AGN} = 10^{7.78}$ K is the temperature change of the cell after an AGN heating event in the FLAMINGO simulation. The amount of material removed is small.

Table 1. Primary gas properties measured within a 3D aperture of R_{500c}. In the TNG simulations, we use all gas cells in this aperture, while in FLAMINGO we exclude recently heated AGN gas cells as defined below.

Property	Definition
M_{gas}	Total hot ($T \geq 10^5$ K) gas mass
T_{sl}	Core-excised, spectroscopic-like temperature, equation (2)
L_X	Observer-frame, soft X-ray luminosity, $[0.5 - 2.0]$ keV
Y_X	X-ray gas thermal energy, $Y_X = M_{gas}T_{sl}$
Y_{SZ}	tSZ thermal energy, $Y_{SZ} = M_{gas}T_{mw}$

2.1.4 Soft-band X-ray luminosity

The X-ray luminosity, L_X, in the observed soft band $[0.5 - 2.0]$ keV is summed for each gas element within R_{500c}. Following the approach of FLAMINGO (J. Braspenning et al. 2024), the contribution of each gas element is calculated by interpolating density, temperature, individual metal abundances of nine elements (H, He, C, N, O, Ne, Mg, Si, and Fe), and redshift using the CLOUDY (G. J. Ferland et al. 2017) photoionization spectral synthesis code. We use luminosities measured in the observer frame and therefore different parts of the rest-frame X-ray spectra will contribute to the $[0.5 - 2.0]$ keV band at different redshifts. For FLAMINGO, these luminosities were included in the halo catalogue, derived from previous calculations using the Spherical Overdensity and Aperture Processor (R. McGibbon et al. 2025). For TNG, we use the density, temperature, and metallicity of each gas cell to calculate each cell's luminosity by interpolating the FLAMINGO X-ray tables. Finally, we sum the contributions of all gas cells within a 3D aperture of R_{500c}. We compared the results of these calculations to the halo luminosities reported by D. Nelson et al. (2024) for TNG300 and TNG-CLUSTER at $z = 0$, finding good agreement with CLOUDY-based measurements being, on average, lower by 0.1 dex at the group scale and 0.05 dex at the cluster scale. The X-ray luminosity–halo mass scaling relation in FLAMINGO has been shown to be in agreement with a variety of observational measurements (J. Braspenning et al. 2024). D. Nelson et al. (2024) demonstrate similar agreement for TNG-CLUSTER, with a caveat discussed below.

2.1.5 Hot gas thermal energy

The tSZ effect is a distortion of the CMB spectrum that arises from inverse Compton scattering of CMB photons off free electrons in the ICM (R. A. Sunyaev & Y. B. Zeldovich 1972). The total tSZ flux density from a cluster is given by the angular integral of the tSZ sky brightness (A. C. da Silva et al. 2004; A.-R. Pop et al. 2022).

For a spherical halo isolated on the sky with no gas beyond R_{500c}, the angular integral of the tSZ signal is equivalent to a volume integral within the halo, yielding a measure of intrinsic gas thermal energy

$$Y = \frac{k_B \sigma_T}{m_e c^2} \int T_e n_e dV \ \rightarrow \ \frac{k_B \sigma_T}{m_e c^2} \sum_i n_{e,i} T_i \frac{m_i}{\rho_i}, \quad (4)$$

where $\sigma_T = 6.65 \times 10^{-25} cm^2$ is the Thomson cross section and m_e the electron mass. The discrete sum is over gas cells contributing to M_{gas} with electron density $n_{e,i}$, temperature T_i, mass m_i, and gas density ρ_i. Because $n_{e,i} \propto \rho_i$, the gas density cancels up to a factor that is effectively constant for the ionization state in the ICM. Thus, $Y \propto \sum_i m_i T_i$ and so we define the tSZ thermal energy as

$$Y_{SZ} = M_{gas}T_{mw} \quad (5)$$

and express this quantity in units of $M_\odot$ K.

X-ray observations also produce estimates of this quantity using T_{spec} and gas masses derived from X-ray imaging (A. M. C. Le Brun et al. 2017; D. Nagai et al. 2007; N. A. Henden et al. 2019). We therefore also employ the X-ray thermal energy

$$Y_X = M_{gas}T_{sl}. \quad (6)$$

Note that in FLAMINGO, these gas properties are calculated using all gas cells within a 3D aperture of R_{500c} with the exclusion of recently heated AGN cells. In ILLUSTRISTNG, we use all particles in the same aperture, irrespective of temporal proximity to feedback.

Observational scaling relation studies sometimes use properties within a fixed metric aperture. KLLR parameters in that case would behave similarly to those within R_{500c} across only a limited mass range. Extending fixed aperture measurements to the full mass range considered here would result in a mixture of different dynamical regimes across the mass spectrum. A choice of 1 Mpc, for example, would capture the virialized regions of the most massive clusters but extend into the infall regions of low-mass groups. A fixed aperture approach would thus increase the complexity and limit the interpretability of the KLLR analysis.

2.2 Self-similar MPRs

The self-similar model of structure formation (N. Kaiser 1986) assumes that haloes: (i) are in virial and hydrostatic equilibrium, (ii) possess a common internal structure in scaled units, and (iii) have a constant internal baryon fraction. While highly idealized, this model forms a baseline to which more sophisticated models can be compared. Self-similar scaling expectations for a certain property have fixed slopes in mass and redshift, but the normalizations are arbitrary.

The self-similar gas temperature will satisfy the virial theorem, $k_B T \propto \frac{GM_{500c}}{R_{500c}}$, which results in the following mass scaling and redshift evolution,

$$T \propto E(z)^{2/3} M_{500c}^{2/3}, \quad (7)$$

where we choose to quantify the redshift dependence using the evolution of the *Hubble* parameter $E(z) = \sqrt{\Omega_m(1+z)^{3/2} + \Omega_\Lambda}$ assuming a flat Lambda-cold dark matter (ΛCDM) cosmology.

For haloes in a self-similar model where larger haloes are simply scaled up from smaller haloes, the total gas mass, M_{gas}, is a constant fraction, f_{gas}, of the total halo mass M_{500c}:

$$M_{gas} = f_{gas}M_{500c} \propto E^0(z)M_{500c}. \quad (8)$$

By the definition of the tSZ and X-ray thermal energies above, the self-similar mass scaling and evolution predicts that,

$$Y_{SZ} \propto Y_X \propto E^{2/3}(z)M_{500c}^{5/3}. \quad (9)$$

The self-similar $L_{X,soft} - M_{500c}$ relation is derived in L. Lovisari & B. J. Maughan (2022) and given by:

$$L_{X,soft} \propto E^2(z)M_{500c} \quad (10)$$

Note that the redshift evolution parametrized by $E(z)$ is a consequence of the definition of halo mass in terms of the critical density.

2.3 Mass proxy quality: E14 model and KLLR method

The MPQ of a single property, gas mass or temperature, for example, is the scatter in true halo mass conditioned on the chosen property. In general, this will be both scale and redshift dependent. The definition generalizes to the case of two or more properties.

When multiproperty statistics take a Gaussian form in log-space, a convolution of the mass-conditioned property likelihood with a quadratic HMF representation in log-mass yields closed form expressions for halo counts and other population statistics. The compact forms presented by E14 are expanded in C. E. Norton et al. (2024), including an explicit function for the halo space density as a function of observed properties. This expression makes clear the degeneracies between population MPR parameters (set largely by astrophysics) with parameters describing the HMF (set by cosmology).

Due to the convolution with a steeply falling HMF needed to predict cluster counts, the mean mass selected by some chosen property is shifted low relative to the simple inverse of that property's MPR. The shift in magnitude scales as the product of the local HMF slope magnitude and the true mass variance of the property. A property that is a more effective mass proxy, one with lower halo mass variance, will thus have a smaller shift is selected mean mass.

2.3.1 MPQ definition

Our measure of MPQ is based on the E14 model expressions for mass variance conditioned on a chosen property.[4] This approach maintains true halo mass as the primary scale variable and allows us to be volume complete above our minimum chosen mass. In Section 3.3, we show that the MPQ of equation (11) matches direct estimates to better than 20 per cent for the three mass proxies with likelihoods closest to lognormal.

Using $\mu = \ln(M_{500c}/M_{\mathrm{fid}})$, where M_{fid} is a fiducial mass scale (e.g. $10^{14}\,M_\odot$), the MPQ for a single observable property, denoted by the subscript a, is then the ratio

$$\sigma_{\mu|a}(\mu, z) = \frac{\sigma_a(\mu, z)}{|\alpha_a(\mu, z)|}, \tag{11}$$

of the standard deviation in the property at fixed mass, $\sigma_a(\mu, z)$, to the magnitude of the local MPR slope, $\alpha_a(\mu, z)$.

Better mass proxies have steeper slopes and/or smaller MPR scatter. In Section 4, we demonstrate that the slope and scatter of many of hot gas MPRs generally are dependent on both scale (halo mass or property value) and redshift.

A benefit of this approach is that the MPQ can be generalized to complex cases involving more than one property. For a set of properties, s, the mass variance about the selected mean is

$$\sigma_{\mu|s}^2(\mu, z) = \left(\boldsymbol{\alpha}(\mu, z)^T \boldsymbol{C}^{-1}(\mu, z)\boldsymbol{\alpha}(\mu, z)\right)^{-1}, \tag{12}$$

where $\boldsymbol{\alpha}(\mu, z)$ is the vector of MPR slopes, formally defined in equation (A1), and $\boldsymbol{C}(\mu, z)$ is the mass- and redshift-dependent log-property covariance matrix given in equation (A4). The implied mass scatter is the square root of the variance, as usual. A smaller value of the mass scatter corresponds to better MPQ.

2.3.2 KLLR: continuous parameters from discrete populations

When sample sizes are small, under 100 or so, it is common practice to use simple linear regression of fixed slope to analyse cluster scaling relations. Such a fitting scheme does not allow for the possibility for a halo property to be a low-quality mass proxy for estimating group

[4]Because the implied mass variance is small for all properties studies here, the HMF curvature term in the mass variance can be ignored. See equations (4) and (11) in E14.

masses and improve in quality at the cluster scale, or vice versa. When sample sizes grow beyond a few thousand, a more flexible method can search for scale-dependent behaviour. Our approach uses KLLR (A. Farahi et al. 2022) to obtain scale-dependent MPQs and MPRs.

The KLLR method has previously been used to validate the E14 model for gas properties in the BAHAMAS and MACSIS simulations (A. Farahi et al. 2018), to examine scale-dependent baryon backreaction on DM scalings relations in the ILLUSTRISTNG suite (D. Anbajagane, A. E. Evrard & A. Farahi 2022a), to examine stellar properties (D. Anbajagane et al. 2020) as well as galaxy velocity segregation (D. Anbajagane et al. 2022b) in massive halo samples produced by multiple independent cosmological simulations. KLLR was also applied to various temperature measures in the FLAMINGO simulations by S. T. Kay et al. (2024).

In brief, the KLLR method extracts parameter estimates of a continuous statistical representation by applying a kernel density estimate approach to a discrete sample. Using total halo mass as the scale parameter, linear regression is performed – assuming a lognormal distribution in each bin – on property data weighted by a Gaussian in log mass centred at some scale. Regularly incrementing the centre by shifts small compared with the widths allows near-continuous measures of normalization (equation A1), slope (equation A1), scatter (equation A7), and covariance (equation A4) of a set of properties by minimizing the weighted residual sum of square errors.

In the limit of infinite sample size and zero KLLR kernel width, the method returns unbiased estimates of a model's continuous parameters, as long as the sample properties obey the underlying statistical form of TNG and FLAMINGO. With samples of roughly 2500 and 94 000 at $z = 0$, we must apply a finite width, and choose 0.2 dex in true mass for most measurements. We note that parameter estimates using a finite-width kernel can be biased if the slope and scatter of MPRs have non-zero curvature in log-mass or if the population has a strong density gradient in mass. As discussed below, we employ a narrower KLLR kernel of 0.1 dex when assessing the shape of the halo mass-conditioned property distributions in Section 4.3. More details of the method can be found in Appendix A.

2.4 Simulations

In this section, we describe the hydrodynamical simulations used in this work and the halo samples used in our analysis. A comparative summary of key simulation parameters is provided in Table 2. Table 3 provides counts of haloes above $10^{13}\,M_\odot$ for each simulation set at the redshifts of our study.

2.4.1 ILLUSTRISTNG: TNG300 and TNG-CLUSTER

We use the ILLUSTRISTNG cosmological hydrodynamic simulations (F. Marinacci et al. 2018; J. P. Naiman et al. 2018; D. Nelson et al. 2018; A. Pillepich et al. 2018b; V. Springel et al. 2018) which are run using the AREPO moving-mesh code with subgrid prescriptions incorporating gas radiative mechanisms, metal-dependent radiative cooling and heating, multimode AGN feedback sourced by SMBHs, and more. We use the flagship 302.6 Mpc periodic box, TNG300, as well as the zoom-in 1003.8 Mpc simulation, TNG-CLUSTER (D. Nelson et al. 2024), which adopt a flat ΛCDM cosmology with characteristics summarized in Table 2. As described in A. Pillepich et al. (2018b), the galaxy stellar mass function and cluster gas fraction produced in ILLUSTRISTNG are in reasonable agreement with observational constraints, although J. Schaye et al. (2023) present criticisms of the former.

Table 2. Cosmological and simulation parameters for the three simulations; the dimensionless *Hubble* constant $h \equiv H_0/100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, the normalized density of matter Ω_m, the normalized density of baryons Ω_b, the normalized density of vacuum energy Ω_Λ, the normalization of the power spectrum σ_8, the power-law index of the primordial matter power spectrum n_s, the box side-length L, the initial number of dark matter particles N_DM, the dark matter particle mass m_DM, the baryonic particle mass m_baryon, and the gravitational softening length at $z=0$ $\epsilon^{z=0}$.

Simulation	h	Ω_m	Ω_b	Ω_Λ	σ_8	n_s	L (Mpc h^{-1})	N_DM	m_DM (M$_\odot$)	m_baryon (M$_\odot$)	$\epsilon^{z=0}$ (kpc)
TNG300	0.6774	0.3089	0.0486	0.6911	0.8159	0.9667	205	2500^3	5.9×10^7	1.1×10^7	1.48
TNG-CLUSTER*	0.6774	0.3089	0.0486	0.6911	0.8159	0.9667	680	$8192^{3\dagger}$	6.1×10^7	1.2×10^7	1.48
FLAM-L1_m8	0.681	0.306	0.0486	0.694	0.807	0.967	1000	3600^3	7.06×10^8	1.34×10^8	2.85

Notes. * Zoom simulation.
$\dagger$ Effective resolution.

Table 3. Sample sizes for haloes with $M_\mathrm{500c} \geq 10^{13}\,\mathrm{M_\odot}$ in the TNG300-1, TNG-CLUSTER, and FLAMINGO-L1_m8 (FLAM-L1_m8) simulations.

Redshift	TNG300	TNG-CLUSTER	FLAM-L1_m8
2	299	320	11 087
1.0	1 290	351	47 392
0.5	2 015	352	74 300
0	2 548	352	93 815

Haloes are found using a standard friends-of-friends (FoF) percolation algorithm with a linking length of $b = 0.2$. The smaller TNG300 box has a limited number of large mass haloes; more specifically, 16 haloes with $10^{14.5} < M_\mathrm{500c}/\mathrm{M_\odot} < 10^{15}$ and only 1 halo with $M_\mathrm{500c} > 10^{15}\,\mathrm{M_\odot}$ at $z = 0$. Our study focuses on haloes in the mass range $10^{13} \lesssim M_\mathrm{500c}/\mathrm{M_\odot} \lesssim 10^{15}$ and uses a localized scaling approach to the MPRs which requires us to have an appreciable number of haloes outside this target range in order to avoid edge effects (Appendix A). For these reasons, we make use of the TNG-CLUSTER ensemble (D. Nelson et al. 2024) that performs 'zoom' resimulations on 352 cluster haloes extracted from a 1003.8^3 Gpc^3 N-body realization in order to boost the number statistics of massive clusters. Relative to TNG300, the cosmological parameters as well as the spatial resolution remain unchanged. This less computationally expensive simulation provides 204 clusters with $10^{14.5} < M_\mathrm{500c}/\mathrm{M_\odot} < 10^{15}$ and 31 clusters with $M_\mathrm{500c} > 10^{15}\,\mathrm{M_\odot}$ at $z = 0$. The TNG-CLUSTER sample is volume complete above $M_\mathrm{200c} \geq 10^{15}\,\mathrm{M_\odot}$ and randomly samples lower mass haloes such that the number is uniform in log-mass in the combined TNG-300 and TNG-CLUSTER samples.

The fact that the two TNG simulations were realized by the same code using the same mass resolution enables us to combine their halo samples for collective study of the supersample. However, the runs differ in an important parameter, α, controlling the star formation rate of cold gas. A change introduced for TNG50 (D. Nelson et al. 2019) was retained for the TNG-CLUSTER runs. In Paper II, we discuss how this parameter boosts the total stellar mass of TNG-CLUSTER haloes by ~ 30 per cent (a shift that is comparable to the impact of changing by one level of resolution in the TNG model; A. Pillepich et al. 2018b) with respect to those in TNG300 haloes at $z = 0$. Other effects on stellar property covariance structure will be discussed there, but we also document effects on hot gas phase structure that affect T_sl below.

2.4.2 FLAMINGO: L1_m8

We also study haloes in the large-volume hydrodynamical simulations of the FLAMINGO project (R. Kugel et al. 2023; J. Schaye et al. 2023) based on the SPHENIX smoothed particle hydrodynamics (SPH) scheme (J. Borrow et al. 2022) using the SWIFT

code (M. Schaller et al. 2024). The flagship simulations utilize subgrid prescriptions calibrated to reproduce the galaxy stellar mass function and cluster gas fractions at $z = 0$. All simulations assume a spatially flat ΛCDM cosmology with initial conditions created using MONOFONIC (O. Hahn, C. Rampf & C. Uhlemann 2021; W. Elbers et al. 2022) that include neutrinos. We employ the highest resolution simulation, FLAMINGO-L1_m8 (FLAM-L1_m8 in Table 2), which is run in a 1 Gpc box with purely thermal AGN feedback. In addition to the flagship simulations, the suite provides variations in cosmologies, AGN feedback models, and gas fractions. Cosmic structure is identified using an updated version of the Hierarchical Bound Tracing algorithm (HBT-HERONS, V. J. Forouhar Moreno et al. 2025) which uses a history-based approach to identify the subhaloes using an iterative unbinding procedure on particles within FoF groups, and then tracks their evolution as they merge.

Particular advantages of this simulation are: (i) it supplies statistics for nearly 100 000 haloes across our target mass range (Table 3), (ii) it employs significantly different subgrid models, for example a single-mode thermal AGN feedback as opposed to ILLUSTRISTNG's two-mode AGN feedback, as well as an SPH solver that is different from the quasi-Lagrangian mesh used in ILLUSTRISTNG, offering an opportunity to assess the robustness of population statistics between two independent implementations, and (iii) FLAMINGO reproduces observed stellar mass functions in the Galaxy and Mass Assembly survey (S. P. Driver et al. 2022) as well as the cluster gas fractions in HSC-XXL[5] (D. Akino et al. 2022) thanks to careful calibration which is detailed in R. Kugel et al. (2023).

2.4.3 Halo sample sizes

We employ samples of haloes with $M_\mathrm{500c} \geq 10^{13}\,\mathrm{M_\odot}$ at four specific redshifts. The mass range represents poor groups to the richest clusters at low redshift to massive protoclusters at $z = 2$. Counts for each simulation and redshift are given in Table 3. To avoid edge effects in KLLR statistics, we employ haloes with $M_\mathrm{500c} \geq 10^{12.5}\,\mathrm{M_\odot}$ in all analysis.

At the final epoch, there are a few thousand haloes realized in the combined ILLUSTRISTNG simulations and nearly one hundred thousand haloes realized in FLAMINGO. The ratio of TNG300 to FLAMINGO-L1_m8 counts is fairly consistent with the volume ratio of the realizations (36), while the TNG-CLUSTER sample, being biased to very high mass at $z = 0$ provides nearly the same number of haloes at $z = 2$ (320) as $z = 0$ (352).

In the KLLR analysis of each sample, the upper bound on the centre of the mass range is set by the 21st most massive halo. More massive

---
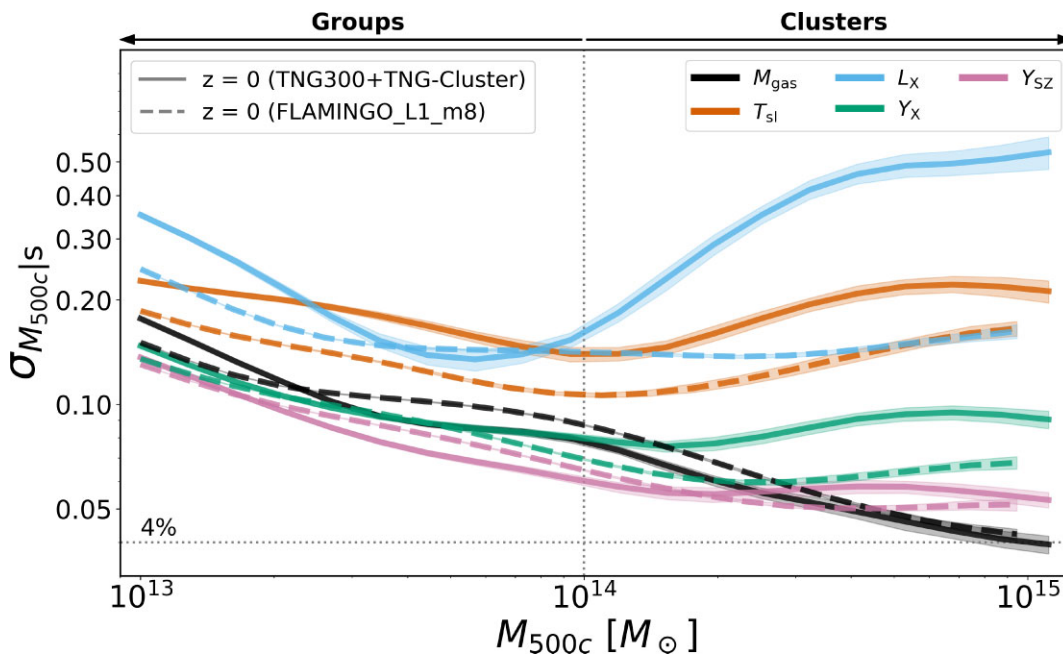[5]Hyper Suprime-Cam and XMM-XXL collaboration.

Figure 1. The MPQ, quantified by the implied halo mass scatter, equation (11), for gas properties within a 3D aperture of R_{500c} listed in the legend derived from TNG (solid lines) and FLAMINGO (dashed) halo populations. Shaded regions are 1σ uncertainties based on 1000 bootstrap samples. We indicate the low-redshift mass regimes of groups and clusters as roughly divided by a halo mass of $10^{14}\,M_\odot$. All MPQs show moderate to strong mass dependence. The relative ordering of MPQ is fairly consistent between the two simulation methods, with thermal energy (Y_{SZ}, violet) being the best mass proxy below $10^{14.5}\,M_\odot$ and hot gas mass (M_{gas}, black) best above this scale, reaching a minimum of 4 per cent halo mass scatter at $10^{15}\,M_\odot$. See the text for further discussion.

haloes contribute to the KLLR property estimates. The similarity of the TNG-CLUSTER and FLAMINGO realization volumes means that the upper mass limits of the simulations are similar at all redshifts, lying just below (FLAMINGO) or above (TNG) a value of $10^{15}\,M_\odot$ at redshift zero.

From this point onwards, we will refer to the joint TNG300 + TNG-CLUSTER as TNG and FLAMINGO-L1_m8 as FLAMINGO. Where appropriate we will explicitly distinguish between haloes in TNG300 and TNG-CLUSTER samples.

3 MASS PROXY QUALITY ASSESSMENT

Here, we present the mass and redshift dependence of the MPQs (Section 2.3) for the five primary properties listed in Table 1. We first examine proxy quality at $z = 0$, then explore evolution with redshift. The underlying MPR behaviours for the five gas properties are presented in Section 4, and their correlations in Section 5.

The MPQs are derived by applying equation (11) on the KLLR estimates of MPR scatter and slope. KLLR was performed using a kernel width of 0.2 dex with centres beginning at $M_{500c} = 10^{13}\,M_\odot$ and incrementing by ~ 0.1 dex up to the maximum value set by the 21st most massive halo at each epoch. Unless otherwise stated, all shaded regions in the figures below indicate 1σ estimated uncertainties using 1000 bootstrap samples.

3.1 Redshift zero behaviours

Fig. 1 presents KLLR estimates of the logarithmic true halo mass scatter, $\sigma_{\mu|s}(\mu, z = 0)$, for M_{gas} (black), T_{sl} (orange), L_X (blue), Y_X (green), and Y_{SZ} (violet) in the halo samples of TNG (solid lines) and FLAMINGO (dashed lines).

The MPQs of all properties show a dependence on halo mass, with generally higher values of $\sigma_{\mu|s}$ seen at the group scale. Some MPQs

decrease monotonically with halo mass, while others decline toward the group/cluster boundary then rise again at the highest masses.

The amount of gas and its thermal energy content are two key bulk measures accessible by X-ray and SZ observations. Fig. 1 shows that both simulations find Y_{SZ} to be the best mass proxy at the group and low-mass cluster scale, with values starting at ~ 13 per cent declining to 6 per cent near $10^{14.5}\,M_\odot$. Above this halo mass scale, M_{gas} is a slightly better proxy, with $\sigma_{\mu|s}$ declining steadily to a value of 4 per cent at $10^{15}\,M_\odot$. This minimum value is consistent in both simulation samples.

As explained below, the MPRs of M_{gas} and Y_{SZ} differ in subtle ways. Consistently in TNG and FLAMINGO, both the slope and scatter in the M_{gas} MPR gently decline across the cluster mass range, in a manner that produces the declines seen in Fig. 1. Those of Y_{SZ}, however, tend toward near constant values of 0.1 in scatter and a nearly self-similar slope, $\alpha = 5/3$, yielding the nearly constant MPQ of roughly 5 per cent for cluster-scale halo with $M_{500c} > 10^{14}\,M_\odot$.

These bulk measures integrate over the internal thermal structure of the hot ICM plasma. The phase space structure expressed as filling fractions in the plane of density and temperature/entropy is driven by a number of physical and numerical factors. We now turn to examine two properties that are sensitive to this internal structure.

3.1.1 Non-monotonic MPQ behaviours

At all mass scales and in both simulation samples, the two least effective mass proxies are X-ray luminosity, L_X, and temperature, T_{sl}. Specific behaviours emerge for the two simulation methods, particularly for L_X at the highest halo masses.

In FLAMINGO, the MPQs of L_X and T_{sl} start near a value of 0.2 and decline at different rates with increasing halo mass. In T_{sl}, the mass scatter is minimized near 10 per cent at $10^{14}\,M_\odot$ at which scale the

value for L_X is significantly higher, ~ 15 per cent. For T_{sl}, the mass scatter rises at larger halo masses, while that of L_X remains nearly flat. As a result, the MPQ of both properties track each other above $10^{14.5}\,M_\odot$, gently rising to a value 0.16 at $10^{15}\,M_\odot$.

For TNG, the MPQ of T_{sl} has a similar shape to that of FLAMINGO, but its magnitude is shifted consistently high, by ~ 5 per cent, throughout the halo mass range shown. The MPQ of L_X is strongly mass-dependent, falling from 35 per cent to 16 per cent in the group mass regime before rising again to values above 50 per cent at the highest masses.

Differences in internal thermal structure, particularly near the site of strong AGN feedback within the core region, are likely to be driving the sample-based MPQ differences for L_X and T_{sl} seen in Fig. 1. We show in Appendix B that core-excision significantly lowers the mass scatter for X-ray luminosity in TNG at high halo masses, bringing it into better agreement with FLAMINGO sample behaviour.

In Appendix C, we show that the MPQs for mass-weighted temperature, T_{mw}, in the two simulations are more consistent in magnitude compared to the case of T_{sl}. Again, the density weighting of the latter makes it more susceptible to thermal structure of the plasma. Note that we excise the core gas contribution when determining T_{sl}, so the phase structure contributions lie at radii beyond $0.15R_{500c}$. We defer a detailed examination of thermal phase structure in these simulations to future work.

3.1.2 Thermal energy measures

At the group scale, Y_X and Y_{SZ} perform better than M_{gas} for both simulation samples. The Y_X MPQ is nearly identical to that of M_{gas} in TNG for haloes with masses $3 \times 10^{13}\,M_\odot < M_{500c} < 10^{14}\,M_\odot$. A the cluster scale, Y_X and Y_{SZ} exhibit a higher mass scatter than M_{gas} in both simulations for haloes larger than $3 \times 10^{14}\,M_\odot$, making M_{gas} the best mass proxy for the largest haloes. The primary reason for this is the increased scatter of Y_X at the cluster scale and the mass-independent scatter of Y_{SZ}, while the heightened correlation between M_{gas} and T_{sl} or T_{mw} increases the mass scatter as a secondary effect (see Section 5).

The increase in Y_X mass scatter at the cluster scale is more pronounced in TNG compared to that in FLAMINGO. The reason for this is that TNG-CLUSTER haloes contain more hot gas at high densities than their TNG300 counterparts, which drives extra variance in T_{sl} and, thereby, Y_X. The MPQ of Y_X in the TNG300 sample alone agrees more closely with values seen in FLAMINGO at cluster masses.

3.2 MPQ evolution with redshift

Fig. 2 shows true halo mass scatter, $\sigma_{\mu|s}(\mu, z)$, at redshifts $z = 0.5$, 1, and 2 (top to bottom). Colours and scales are consistent in each panel and consistent with that of Fig. 1.

One way to consider this evolution is to examine the highest mass scale representing rare peaks in the initial density field able to have collapsed at each redshift (J. M. Bardeen et al. 1986; J. R. Bond et al. 1991). Effectively the mass limit represents a fixed comoving number threshold of 20 objects per cubic gigaparsec. The MPQs for Y_{SZ} and M_{gas} at this limit remain close to their $z = 0$ values, lifting gently above the 4 per cent dotted line at $z \geq 1$.

In general, the redshift dependence of MPQs is weaker than the mass dependence. An exception is the case of L_X mass scatter in TNG. We anticipated that the large discrepancy with FLAMINGO at $z = 0$ is due to the inclusion of dense hot gas residing near large
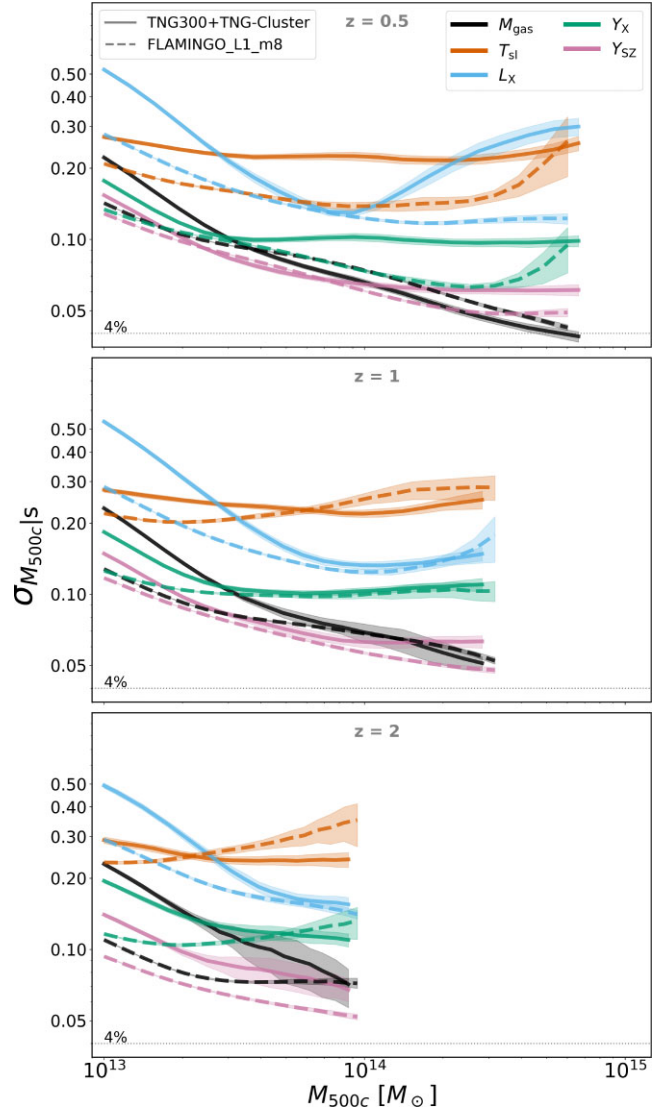

Figure 2. Halo mass and redshift dependence of the MPQs using the same format as in Fig. 1 for redshifts $z = 0.5$ (top), 1 (middle), and 2 (bottom). At each redshift, lines extend up to the mass of the 21st ranked halo in each simulation sample.

central galaxies. This effect is sensitive to redshift. At $z > 1$, the MPQs for L_X of the two simulations lie in agreement at the highest sampled masses. The MPQ values diverge toward the lowest masses, as they do at $z = 0$, but at higher absolute values.

The redshift evolution of the hot gas mass MPQ behaves differently in the two simulation samples. While similar values are found at $z = 0$ across the entire halo mass range, by $z = 2$ a difference of a factor of two emerges at $10^{13}\,M_\odot$. We show in the following section that this is primarily driven by redshift-independent scatter in M_{gas} at this mass scale in the TNG halo sample.

At the mass scale of $10^{14}\,M_\odot$, haloes are sufficiently numerous to be accessible to observations across the full redshift range, $0 < z < 2$, studied here. At that scale, both simulations provide a reasonably consistent picture of MPQ. The measures Y_{SZ}, M_{gas}, and Y_X provide 0.1 or better total mass scatter across all redshifts, while T_{sl} and L_X have scatter greater than 0.1.

The relatively poor performance of X-ray luminosity should not be interpreted as a mark against cluster samples selected by X-ray
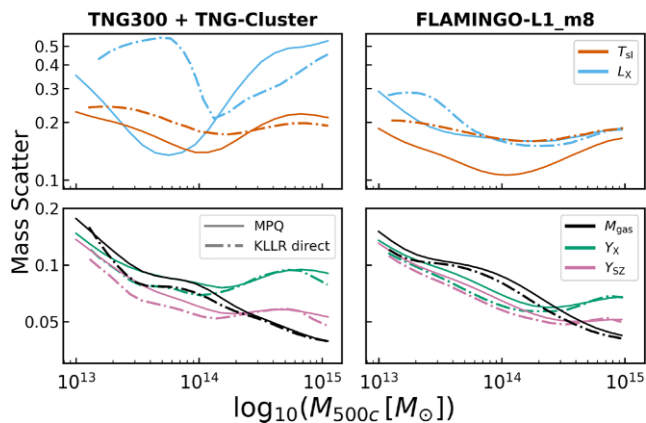
Figure 3. Comparison of $z = 0$ scatter in true halo mass inferred by the MPQ, equation (11) (solid lines), with values directly measured by KLLR applied to the conditional likelihood, $P(M|S)$, of each hot gas property, S (dot–dashed). The lower panels show good agreement for properties that adhere well to lognormality, while the agreement degrades in the upper panels for properties with more complex likelihood shapes (see Section 4.3).

flux. At 10^{14} M$_\odot$, we note that there is a nearly redshift-independent value of the MPQ of ~ 0.15, consistent in both simulated samples. As we show in Appendix B, the divergence in behaviors at higher masses is due to excess scatter in core emission in TNG. This is not true at the group scale, however.

The MPQ of thermal SZ is also nearly redshift independent at 10^{14} M$_\odot$, with values $6 - 8$ per cent. Projection effects will broaden this intrinsic measure, however, in a manner that will generally depend on cluster mass, redshift, and angular scale of the measurement (e.g. L. D. Shaw, G. P. Holder & P. Bode 2008; N. Gupta et al. 2017).

3.3 MPQ versus direct true mass scatter measurement

We now examine how well the MPQ measure of equation (11) reflects direct measurement of true mass scatter obtained by performing the KLLR fits to the MPR inverse, $\Pr(M|S, z)$. We perform this regression for each gas property with minimum values set by requiring the sample be mass complete above $M_{500c} = 10^{13.25}$ M$_\odot$. A KLLR kernel width of 0.2 dex is applied to all. The log-mean mass of each property is used to map each to a common halo mass axis.

Fig. 3 compares the $z = 0$ MPQ estimates (solid lines) with the direct halo mass scatter measurements (dot–dashed) for TNG (left) and FLAMINGO (right). Lower panels show cases of properties which adhere closely to lognormal likelihood shapes (M_{gas}, Y_{SZ}, and Y_X). The MPQ values tend to lie somewhat higher than the direct measures, but the differences are consistently below 20 per cent. The upper panels show that larger discrepancies exist for L_X and T_{sl}. Both of these properties have more complex conditional likelihood shapes, as discussed in Section 4.3, so the larger disagreement is expected. Importantly, the mass-dependent shapes and relative ordering of MPQ are consistent in both measures, with the exception of the behaviour of X-ray luminosity in the TNG sample. We show in Appendix B that core-excised X-ray emission adheres more closely to a lognormal shape in both simulation samples.

We note that a similar assessment of the E14 model was performed by A. Farahi et al. (2018), but that work examined the mean halo mass conditioned on either hot gas mass or total stellar mass. A. Farahi et al. (2018) demonstrate that the second-order HMF correction term for mean halo mass achieves a subpercent level accuracy relative to

direct KLLR measures (see their fig. 8) at the cluster scale. This work is the first to test E14 estimates of the second moment.

4 MASS-PROPERTY RELATION ANALYSIS

We now turn to the mass and redshift dependence of the MPRs themselves. In Section 4.1, we present the slope and scatter of each property in light of the MPQ trends seen above. Property normalizations, framed as deviations from simple self-similar expectations, are discussed in Section 4.2.

We selectively compare our results with those of past simulation studies in this section, offering a more extended review of MPR slope and scatter in Section 6. In future work, we plan to investigate the impact of different AGN and supernovae feedback models and to compare them with observations using the physics variations provided by the FLAMINGO simulation suite (J. Schaye et al. 2023).

4.1 Slope, scatter, and their relationships to MPQ

In Figs 4 and 5, we present the mass and redshift dependent slopes and scatters – resulting from equation (A1) for the KLLR slope and equation (A7) for the KLLR scatter – , respectively, for the five primary gas properties at four redshifts. The column on the left presents results for the TNG halo population and the right column shows those of FLAMINGO. Horizontal dotted lines in each panel of Fig. 4 display the self-similar model expectations for the property slopes (Section 2.2).

4.1.1 Hot gas mass

For the hot gas mass–halo mass relation, the slope is steeper than the self-similar value of 1 throughout the mass range across all redshifts. In TNG, the slope decreases from 1.7 to 1.02 as mass increases from $M_{500c} = 10^{13}$ M$_\odot$ to 10^{15} M$_\odot$ at $z = 0$, while in FLAMINGO, it is systematically steeper by $10 - 20$ per cent for all redshifts, approaching a value of 1.09 for the largest haloes at $z = 0$. This significant deviation from the self-similar expectation can be attributed to two primary factors: (1) the conversion of gas content into stars through radiative cooling and star formation (D. Nagai et al. 2007; N. Battaglia et al. 2013; S. Planelles et al. 2014; A. M. C. Le Brun et al. 2017; N. Truong et al. 2018), and (2) the ejection of gas by AGN feedback events, which occurs more efficiently in lower mass haloes due to their shallower potential wells (E. Puchwein et al. 2008; D. Fabjan et al. 2010; I. G. McCarthy et al. 2011; M. Gaspari et al. 2014; A. M. C. Le Brun et al. 2017; N. Truong et al. 2018; M. Ayromlou, D. Nelson & A. Pillepich 2023). Specifically, M. Ayromlou et al. (2023) show that the ratio of closure radius R_c, the radius within which all baryons associated with a halo can be found, to R_{200c} for TNG groups is more than three times larger than that for clusters. The deeper potential well of clusters helps retain their gas content within R_{200c}.

Both simulations show a $10 - 20$ per cent decrease in slope with increasing redshift for haloes with mass $M_{500c} \lesssim 10^{14}$ M$_\odot$, while the slope remains approximately redshift-independent for more massive haloes. This trend aligns with previous numerical studies (A. M. C. Le Brun et al. 2017; A. Farahi et al. 2018; N. A. Henden et al. 2019), which link it to the increased binding energy at a fixed halo mass with increasing redshift, $E_{bind}(z) \propto M_{500c}^{5/3} E(z)^{2/3}$. At fixed mass, the larger binding energy with increased redshift suggests that it is more difficult to eject gas due to AGN feedback
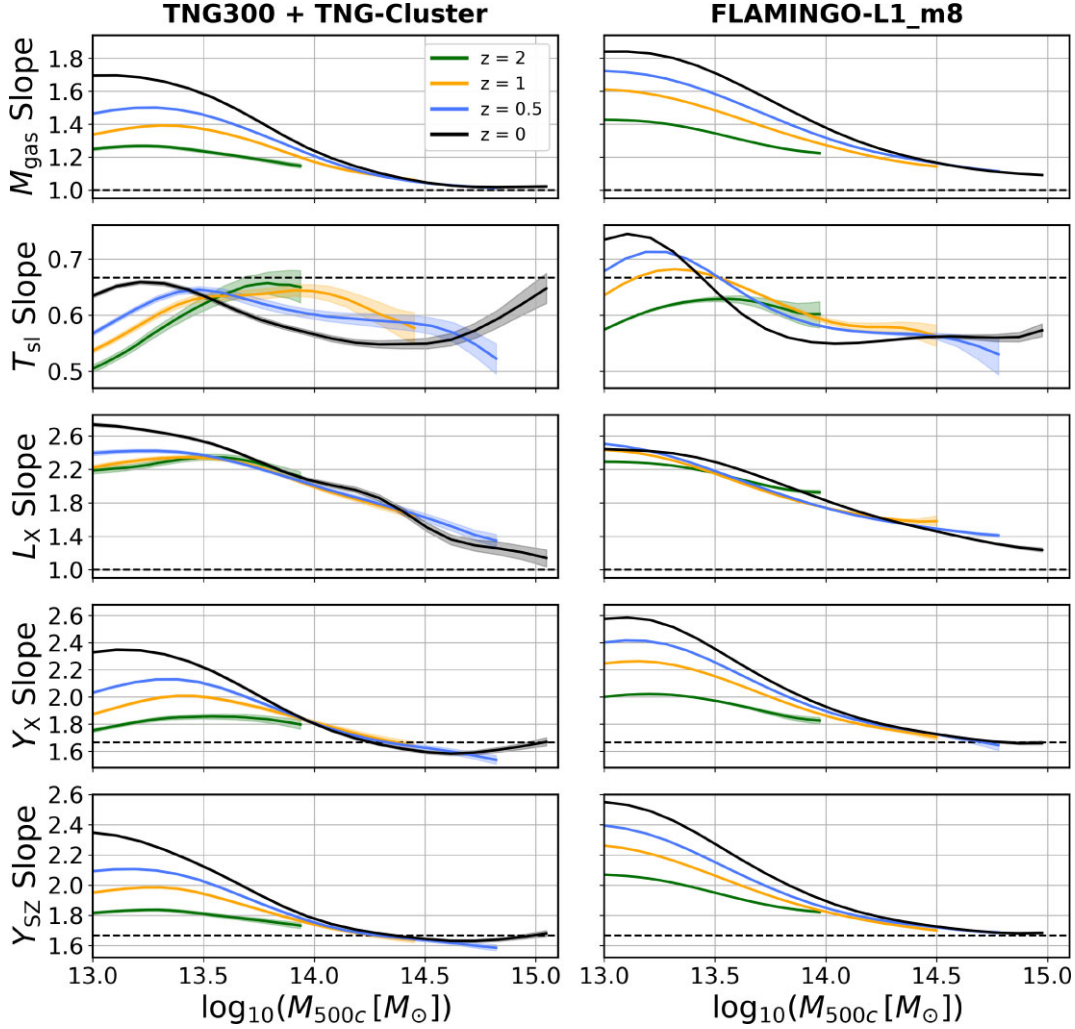
Figure 4. Halo mass and redshift dependence of the MPR slopes for (top to bottom) M_{gas}, T_{sl}, L_X, Y_X, and Y_{SZ}. The KLLR-derived values (equation A1) are shown at redshifts $z = 0$ (black), 0.5 (blue), 1.0 (gold), and 2 (green) for TNG (left) and FLAMINGO (right) halo populations. Dotted lines indicate the self-similar slopes given in Section 2.2. Note that, by definition, the slope in Y_X is the sum of the slopes in M_{gas} and T_{sl}.

at larger redshifts, resulting in a shallower M_{gas} slope. Moreover, M. Ayromlou et al. (2023) show that R_c/R_{200c} decreases with increasing redshift, suggesting that haloes retain an increasing fraction of their hot gas at higher redshifts. N. Truong et al. (2018) observe a slight increase in the $M_{gas}-M_{500c}$ slope for $z > 1$ which they attribute to declining gas fractions in lower mass haloes at those redshifts.

The intrinsic scatter in the $M_{gas}-M_{500c}$ relation, shown in the top panels of Fig. 5, decreases monotonically from ~ 30 per cent for low-mass groups, reaching ~ 4 per cent for the most massive haloes. Consistent with other simulation studies (A. M. C. Le Brun et al. 2017; A. Farahi et al. 2018; N. A. Henden et al. 2019), FLAMINGO shows a 10 per cent reduction of the scatter only at the group scale with increasing redshift, likely due to the reduced influence of non-gravitational physics on galaxy formation as the binding energy rises at fixed mass. In contrast, the scatter is redshift-independent in TNG throughout the mass range corroborating the results by D. J. Barnes et al. (2017a) that show no redshift dependence in the scatter.

The declining MPQ for M_{gas} at $z = 0$ with increasing halo mass is primarily due to the falling M_{gas} scatter, with the mass-

dependent slope having a softening effect on the magnitude range from groups to clusters. In FLAMINGO, the decrease in slope with redshift compensates the decrease in scatter, rendering the MPQ to be approximately redshift-independent. Conversely, in TNG, the lower slope and constant scatter at the group scale with increasing redshift lead to a rise in mass scatter for $z > 0$. Though FLAMINGO's slope is larger than that of TNG, the scatter at most masses is also slightly larger, leading to an agreement in the M_{gas} MPQ between the two simulations.

Overall, the finding of very small, redshift-independent scatter in M_{gas} at high cluster masses supports the use of X-ray gas mass fraction as a cosmological probe (A. B. Mantz et al. 2022).

4.1.2 Gas temperature

The $T_{sl}-M_{500c}$ relation lacks a clear monotonic trend in both the slope and the scatter compared to the other properties, the range of both parameters is relatively modest. At $z = 0$, the slope in TNG is slightly shallower than the self-similar expectation of 2/3, but increases toward this value for haloes with $M_{500c} > 10^{14.5}\,M_\odot$, in agreement with A.-R. Pop et al. (2022) who use a smoothly broken
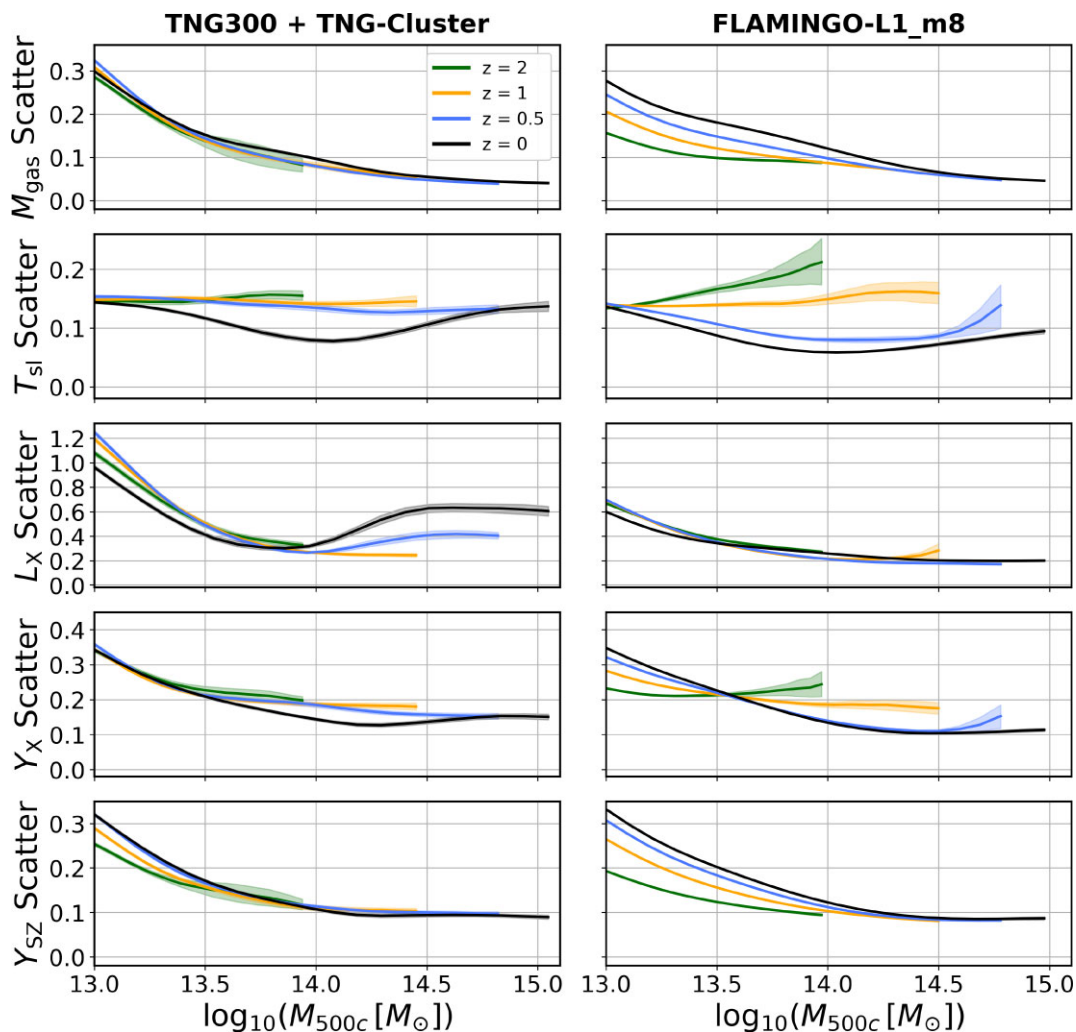
Figure 5. Halo mass and redshift dependence of the MPR intrinsic scatter values (equation A7). Format is identical to that of Fig. 4.

power law (BPL) and S. T. Kay et al. (2024) who implement KLLR on the lower resolution FLAMINGO-L1_m9. In contrast, in FLAMINGO, the slope is slightly steeper than self-similar for $M_{500c} \lesssim 10^{13.5} \, M_\odot$, becoming shallower for larger haloes, approaching a value of 0.57 for $M_{500c} > 10^{14.5} \, M_\odot$. As shown in Appendix C, many of these features are also apparent in the mass-weighted temperature, T_{mw}, although the scatter is considerably smaller, consistent with S. T. Kay et al. (2024).

Departures from the self-similar expectation, where the temperature reflects only the depth of the potential well, can be physically linked to the steeper than self-similar slopes observed for $M_{gas}-M_{500c}$ seen above. Radiative cooling removes low entropy gas through the formation of stars, decreasing gas density and increasing temperature. Moreover, AGN feedback ejections remove low entropy gas and increase the average entropy, thereby increasing the temperature while lowering the gas fraction. This increase in temperature flattens the slope of the temperature–halo mass scaling relation.

The slopes of both the $T_{sl}-M_{500c}$ and $T_{mw}-M_{500c}$ relations slightly increase with redshift in the mass range $10^{13.5} \, M_\odot < M_{500c} < 10^{14.5} \, M_\odot$ and decrease by ~ 15 per cent for haloes with $M_{500c} < 10^{13.5} \, M_\odot$. This aligns with the modest increases reported in the slope

of the $T_{mw}-M_{500c}$ relation by A. M. C. Le Brun et al. (2017) and the slope of the $T_{sl}-M_{500c}$ relation reported by N. Truong et al. (2018). Other numerical studies (D. J. Barnes et al. 2017a; N. A. Henden et al. 2019) report no significant redshift-induced changes in the temperature–halo mass relation. Because the effectiveness of AGN feedback in expelling gas from haloes of a fixed mass diminishes with redshift, the departure of the temperature slopes from self-similarity become less pronounced with increasing redshift.

At $z = 0$, the scatter in T_{sl} exhibits weak mass dependence, decreasing from 14 per cent at the group scale to ~ 8 per cent at $M_{500c} = 10^{14} \, M_\odot$, then increasing for larger haloes. These values are consistent with the intrinsic scatter of the temperature–halo mass relation reported by D. Nagai et al. (2007), A. M. C. Le Brun et al. (2017), and N. A. Henden et al. (2019) who measure a scatter of 14 per cent, 5 per cent and $9 - 12$ per cent, respectively.

In TNG, the scatter in T_{sl} increases at $z = 0.5$ relative to $z = 0$ but then remains both mass- and redshift-independent for $z \geq 0.5$. In FLAMINGO, the scatter closely follows that of TNG at $z = 0$, but the increase with redshift is delayed to $z > 0.5$. In contrast, we show in Appendix C that the T_{mw} scatter is redshift-independent for $M_{500c} \gtrsim 10^{13.5} \, M_\odot$. Above $10^{14} \, M_\odot$, the scatter

lies between 5 and 7 per cent independent of redshift in both simulation samples. From the T_{sl} panels of Fig. 5, we see that the two simulations produce remarkably similar scale-dependent scatter at $z = 0$, with non-monotonic behaviour minimized at 10^{14} M$_\odot$. In TNG, the scatter is roughly 40 per cent higher at the upper mass limit of 10^{15} M$_\odot$. At larger redshifts, the scatter at 10^{13} M$_\odot$ remains consistent for both simulations, with amplitude close to 0.15, but TNG haloes show no mass or redshift dependence about this value, whereas the FLAMINGO samples show mild dependence on both.

Overall, the mild scale behaviour of the slope and scatter are reflected in the approximately flat T_{sl} MPQ seen in Figs 1 and 2. For $z < 1$, the lower intrinsic scatter of T_{sl} in FLAMINGO leads to a reduced mass scatter compared to TNG values. This difference is not apparent at $z \geq 1$.

4.1.3 X-ray luminosity

The slope of the L_X–M_{500c} relation is significantly steeper than the self-similar expectation of 1, particularly at the group scale, consistent with the bolometric luminosity–total mass relation slopes reported previously (D. J. Barnes et al. 2017a; A. M. C. Le Brun et al. 2017; N. Truong et al. 2018; N. A. Henden et al. 2019; A.-R. Pop et al. 2022). At $z = 0$, the slope decreases approximately linearly from the mid-group scale to massive clusters, from ∼ 2.3 to ∼ 1.2, approaching the self-similar expectation of one for the most massive haloes. The nearly redshift-independent, linearly declining slope for halo mass scales > $10^{13.5}$ M$_\odot$ indicates that a quadratic form for the $L_X − M$ relation would sufficiently capture the scale dependence.

Since the X-ray luminosity is a volume integral of the square of the gas density, the departure from self-similarity in M_{gas}, is reflected in the steep L_X scaling with halo mass. In Appendix B, we show that, unlike the scatter, the slopes are largely unchanged when cores are excised. The scatter is dramatically reduced for the TNG sample, but less so for FLAMINGO (see Fig. B2).

Strikingly, L_X parameters exhibit some of the weakest redshift evolution among all gas properties, particularly in FLAMINGO. We note that the exclusion of recently heated gas particles in that simulation will reduce upward excursions in L_X. For TNG-CLUSTER haloes at $z = 0$, there is a subpopulation with X-ray luminosities nearly 1 dex higher than those observed (D. Nelson et al. 2024), which can be attributed to the abundance of cool cores in TNG-CLUSTER haloes (K. Lehle et al. 2024): 24 per cent and 60 per cent of the 352 TNG-CLUSTER haloes are strong and weak cool cores, respectively. This tail drives the TNG scatter in Fig. 5 upward to a value of 0.6 at the massive cluster scale, significantly higher than the 0.2 value for the FLAMINGO sample. The scatter in TNG-CLUSTER for high halo masses is significantly larger (∼ 50 per cent) than that in TNG300 at the same mass scale. In Fig. B2, we observe that $L_{X,ce}$ does not exhibit the same upturn in intrinsic scatter as L_X for high-mass haloes which are dominated by the TNG-CLUSTER sample, providing further evidence that the increased scatter is a result of the abundance of luminous cool cores.

Fig. 4 shows that TNG slopes lie slightly above those of FLAMINGO, consistent with the slight elevation in M_{gas} slopes. For poor groups, the TNG population shows a steepening slope over time that is more pronounced than that seen for FLAMINGO. The lack of redshift evolution in slope at scales > $10^{13.5}$ M$_\odot$, noted above, contrasts with prior findings (A. M. C. Le Brun et al. 2017; N. Truong et al. 2018; N. A. Henden et al. 2019) of slopes that increase with redshift. We note that fixed-slope linear regression will be probing lower halo mass

ranges at larger redshift, so the drift reported may simply reflect population evolution in the HMF. Consistent with our findings, D. J. Barnes et al. (2017a) report no change in the slope of the core-excised bolometric luminosity with redshift.

Despite similar slopes, the intrinsic scatter in Fig. 5 differs between the two simulations. They are closest at the group-cluster transition scale of 10^{14} M$_\odot$, where values $0.20 − 0.25$ are seen at $z \geq 0.5$. The TNG population scatter is consistently larger at the group scale at all redshifts, reaching values above one at 10^{13} M$_\odot$.

As noted in Section 3, the late-time ($z < 1$) increase in TNG scatter at high halo masses degrades the MPQ of L_X on those scales. The scatter in core-excised luminosity does not exhibit this upturn and aligns better with the scatter in FLAMINGO haloes. For FLAMINGO clusters, the scatter is essentially flat at ∼ 20 per cent above $M_{500c} \gtrsim 10^{14}$ M$_\odot$. In Fig. 1, the MPQ of L_X in FLAMINGO gently rises with mass at these scales because the slope gently falls.

4.1.4 Gas thermal energy

The total thermal energy measure, Y_{SZ}, is a fundamental measure of halo gas since pressure gradients strive to balance gravity as a halo dynamically relaxes (E. T. Lau, A. V. Kravtsov & D. Nagai 2009). As the product of gas mass and mass-weighted temperature, the scatter in Y_{SZ} is sensitive to the correlation between T_{mw} and M_{gas} conditioned on halo mass. Similarly, the scatter in Y_X depends on the correlation between T_{sl} and M_{gas}. We discuss property correlations in Section 5.

At group-scale masses, the Y_X–M_{500c} and Y_{SZ}–M_{500c} relations display slope and scatter behaviours similar to those of the M_{gas}–M_{500c} relation. Both relations have slopes that deviate increasingly from self-similar over time. At $z = 0$, the maximum slopes of Y_X (Y_{SZ}) in TNG and FLAMINGO are 2.33 (2.34) and 2.58 (2.55), respectively. Similar to M_{gas}, slopes at all redshifts are consistently larger by ∼ 10 − 20 per cent in FLAMINGO relative to TNG.

The scatter in Y_X and Y_{SZ} decreases from ∼ 30 per cent at $M_{500c} = 10^{13}$ M$_\odot$ to ∼ 10 per cent and 12 per cent at $M_{500c} = 10^{14}$ M$_\odot$, respectively. While this scatter is larger than the value of 0.053 found in D. Nagai et al. (2007), it aligns with A. M. C. Le Brun et al. (2017), D. J. Barnes et al. (2017a), and N. A. Henden et al. (2019) who measure values of 10 per cent, 12 per cent and 14 per cent, respectively. For more massive haloes, the scatter in Y_X plateaus at ∼ 10 − 15 per cent for haloes with $M_{500c} > 10^{14}$ M$_\odot$. In FLAMINGO, the redshift dependence of the Y_X scatter is confined to groups, decreasing with increasing redshift. The larger scatter in Y_X at the massive cluster scale allows M_{gas} to be a superior mass proxy for haloes with $M_{500c} > 10^{14.5}$ M$_\odot$.

Unlike the previous properties, Fig. 4 shows that the SZ signal, Y_{SZ}, scales in a nearly self-similar manner at high halo masses. Above $10^{14.5}$ M$_\odot$ in both simulations, the slopes of both gas thermal energy measures lie within 0.05 of the 5/3 self-similar value. This finding is consistent with several existing results (A. M. C. Le Brun et al. 2017; D. J. Barnes et al. 2017a; N. Truong et al. 2018; N. A. Henden et al. 2019; A.-R. Pop et al. 2022).

At high masses, Fig. 5 shows that the scatter in Y_{SZ} in both simulated samples are remarkably mass and redshift independent with values close to 0.1. Hence, the MPQ of the SZ thermal energy at $z = 0$ (Fig. 1) is nearly constant, taking values of $6 − 8$ per cent for high-mass haloes at all redshifts in both populations.

This finding reinforces observational efforts to identify clusters using the SZ effect (N. Sehgal et al. 2011; T. de Haan et al. 2016; Planck Collaboration XXIV 2016; S. Bocquet et al. 2019, 2024). A

recent Atacama Cosmology Telescope (ACT-DR6) sample of nearly 10 000 clusters is nearly complete for haloes with $10^{14.3}\,\mathrm{M}_\odot$ for $z \leq 1$ (Bolliet et al., in preparation). We find that the variance in Y_{SZ} above this mass scale is $(0.098 \pm 0.004)^2$ and $(0.082 \pm 0.002)^2$ for TNG and FLAMINGO, respectively. Based on older simulations, the value assumed in the ACT analysis was $(0.2)^2$. We anticipate that future SZ survey analyses could benefit from more precise estimates of this variance, particularly when verified by multiple simulations.

4.2 MPR mean behaviours

Over the two orders of magnitude in halo mass we study, the log-mean values of most properties varies by several orders of magnitude (see Appendix E). To enhance mass- and redshift-dependent features, we reduce the dynamic range by presenting in Section 4.2.1 each property's mass-dependent mean relative to the simple self-similar power-law form. We use reference normalizations from KLLR at $10^{15}\,\mathrm{M}_\odot$ and $z = 0$ derived separately for each simulated halo sample.

In Section 4.2.2, we present absolute $z = 0$ normalizations for each simulated sample, finding agreement at the few percent level for thermal energies at the highest masses despite $\sim 15 - 30$ per cent shifts in hot gas mass.

4.2.1 Deviations from self-similarity

For each property S_a, let $\langle S_a(\mu, z)\rangle$ be the exponential of the log-mean MPR derived from KLLR analysis of a given halo sample. We define the $z = 0$ self-similar relation, $S_{a,\mathrm{SS}}(\mu, 0)$, as the SPL scaling that matches the KLLR normalization for each population found at $M_{500\mathrm{c}} = 10^{15}\,\mathrm{M}_\odot$ (because the highest mass halo used in the FLAMINGO analysis is just below $10^{15}\,\mathrm{M}_\odot$, we linearly extrapolate the normalization to $10^{15}\,\mathrm{M}_\odot$), but with self-similar slope given in Section 2.2. We then define deviations from self-similar scaling at any redshift by

$$\Delta_{\mathrm{SS},a}(\mu, z) = \frac{\langle S_a(\mu, z)\rangle}{E^\beta(z)S_{a,\mathrm{SS}}(\mu, z = 0)} \tag{13}$$

where β denotes the self-similar redshift evolution exponent described in Section 2.2. We choose the $10^{15}\,\mathrm{M}_\odot$ reference values at $z = 0$, since the hot gas scalings are tending toward self-similar slopes at the highest masses.

Fig. 6 presents the mass and redshift dependence of $\Delta_{\mathrm{SS},a}(\mu, z)$ for the five gas properties in each simulation (Fig. E1 offers the original KLLR fits). In magnitude, the shifts vary least for T_{sl} (-0.1 to 0.15 dex) and most for L_{X} (-2 to 0 dex) with gas mass and thermal energy being intermediate.

Across a majority of the mass and redshift ranges shown there is striking similarity in the behaviours exhibited by the two simulation samples. This is not surprising given that the same astrophysical ingredients underlie both numerical methods. Star and SMBH formation removes hot gas and heats the remainder. For haloes of low mass, the actions develop over time to both reduce the gas mass and elevate the temperature. Above $10^{14}\,\mathrm{M}_\odot$, reduced efficiency of star and SMBH formation produces a higher gas mass and a progressively lower temperature. These trends are apparent in the slopes shown in Fig. 4.

Above $M_{500\mathrm{c}} > 10^{14}\,\mathrm{M}_\odot$, the M_{gas} mean values are independent of redshift, whereas below this mass scale the gas mass increases toward larger redshifts. At the group scale, the higher binding energy at earlier times allows for more gas to be retained, in agreement with
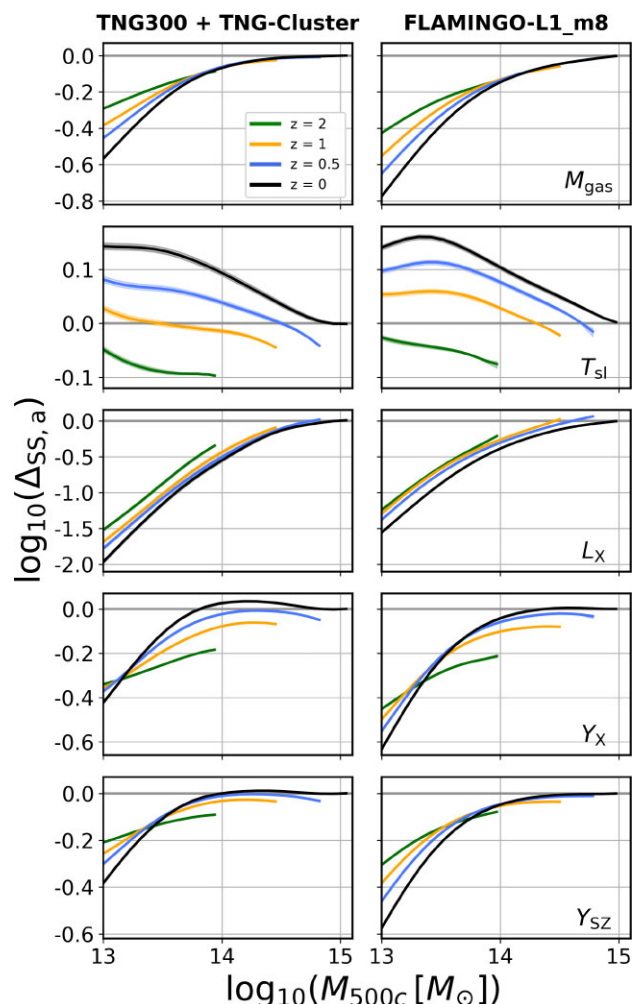

Figure 6. Hot gas property mean values quantified in terms of differences from the self-similar behaviour $\Delta_{\mathrm{SS},a}(\mu, z)$, equation (13). Note the logarithmic shifts are presented in dex. For a non-scaled version, see Fig. E1.

previous works (A. M. C. Le Brun et al. 2017; D. J. Barnes et al. 2017a).

In contrast, T_{sl} evolves at all mass scales such that the ICM plasma in haloes of fixed mass is cooler, by up to 0.2 dex at $z = 2$ compared to the scaled $z = 0$ expectations. While some of this effect may be driven by increased merger frequency and lack of complete virialization at high redshift (A. M. C. Le Brun et al. 2017), much is likely due to the heating effects of AGN feedback over time.

The X-ray luminosity shows large deviations in mass, with low-mass haloes depressed by an order of magnitude relative to self-similar expectations. The redshift evolution, however, adheres much more closely to self-similar expectations, in that the mass-dependent shape is consistently reproduced, especially for TNG haloes at $z \leq 1$.

For both $Y_{\mathrm{X}}-M_{500\mathrm{c}}$ and $Y_{\mathrm{SZ}}-M_{500\mathrm{c}}$ relations, there is a interesting sign change in the evolution at group scales, occurring at slightly different masses for the two simulations. At low masses, the evolution with redshift is positive, with higher thermal energies relative to the self-similar scaled $z = 0$ values, while for larger haloes with $M_{500\mathrm{c}} \gtrsim 10^{13.5}\,\mathrm{M}_\odot$, it becomes negative. Similar behaviour is seen by A. M. C. Le Brun et al. (2017). This sign change is a consequence of the mass-dependent evolution of M_{gas} and the negative evolution of T_{sl} and T_{mw}. At low masses, the positive evolution of M_{gas} more than

Table 4. Normalizations, in decimal exponent format, of the scaling relations at five halo mass scales for TNG and FLAMINGO (FLAM) at $z = 0$. Values at 10^{15} M$_\odot$ are used to set the zero-points for the deviations from self-similarity shown in Fig. 6.

$\log_{10}(M_{500c}/[\,\mathrm{M_\odot}])$	$\log_{10}(M_{gas}/[\,\mathrm{M_\odot}])$		$\log_{10}(T_{sl}/[\mathrm{K}])$		$\log_{10}(L_X/[\mathrm{erg\ s^{-1}}])$		$\log_{10}(Y_X/[\,\mathrm{M_\odot \cdot K}])$		$\log_{10}(Y_{SZ}/[\,\mathrm{M_\odot \cdot K}])$	
	TNG	FLAM	TNG	FLAM	TNG	FLAM	TNG	FLAM	TNG	FLAM
13.0	11.58	11.31	6.65	6.72	40.97	40.96	18.24	18.03	18.29	18.07
13.5	12.40	12.19	6.98	7.07	42.28	42.14	19.38	19.27	19.40	19.27
14	13.08	12.93	7.27	7.35	43.39	43.13	20.36	20.28	20.34	20.27
14.5	13.64	13.54	7.55	7.63	44.31	43.90	21.18	21.17	21.19	21.15
15.0	14.15	14.08	7.84	7.91	44.94	44.52	21.99	21.99	22.01	21.98

compensates for the negative evolution in T_{sl} (or T_{mw} for Y_{SZ}), and at larger masses, M_{gas} evolves self-similarly, while T_{sl} still evolves negatively.

Observational constraints regarding redshift evolution remain mixed and are often limited by small sample sizes, particularly at high redshift. In a joint analysis of 14 cluster samples focusing on high-mass groups and clusters and extending to $z \sim 1.5$, A. Reichert et al. (2011) find no deviation from self-similar evolution in the $M–T$ relation but report a significant deviation for the $M - L_X^{bol}$ relation ($\propto E(z)^{-0.81\pm0.12}$ compared to the SS value of $-7/4$). Additionally, M. Sereno & S. Ettori (2015) observe that the evolution of the $Y_{SZ}–M_{500c}$ relation is compatible with the SS model, while that of $L_X–M_{500c}$ exhibits a steeper evolution than the theoretical prediction. A more recent analysis of a large sample of groups and clusters by I. N. Chiu et al. (2022) reports that the $M_{gas}–M_{500c}$, $L_X–M_{500c}$, and $Y_X–M_{500c}$ relations are consistent with SS evolution, while the $T_X–M_{500c}$ relation showed a shallower than SS behaviour in redshift at $\sim 2\sigma$ significance.

4.2.2 MPR absolute normalizations

Fig. E1 normalizes mean behaviours separately for each simulated halo sample. In Table 4, we present absolute log-mean values of the gas properties at $z = 0$ for TNG and FLAMINGO (black lines in Fig. E1) at five mass scales.

For the lowest mass groups, TNG haloes contain 60 per cent (0.27 dex) more gas than those in FLAMINGO, and this difference decreases with scale to 16 per cent for the largest haloes. In contrast, the spectroscopic-like temperatures in FLAMINGO are approximately 15 per cent higher than TNG independent of halo mass. The thermal energy product, Y_X, thus begins higher in TNG at the group scale but the difference drops to < 0.01 dex at 10^{15} M$_\odot$. The direct thermal energy measure, Y_{SZ}, also begins higher in TNG, by 45 per cent, but the difference at cluster scales declines to 7 per cent at 10^{15} M$_\odot$.

Surprisingly, the X-ray luminosity differences deviate strongly from the pattern seen in M_{gas}. At the group scale, TNG haloes are only 4 per cent more luminous than those of FLAMINGO, suggesting significant structural differences in the hot gas interiors of low-mass haloes as well as differences in metallicity which is shown to be too high in FLAMINGO haloes (J. Braspenning et al. 2024). At the high-mass end, this difference grows dramatically, to a magnitude of 0.42 dex (a factor of 2.6) at 10^{15} M$_\odot$. This upward pull in the TNG mean L_X is driven by a significant increase in positive skewness of the mass-conditioned distribution, $\mathrm{Pr}(L_X|M_{500c})$ (D. Nelson et al. 2024), a topic we discuss further in Section 4.3 below. Excising emission from the core, the inner $0.15R_{500c}$ region, cuts the gap at high halo masses roughly in half, to 0.24 dex, which lies closer to the simple expectation of twice the gap in M_{gas}, or 0.14 dex.

4.3 Conditional-likelihood shapes

Modelling the count of DM haloes as a function of an observable property and utilizing it for cosmological inference requires careful handling of the MORs and selection effects (R. E. Angulo et al. 2012; A. B. Mantz 2019; R. Kugel et al. 2025). Both are sensitive to the form of the mass-conditioned likelihood, $\mathrm{Pr}(S|M, z)$ (e.g. L. D. Shaw, G. P. Holder & J. Dudley 2010; B. M. S. Erickson et al. 2011; M. Costanzi et al. 2019), which is often assumed to be lognormal. Indeed, lognormal covariate behaviour is an implicit assumption of the E14 multiproperty population model.

Although some observations support the lognormality of conditional statistics (e.g. G. W. Pratt et al. 2009; A. Mantz et al. 2010b; N. G. Czakon et al. 2015; A. B. Mantz et al. 2016a), sample size tend to be too small to make sensitive tests. Simulations of large cosmic volumes, however, create samples with substantial statistical power.

Lognormality in halo mass-conditioned deviations was found in both hot gas and total stellar mass within R_{500c} for 22,000 haloes above 10^{13} M$_\odot$ in the MACSIS and BAHAMAS simulations (A. Farahi et al. 2018). Lognormality in total stellar mass was also confirmed using samples of thousands of haloes with $M_{200c} > 10^{13.5}$ M$_\odot$ in the TNG300 and MAGNETICUM simulations (D. Anbajagane et al. 2020). However, that work also identified a form with modest skewness, common to all three simulations sets mentioned above, for the likelihood of satellite galaxy number conditioned on halo mass, with low satellite galaxy number associated with early forming haloes.

In Fig. 7, we present normalized likelihood distributions of residuals in three properties at $z = 0$. Of the three, two adhere closely to log-normality while the third does not. Additional property likelihood forms are presented in Fig. D1. The residuals are integrated across the full mass range of halo mass using equal weight per halo. Each halo's contribution is centred at the KLLR mean and normalized by the KLLR standard deviation at the halo's total mass value. In the presence of curvature in the slope or scatter as a function of μ, and also in the presence of a declining HMF with increasing mass, the KLLR estimates of the first two moments can be biased by small amounts. To minimize this bias, we employ a narrower KLLR kernel width of 0.1 dex to produce these likelihood shapes.

Each panel in Fig. 7 lists the first four moments of the measured residual distribution and dashed lines show a Gaussian reference. Means and standard deviations of the normalized residuals are very close to the expected values of 0 and 1, respectively. The skewness and the kurtosis are the key unconstrained moments shown here that measure deviations from lognormality – our definition of kurtosis is shifted so that a lognormal distribution has a kurtosis of 0.

The residuals in both $\ln M_{gas}$ and $\ln Y_{SZ}$ are well fitted by a Gaussian, particularly in FLAMINGO, while in TNG, there is slightly reduced support at high values. The skewness is therefore negative
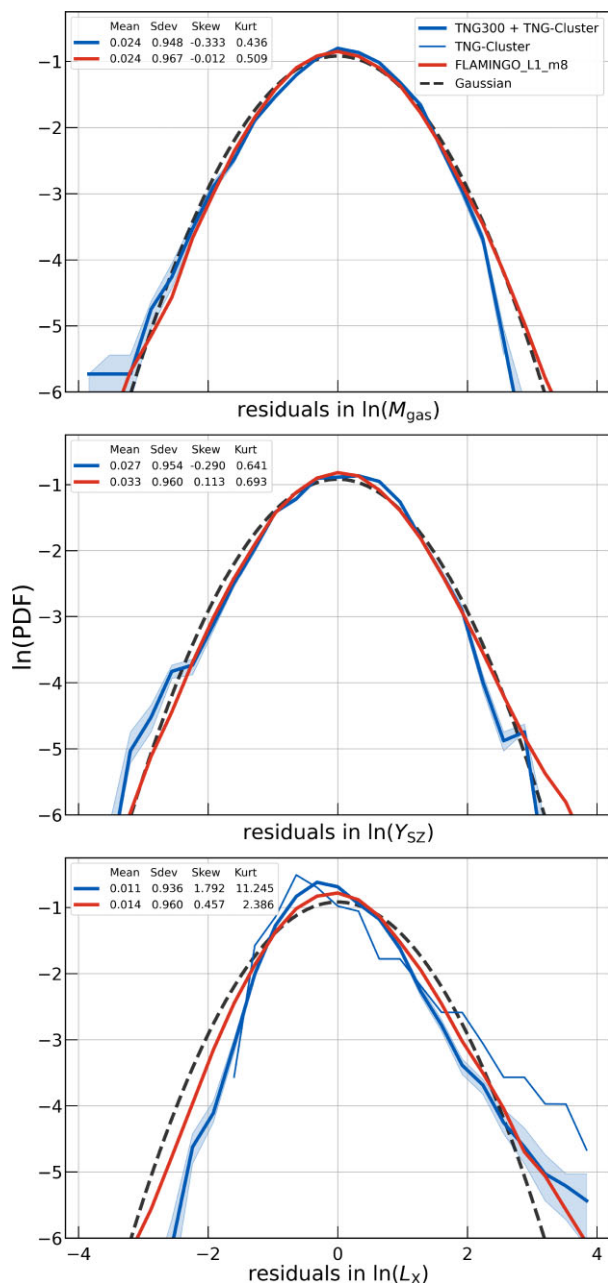
Figure 7. Probability density functions of normalized deviations from mean KLLR behaviour at $z = 0$ formed by marginalizing across halo mass for M_{gas} (top panel), Y_{SZ} (middle), and L_X (bottom). Deviations are expressed in units of the KLLR standard deviation at each halo mass. A KLLR kernel width of 0.1 dex is used to reduce numerical effects on low-order moment estimates, values for which are given in the panel insets.

in TNG with value close to -0.3 in both properties. The kurtosis is slightly larger than zero in both populations.

The choice of radial scale is important in these measures, as R_{500c} is well within the zone where virial and hydrostatic equilibrium are expected to generally hold. For FLAMINGO haloes, R. Kugel et al. (2024) observe a significant tail in the Compton-Y signal at a much larger scale of $5 \times R_{500c}$. The larger aperture reaches out into the infall regime where surrounding massive haloes can be encountered. The extra thermal energy from this two-halo term will boost the positive tail of residuals around the mean. R. Kugel et al. (2024)

show that the significance of this tail decreases when the aperture is reduced to R_{500c}, consistent with our findings.

In contrast to the gas mass and thermal energy, the overall shape of the X-ray luminosity likelihood has substantial skewness and kurtosis. The distributions feature a high-value tail in both TNG and FLAMINGO, the latter previously reported by R. Kugel et al. (2024), with less support at low values. The non-Guassianity is more extreme in TNG haloes. The thin line in the bottom panel of Fig. 7 shows the likelihood shape for the massive TNG-CLUSTER sample only. This subpopulation features a more extreme tail, with 3σ upward fluctuations roughly 10 times more likely than for the full TNG300 sample.

We note in Appendix D that core-excision does not strongly alter these shapes. While the L_X variance in TNG-CLUSTER haloes is significantly reduced by core excision (see Fig. B2) this is not the case for the much more numerous TNG300 haloes. This finding suggests that other factors that affect the phase structure outside the core are responsible for driving the non-normal shape of the X-ray luminosity likelihood.

5 MASS-CONDITIONED PROPERTY CORRELATIONS

Property correlations at fixed halo mass are higher order statistical features of the group and cluster population that can have important consequences. For example, B. Nord et al. (2008) illustrate that, if covariance is not properly taken into account, selection effects can mimic redshift evolution in the $L_X - T$ scaling relation derived from flux-limited samples. For massive clusters, H.-Y. Wu et al. (2015) and A. Farahi et al. (2018) show that total baryon mass serves as a superior mass proxy compared to both stellar mass and gas mass alone, due to strong anticorrelation between the stellar and gas masses at a given halo mass.

Additionally, D. Anbajagane et al. (2020) find an anticorrelation between the number of satellite galaxies and the brightest central galaxy (BCG) stellar mass across multiple simulations, with early forming systems preferring higher BCG mass and fewer satellite galaxies. Relatedly, A. Farahi, M. Ho & H. Trac (2020) find that the total stellar mass–BCG stellar mass and total stellar mass–satellite stellar mass correlations in simulated halo samples become stronger when conditioned on both the magnitude gap – the difference between the fourth and brightest cluster member – in addition to halo mass. The jointly conditioned properties show reduced intrinsic scatter relative to halo mass conditioning alone.

Measuring these correlations can be challenging due to the need for high signal-to-noise measurements, trustworthy estimates of observational errors, and careful modelling of survey selection and other sources of systematic uncertainty. None the less, recent studies have successfully measured the covariance matrix of several X-ray properties (A. Mantz et al. 2010b; A. B. Mantz et al. 2016a, b; S. Andreon et al. 2017; M. Sereno et al. 2020; D. Akino et al. 2022; H. Poon et al. 2023) as well as multiwavelength properties (A. Farahi et al. 2019; S. L. Mulroy et al. 2019). The latter work used 41 clusters from the LoCuSS[6] X-ray-selected sample and analysed up to 11 optical and hot gas properties conditioned on weak lensing mass.

Lines in Fig. 8 present KLLR mass-conditioned correlations (equation A6) for the five principal hot gas properties at redshifts

---

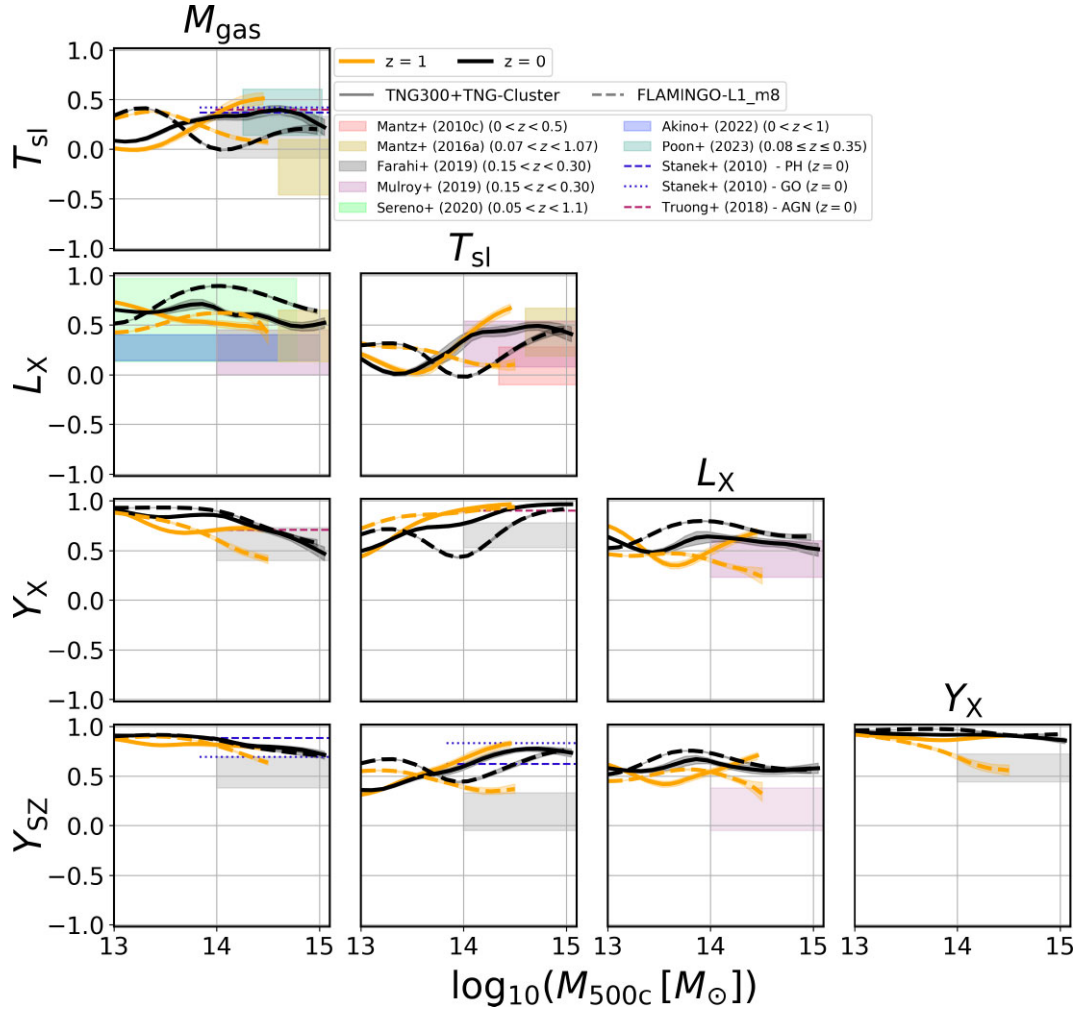[6]Local Cluster Substructure Survey.

Figure 8. KLLR hot gas property correlations at redshifts zero (black) and one (orange) in TNG (solid) and FLAMINGO (dashed) simulations. Observational estimates are shown with shading with vertical extent given by the 1σ reported error and horizontal extent given by the mass range used for the analyses. Correlation measures in past numerical studies, with no errors reported, are given by the dashed and dotted lines, with horizontal extent given by mass ranges used. See the text for discussion.

of zero and one (we omit other redshifts for clarity). Shaded regions show observational constraints, with the caveat that projection effects may drive deviations from the intrinsic, spherical values we measure in the simulations.

Strongly positive correlations exist between $M_{\rm gas}$ and its direct derivative quantities, $L_{\rm X}$, $Y_{\rm SZ}$, and $Y_{\rm X}$. The SZ thermal energy, $Y_{\rm SZ}$, and $M_{\rm gas}$ have correlation coefficients in the range $0.6 - 0.9$, with only weak dependence on halo mass and redshift.

The $Y_{\rm X}$ and $M_{\rm gas}$ correlation has scale-dependence that reflects the mass-dependent scatter of gas mass and temperature. At group scales, the MPR scatter in gas mass dominates that of temperature, which drives the correlation to high values. For cluster-scale haloes above $10^{14}\,{\rm M}_\odot$, the MPR scatter in in $T_{\rm sl}$ is the dominant contributor to $Y_{\rm X}$ scatter. The relatively weak correlation of $T_{\rm sl}$ with $M_{\rm gas}$ thus results in a decline in the $Y_{\rm X} - M_{\rm gas}$ correlation, while the $Y_{\rm X} - T_{\rm sl}$ grows, at these mass scales. While qualitatively similar, there are differences in detail between TNG and FLAMINGO behaviours.

These differences also emerge in the $Y_{\rm X}$ and $Y_{\rm SZ}$ correlation (lower right of the triangle plot) which is large, $r \sim 0.9$ independent of halo mass. A weakening trend is apparent in FLAMINGO haloes at $z = 1$ that mimics the weakening in the $Y_{\rm X} - M_{\rm gas}$ correlation.

The $M_{\rm gas}$–$L_{\rm X}$ correlation is positive and remains above 0.4 at all mass scales and for all redshifts. At $z = 0$, both simulations predict a mild decrease with system mass in the correlation for clusters. At the group scale, the correlation increases from ~ 0.5 to ~ 0.9 in FLAMINGO and from 0.65 to 0.78 in TNG. Both simulation samples show a slight decline in correlation at $z = 1$. The $T_{\rm sl}$–$L_{\rm X}$ correlation is mildly positive across the mass range at all redshifts, with dips toward zero at different halo mass scales. This correlation rises to ~ 0.4 at the cluster scale. These correlation values agree with the observational constraints of A. B. Mantz et al. (2016a) and M. Sereno et al. (2020) and are only slightly larger than the results reported in S. L. Mulroy et al. (2019) and D. Akino et al. (2022).

Within this lower part of the triangle, observational estimates, shown in shaded colours with 1σ uncertainties, are largely in agreement with the simulation values. The LoCuSS sample (A. Farahi et al. 2019; S. L. Mulroy et al. 2019) finds correlation coefficients of $0.33^{+0.21}_{-0.25}$ for $T_{\rm X,ce}$–$L_{\rm X,RASS}$, and $0.13^{+0.20}_{-0.22}$ for $T_{\rm X,ce}$-$M_{\rm gas}$. Using 238 clusters drawn from the ROSAT All-Sky Survey (RASS), A. Mantz et al. (2010b) measure a correlation of 0.09 ± 0.19 between the scatter of the core-excised temperature and the X-ray luminosity in the *ROSAT* broad band. In contrast, using a sample 19
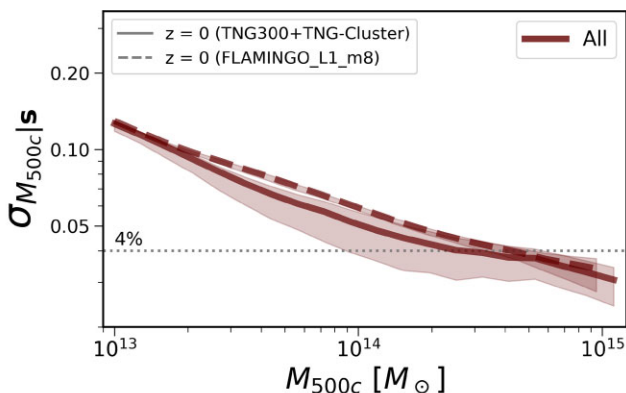
Figure 9. Halo MPQ at $z = 0$ implied by combining the five principal properties listed in Table 1. Because correlations among parameters are mostly positive, the MPQ of the combined properties improves only modestly upon the best single measures of Fig. 1.

X-ray selected clusters, H. Poon et al. (2023) measure a larger value, $0.43^{+0.172}_{-0.293}$, that is consistent within the uncertainties.

Finally, the observed Y_{SZ} correlations show mixed agreement, both internally and with the simulation expectations. The *Planck* measurements, both from the same LoCuSS study (S. L. Mulroy et al. 2019; A. Farahi et al. 2019), are consistent at the 2σ level with the simulation values, but some values are also consistent with zero. Projection effects, particularly in SZ measurements, may be affecting the observational estimates. More accurate estimates of these correlations using larger, more sensitive surveys are clearly needed.

5.1 MPQ from five properties combined

We close this section on correlations by using the full covariance and set of slopes to determine the MPQ achieved by combining all five properties (equation 12), comparing it to scatter in mass estimates produced by novel machine learning (ML) methods in the following section. The result at $z = 0$ is shown in Fig. 9. Strongly negative correlations can greatly reduce the effective halo mass variance while strongly positive correlations are less helpful. This makes qualitative sense, in that the limit of correlation coefficient $r \rightarrow 1$ for any pair of properties implies having a pair of redundant measurements.

There is a remarkable level of consistency in the combined MPQs for TNG and FLAMINGO halo samples, with values dropping gently from 12 per cent for $10^{13}\,M_\odot$ low-mass groups to 3 per cent for $10^{15}\,M_\odot$ high-mass clusters. While improvements in haloes mass scatter may appear relatively modest, the reduction in variance brought about by combining all properties is more substantial, a 43 per cent drop from $(0.04)^2$ to $(0.03)^2$ for high-mass clusters, for example.

5.2 Comparison to machine-learning mass estimates

Recent campaigns for cluster mass estimation have employed various ML models, such as convolutional neural networks (CNNs) and random forests (RFs), trained on X-ray (S. B. Green et al. 2019; M. Ntampaka et al. 2019; M. Ho et al. 2023; S. Krippendorf et al. 2024) and SZ (D. de Andres et al. 2022; D. Wadekar et al. 2023a, b) mock observations from hydrodynamical simulations.

M. Ntampaka et al. (2019) utilized a CNN trained and X-ray mock observations from 329 ILLUSTRISTNG clusters to reproduce cluster

masses with a low bias of -0.02 dex and average intrinsic scatter of 12 per cent. Using a similar architecture, M. Ho et al. (2023) demonstrated that a CNN trained on bolometric X-ray photon maps derived from the MAGNETICUM hydrodynamic simulation achieves a predictive mass scatter of 17.8 per cent, while a CNN trained on a multichannel maps split into low-, medium-, and high-energy bands reduces the scatter to 16.2 per cent. In a study using mock *eROSITA* observations for a cluster sample within a mass range of $10^{13}\,M_\odot < M_{500c} < 10^{15}\,M_\odot$ in the Final Equatorial Depth Survey, S. Krippendorf et al. (2024) reduced the scatter of 45 per cent resulting from the luminosity–mass scaling relations (I. N. Chiu et al. 2022) to 43 per cent by using a CNN augmented with redshift information. They also showed that a CNN can properly handle major contaminators such as AGNs, without a priori filtering. In contrast to CNNs, S. B. Green et al. (2019) used RF regression trained on *Chandra* and *eROSITA* mock X-ray observations of the core-excised luminosity for 1615 clusters from the MAGNETICUM simulations with additional set of morphological parameters, showing a scatter of 15 per cent on predicted masses, a 20 per cent reduction relative to the scatter in the mass–luminosity relation. This level of scatter for the core-excised luminosity is similar to what we find for TNG and FLAMINGO (see Fig. B1). Beyond X-ray images, D. de Andres et al. (2022) used a CNN trained mock *Planck* Compton-y observations from the Three Hundred simulations to estimate the cluster masses, achieving a scatter of 17 per cent.

The levels of scatter in the predicted cluster mass from the ML studies above are large compared to the scatter achieved by the combined set of idealized mass proxies reported in this work. Therefore, we hope that future ML estimates of cluster masses will provide smaller levels of scatter, closer to the value we report here.

6 COMPARISON TO PREVIOUS WORKS

The literature on observed scaling relations of cluster properties is now in its fifth decade (N. A. Bahcall 1974; N. Kaiser 1986; A. E. Evrard & J. P. Henry 1991; M. Markevitch 1998; M. Arnaud & A. E. Evrard 1999; T. A. McKay et al. 2001, 2002; M. C. Novicki, M. Sornig & J. P. Henry 2002; H. J. Eckmiller, D. S. Hudson & T. H. Reiprich 2011; B. J. Maughan et al. 2012; G. Schellenberger et al. 2015; L. Lovisari et al. 2020), and simulation campaigns to theoretically address them began in the 1990s (J. F. Navarro, C. S. Frenk & S. D. M. White 1995; A. E. Evrard et al. 1996; G. L. Bryan & M. L. Norman 1998). In this section, we place our findings into the context of prior studies of MPQ and property scaling relations.

It is important to note that any sample has specific selection criteria that yields varying coverage in mass and redshift. Simulation studies are often complete in a certain volume above some minimum mass, but some approaches, such as zoom-in re-simulations, will trade completeness for uniform coverage across a mass range. In addition, definitions of halo properties, including true mass, can differ by construction. Here, we focus on studies using M_{500c}. Simulation works differ in their physical assumptions for gas evolution, and we include examples of: shock heating under gravity only (GO); preheating (PH), the simple assumption of a rapid rise in entropy at some fixed epoch to mimic galaxy feedback; cooling, star formation and feedback from supernovae (CSF); and SMBH formation and feedback from AGNs. All AGN simulation models also include a CSF treatment.

Observational measures will include projection effects and other real-world features that are often absent from simulation studies.
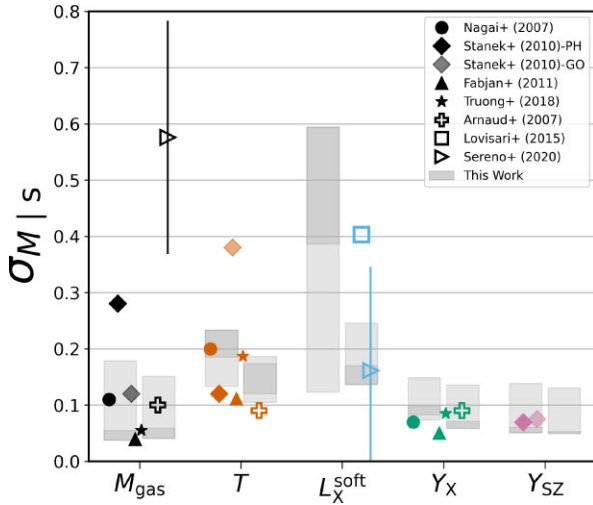
Figure 10. Comparison of MPQ values (the scatter in total halo mass conditioned on each of the listed properties) from prior works (symbols) with those presented here (shaded, left=TNG, and right=FLAMINGO). Dark shading shows the MPQ range for high-mass systems ($M_{500c} \in 10^{14.5} - 10^{15}\,M_{\odot}$), while light shading shows the range of values for all haloes in this study, $10^{13} - 10^{15}\,M_{\odot}$. Filled symbols show values from computational haloes samples at $z = 0$ (typically mass-complete), while open symbols are results from observational cluster samples in the cosmic near-field. For simulations, specific temperature definitions vary (see Table 5).

Caution must be thereby be exercised when making comparisons, particularly between computational and observational samples. We make no attempt to explicitly include redshift evolution; we quote $z = 0$ values for simulations while observed samples have varying depths but tend to be near field, $z < 0.5$. Despite these caveats, the literature review we present here offers insights into the state of our understanding of the astrophysics operating within the massive halo population.

In the figures below, we show values obtained in this study, using dark shading for values found for high-mass haloes, $10^{14.5} - 10^{15}\,M_{\odot}$, and light shading for values across the full halo mass range $10^{13} - 10^{15}\,M_{\odot}$. The former is indicative of the typical median mass scale for the samples in the literature that we include. We note that the lack of error bars in any figures or tables reflects the fact that they were not reported in the original source.

This analysis is not intended to be complete with respect to all published scaling relations for clusters or massive haloes. This exercise simply aims to compare the MPQ and MPR parameters of the current halo samples with those of previous studies. We leave a more complete review of these topics to future work. Indeed, a combined KLLR statistical analysis of all available data from extant simulations and observations would be a first step toward a global 'particle data book' (Particle Data Group et al. 2020) approach to massive halo population statistics.

6.1 Mass proxy quality

Fig. 10 presents MPQ values for the five principal hot gas properties studied in this work. For simulations, the true 3D total mass is used while for observations total masses are derived using hydrostatic X-ray or weak lensing methods. The shaded bands shown for each property (left is TNG, and right is FLAMINGO) provide extremal

values for the MPQs shown in Fig. 1 for high-mass systems ($10^{14.5} - 10^{15}\,M_{\odot}$, dark shading) and for all halo masses (light).

For the most massive simulated haloes at $z = 0$, the finding that hot gas mass, M_{gas}, is a slightly better mass proxy than either thermal energy measures, Y_X or Y_{SZ}, is supported by the studies of D. Fabjan et al. (2011) and N. Truong et al. (2018). The former work employed CSF and a sample of roughly 140 haloes, while the latter used an AGN treatment on 58 resimulated haloes. The midpoints of halo mass ranges in both works lie close to our dark shaded region. Both studies find Y_X to be a slightly worse halo mass proxy compared to M_{gas}.

This finding contradicts the original work of D. Nagai et al. (2007) based on a sample 16 haloes simulated with a CSF approach. That study found Y_X to be superior, with 7 per cent scatter in true mass relative to 11 per cent for gas mass and 20 per cent for gas temperature. While that sample extended into the mass range of groups, no evidence for power-law slope deviation or increased scatter was found, presumably because of the small overall number of systems.

An observational study of ten relaxed clusters by M. Arnaud, E. Pointecouteau & G. W. Pratt (2007) also found that the intrinsic scatter in hydrostatic masses derived from X-ray observations was smallest for Y_X, with slightly larger values for temperature and gas mass. The statistical significance of this comparative result is unclear, and the result reflects values for a small sample selected on criteria that indicate dynamical relaxation.

In simulations, the implied mass scatter for T_{sl} at $z = 0$ ranges between 10 per cent and 20 per cent, with the exception of Millennium Gas simulation haloes under GO treatment of R. Stanek et al. (2010). That work used the same T_{sl} definition used here (and in all other cited studies except D. Nagai et al. 2007, who derived temperatures from fits to synthetic X-ray spectra) and found a mass scatter near 0.4 for the GO halo sample. Under purely gravitational evolution, the internal phase space structure of haloes includes low entropy core gas that would be consumed by star and SMBH formation or heated and ejected by AGN feedback events under more advanced treatments (G. M. Voit 2005). The existence of variable amounts of such high density and low temperature gas in the cores of haloes increases the variance in T_{sl} relative to models for which this gas phase is consumed by compact object formation. Indeed, the preheated (PH) halo sample of R. Stanek et al. (2010), in which a floor is imposed on core entropy, has a much reduced MPQ of 0.12. A value of 0.1 was found for the scatter in hydrostatic mass at fixed X-ray temperature in the small sample of M. Arnaud et al. (2007).

Fewer studies have explicitly reported the scatter in halo mass at fixed thermal energy, but the values shown in Fig. 10 for Y_X and Y_{SZ} lie consistently below 10 per cent. In our study, we find good consistency between MPQ derived from TNG and FLAMINGO samples for Y_{SZ}, with values falling from ~ 15 per cent at the group scale to ~ 5 per cent for high-mass clusters. As previously described, MPQs for Y_X in TNG are higher than those in FLAMINGO in the high-mass cluster regime. For Y_{SZ}, the PH and GO treatments of R. Stanek et al. (2010) produce nearly identical MPQs, and both values lie within the range seem in the full physics simulations used here. This suggests that the MPQ of gas thermal energy may be insensitive to cluster astrophysics, lending further support to the use of SZ selection for cosmological studies.

Because of the sensitivity of X-ray emission measure to internal cluster structure, the MPQ of soft X-ray luminosity is somewhat

Table 5. Intrinsic scatter estimates for hot gas properties studied here. Simulation results are present epoch ($z = 0$) while observational results are for various median redshifts. Our scale-dependent results at $z = 0$ are shown as the black lines in Fig. 5.

Source	$N_{\rm sample}$	$\sigma_{\ln M_{\rm gas}}$	$\sigma_{\ln T}$	$\sigma_{\ln L_X}$	$\sigma_{\ln Y_X}$	$\sigma_{\ln Y_{\rm SZ}}$	Notes
			Simulations				
N. A. Henden et al. (2019)	27	$0.276^{+0.069}_{-0.046}$	–	$0.576^{+0.092}_{-0.069}\,^e$	0.322 ± 0.046	$0.23^{+0.046}_{-0.023}$	CSF + AGN
N. Truong et al. (2018)	58	0.060	0.111^a	0.516^e	0.140	–	CSF + AGN
D. J. Barnes et al. (2017b)	30	0.299	0.138^b	0.691	0.230	–	CSF + AGN
D. J. Barnes et al. (2017a)	1294	0.161 ± 0.023	0.111 ± 0.007^c	$0.345^{+0.023}_{-0.046}\,^f$	0.276 ± 0.023	0.230 ± 0.023	CSF + AGN
S. Planelles et al. (2017)	~ 100	–	–	–	–	0.154	CSF + AGN
A. M. C. Le Brun et al. (2017)	$\sim 10\,000$	0.10	0.10^c	0.25^e	0.10	0.10	CSF + AGN
S. R. Pike et al. (2014)	30	–	0.048 ± 0.005^a	0.283 ± 0.04	–	0.078 ± 0.009	CSF + AGN
S. Planelles et al. (2014)	160	–	0.092^a	–	0.184	0.161	CSF + AGN
V. Biffi et al. (2014)	179	–	0.11^b	0.25^e	–	–	CSF
D. Fabjan et al. (2011)	~ 140	0.042	0.069^d	–	–	0.084	CSF
R. Stanek et al. (2010)	$\sim 10\,000$	0.086 ± 0.001	0.069 ± 0.001^a	0.193 ± 0.002^e	–	0.125 ± 0.002	PH
C. J. Short et al. (2010)	187	–	0.071^a	0.226^e	0.111	–	CSF + AGN
			Observations				
S. Bocquet et al. (SPT 2024)	727	–	–	–	–	0.20 ± 0.05	SZ selection
H. Poon et al. (2023)	19	$0.207^{+0.045}_{-0.033}$	$0.137^{+0.032}_{-0.025}$	$0.219^{+0.062}_{-0.051}\,^\dagger$	$0.412^{+0.100}_{-0.073}$	–	X-ray selection
D. Akino et al. (2022)	136	0.39 ± 0.08	–	$0.73^{+0.12}_{-0.14}$	–	–	X-ray selection
I. N. Chiu et al. (2022)	434	$0.074^{+0.063}_{-0.019}$	$0.069^{+0.061}_{-0.014}$	$0.120^{+0.138}_{-0.060}$	$0.106^{+0.171}_{-0.047}$	–	X-ray selection
M. Sereno et al. (2020)	118	0.253 ± 0.23	–	1.266 ± 0.299	–	–	X-ray selection
I. Chiu et al. (2018)	91	0.11 ± 0.05	–	–	–	–	SZ selection
A. B. Mantz et al. (2016b)	139	–	0.16 ± 0.017	0.43 ± 0.03	0.185 ± 0.016	–	X-ray selected
A. B. Mantz et al. (2016a)	40	0.09 ± 0.02	0.13 ± 0.02	0.24 ± 0.05	–	–	combination
L. Lovisari et al. (2015)	23	–	–	0.564	–	–	X-ray selection
Planck Collab. (2014)	71	–	–	–	–	0.145 ± 0.025	SZ selection
Y.-T. Lin et al. (2012)	94	0.08	–	–	–	–	combination
K. Andersson et al. (2011)	15	–	–	0.253 ± 0.092	–	0.207 ± 0.115	SZ selection
A. Mantz et al. (2010b)	238	–	0.126 ± 0.018	0.414 ± 0.044	–	–	X-ray selection
A. Vikhlinin et al. (2009)	88	–	–	0.396 ± 0.039	–	–	X-ray selection

Notes. ^{a} Spectroscopic-like temperature.
^{b} X-ray spectroscopic temperature.
^{c} Core-excised X-ray spectroscopic temperature.
^{d} Mass-weighted average temperature.
^{e} Bolometric X-ray luminosity.
^{f} Core-excised bolometric X-ray luminosity.

poorer relative to the other measures shown in Fig. 10. In a study of 23 groups of galaxies using hydrostatic mass estimates, L. Lovisari, T. H. Reiprich & G. Schellenberger (2015) find a mass scatter of 40 per cent, noting that hydrostatic mass measurements can introduce significant scatter (J. Braspenning et al. 2024), consistent with the range of values seen in the TNG halo sample but not in FLAMINGO.

A far larger number of studies report scaling relation slope and scatter values from which MPQs could be derived. We quote in this section only studies that explicitly measure variance in total system mass conditioned on an observed property. In the next section, we turn to examining past MPR measures of slope and intrinsic scatter.

6.2 MPR slopes and standard deviations

The majority of sources used in this section are listed in Table 5 where we also give the MPR standard deviation estimate. We will refer to this table in each subsection below. Sample sizes from previous simulations vary from a few dozen haloes to more than 10 000. The

size of the FLAMINGO sample used here is considerably larger than any previously reported.

The observational samples we include vary in size from 15 to 434, with a median value near 70. Many are X-ray selected, while three are SZ-selected samples. The selection is sometimes complex, based on multiple characteristics, so caution must be employed when inter-comparing results.

Once again we emphasize that the values from the numerical literature includes samples with different selected mass ranges. We report simulation slopes at $z = 0$, while for observational studies this is not the case. The TNG and FLAMINGO haloes show little redshift dependence for slopes at cluster mass scales (see Fig. 4), while low-mass groups typically display slopes that decrease with increasing redshift.

Simulations use different codes, physical models, and a variety of fitting methods. Most works fit the scaling relations using an SPL model, while both A. M. C. Le Brun et al. (2017) and A.-R. Pop et al. (2022) employ both SPL and BPL models. The former uses a fixed break scale of $M_{500c} = 10^{14}\,M_\odot$ while the latter fits for this scale. For these sources we report the SPL circle value along with the low- and high-mass slopes of the BPL fits; these are denoted by
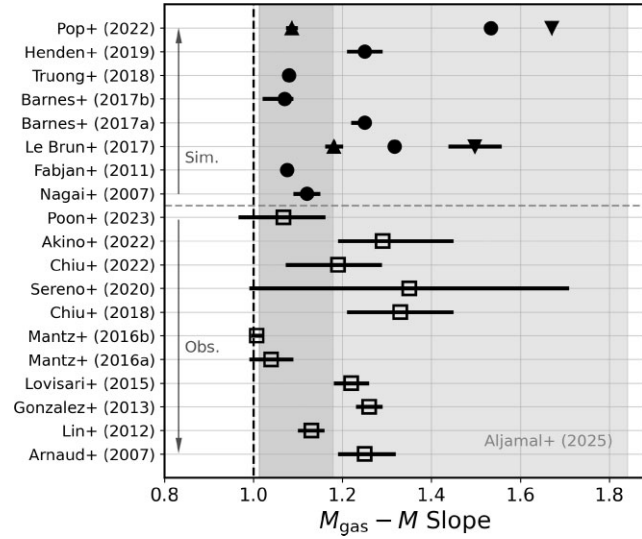
Figure 11. Hot gas scaling relation slopes reported for $z = 0$ simulated halo samples (top) and from local observational cluster samples (bottom). The vertical dashed line represents the self-similar slope of one. The light grey shaded area indicates the extremal range of KLLR slopes from TNG and FLAMINGO samples across the full mass range studied here ($10^{13} - 10^{15}$ M$_\odot$), while the dark shading highlights the slopes observed in the mass range $10^{14.5}$ M$_\odot \leq M_{500c} \leq 10^{15}$ M$_\odot$. For A. M. C. Le Brun et al. (2017) and A.-R. Pop et al. (2022), we report the SPL slopes (circle) as well as BPL slopes for the low (downward-pointing triangle) and high (upward-pointing triangle) haloes.

a circle, downward-pointing triangle, and upward-pointing triangle, respectively, in the figures below.

Subsections that follow present results for individual hot gas properties. Slope estimates are summarized in Figs 11–14. In each figure, the vertical dashed line represents the self-similar expected slope. The light shaded area indicates the range of KLLR slopes observed across the full mass range studied in this work, while the darker grey shade highlights the slope at high masses, $\log_{10}(M_{500c}/\text{M}_\odot) \in 14.5 - 15$. Since TNG and FLAMINGO slopes shown in Fig. 4 are similar, we base these ranges on the minimum and maximum values from the combined set. Note that some works that present slopes but not intrinsic scatter are not listed in Table 5.

6.2.1 Hot gas mass

Fig. 11 shows slopes of the hot gas mass scaling relation reported in prior literature compared to the ranges we find within the high and full mass ranges applying KLLR to TNG and FLAMINGO samples.

All existing evidence supports a scaling of gas mass with total mass that is steeper than self-similar, with magnitudes ranging from 1 to 1.3 The high-mass portion of BPL fits lie in the range we find here, as do most previous simulation works that focus on cluster-scale haloes.

While D. J. Barnes et al. (2017a) report a slope of $1.25^{+0.01}_{-0.03}$ in their full BAHAMAS + MACSIS sample, they find a slope of 1.02 ± 0.03 in their hot sample (clusters with core-excised temperature $k_B T \geq 5$ keV), aligning more closely with the trend of decreasing slope with halo mass scale seen in Fig. 4. The 27 haloes of the FABLE simulation (N. A. Henden et al. 2019) are uniformly selected in halo mass across the two orders of magnitude we study here. That study finds a slope 1.25 with no evidence for steepening at the group scale.

Supporting the trend of steeper M_{gas} slope at the group scale, A. M. C. Le Brun et al. (2017) find a slope 1.50 ± 0.06 for haloes with $M_{500c} \leq 10^{14}$ M$_\odot$. Similarly, A.-R. Pop et al. (2022) find a break in the M_{gas} scaling relation at $M_{500c} \sim 5.7 \times 10^{13}$ M$_\odot$, with slopes of 1.671 ± 0.002 for smaller haloes and $1.086^{+0.014}_{-0.013}$ for larger haloes.

On the observational side, estimates are mostly consistent with a value near 1.2 with the exceptions of the Weighing the Giants (WtG) study of 139 clusters (A. B. Mantz et al. 2016b) and an independent analysis of 40 relaxed, massive clusters (A. B. Mantz et al. 2016a) using X-ray hydrostatic masses. In the WtG study, weak lensing masses were available for only 27 systems. For the remaining 112 clusters, total mass was inferred from gas mass using a fixed gas fraction overdensity model (see equation 1 of A. B. Mantz et al. 2016b). Given the prevalence of this mass definition, a slope of one is to be expected by construction. The focus by A. B. Mantz et al. (2016a) on high-mass, relaxed systems may reflect our finding that slopes approach one as masses approach 10^{15} M$_\odot$.

Reports of intrinsic scatter in hot gas mass are listed in Table 5, seven from simulation campaigns and five from observational samples. The values reported in these studies occupy the full range we find in TNG and FLAMINGO samples, from a high value of 30 per cent at the group scale to a low of 4 per cent for the highest mass clusters. Previous simulations have noted a similar scale dependence to the scatter in hot gas mass (D. J. Barnes et al. 2017a; A. M. C. Le Brun et al. 2017).

For simulations, the high value of 0.30 from D. J. Barnes et al. (2017b) was derived using both gas and total hydrostatic masses from synthetic X-ray observations. A similar value was found by N. A. Henden et al. (2019) for the FABLE simulations that uniformly sample halo masses from group to cluster scales. The studies of D. Fabjan et al. (2011) and N. Truong et al. (2018) find values of 0.04 and 0.06, respectively, close to the minima we find here.

On the observational side, values lie in the range 0.1 to 0.2. The high value of $0.207^{+0.045}_{-0.033}$ found by H. Poon et al. (2023) is based on X-ray hydrostatic masses for 19 X-ray-selected clusters. In that work, a relaxed subset of 9 clusters displays smaller scatter of $0.120^{+0.045}_{-0.028}$. This finding is consistent with the result of A. B. Mantz et al. (2016a), who find a hot gas scatter of 0.09 ± 0.02 for 40 relaxed clusters.

6.2.2 Hot gas temperature

Fig. 12 illustrates a trend among observations and simulations indicating that the slope of the temperature–total mass scaling relation is shallower than predicted by the self-similar model. While some works report slopes that are consistent with 2/3 (D. Nagai et al. 2007; A. Vikhlinin et al. 2009; N. A. Henden et al. 2019), most measurements agree with our results over both the full and reduced mass ranges, with the exception of D. J. Barnes et al. (2017b), who find a much lower slope, $0.47^{+0.07}_{-0.02}$, with large uncertainty.

Our findings in Section 4.1.2 show no clear mass dependence for the slope, which is supported previous works (D. J. Barnes et al. 2017a; A.-R. Pop et al. 2022). The method employed by A.-R. Pop et al. (2022) identify a break scale in the temperature–total mass relation at $0.83^{+4.17}_{-0.24} \times 10^{14}$ M$_\odot$ (note the large error bar), with slopes of $0.445^{0.019}_{-0.075}$ below and $0.599^{+0.273}_{0.032}$ above. In contrast, A. M. C. Le Brun et al. (2017) observe a significant decrease in the slope from 0.64 ± 0.03 for haloes with $M_{500c} \leq 10^{14}$ M$_\odot$ to 0.514 ± 0.009 for more massive haloes.
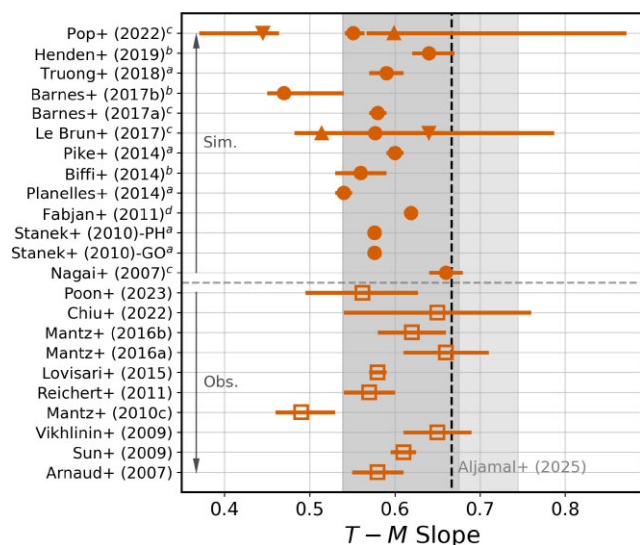
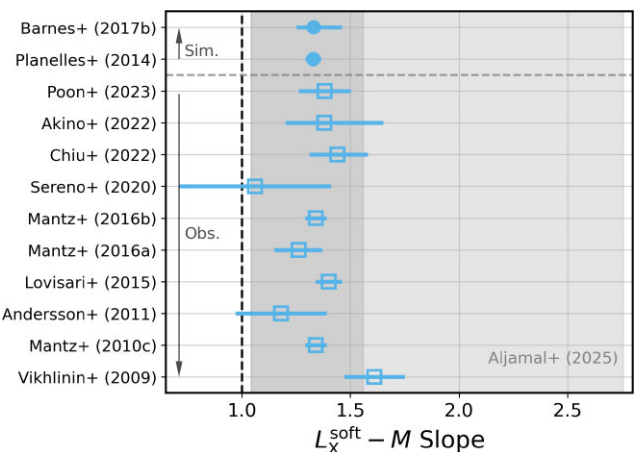**Figure 12.** Temperature–mass scaling relation slopes from simulations and observations. The vertical dashed line represents the slope of 2/3 predicted by the self-similar model. Superscripts for simulations identify the specific temperature definition used in each study; key is given in Table 5. Format is the same as that used in Fig. 11.

For intrinsic scatter, our results in Fig. 5 show that the intrinsic scatter in T_{sl} varies only moderately with mass at $z = 0$, with values ranging from 0.06 to 0.15. Simulation values listed in Table 5 are mainly consistent within this range. Remarkably, multiple observational studies find consistency with a value of approximately 13 per cent.

6.2.3 Soft X-ray luminosity

Fig. 13 presents four observational and two numerical measurements of the slope for the soft-band luminosity–total mass scaling relation. We found only a small number of simulation studies reporting the scaling for soft-band X-ray luminosity. Both works find values near 1.3, but note that we inferred the 1.33 ± 0.05 value from S. Planelles et al. (2014) by combining their reported slopes for the luminosity–temperature (2.46 ± 0.05) and temperature–total mass (0.54 ± 0.01) relations. Seven other simulations, listed in Table 5, employed bolometric X-ray luminosity. We discuss their estimates of intrinsic scatter below, as this feature is less dependent on passband than is the scaling relation slope.

Observational measurements also find steeper than self-similar slopes, with values that agree with those found in our study using the high-mass range of $10^{14.5} - 10^{15}\,M_\odot$. All studies are consistent with a central value of 1.3, though the work of A. Vikhlinin et al. (2009) lies in 2σ tension with this value.

The intrinsic scatter estimates are generally consistent with a range of 0.2 to 0.5, with come simulation estimates lying even higher. The presence of a cool core has long been known to increase variance in the $L_X - T$ relation (A. C. Fabian et al. 1994) yet D. J. Barnes et al. (2017a) still find a scatter of $0.345^{+0.023}_{-0.046}$ for core-excised L_X in the MACSIS halo sample.

6.2.4 Hot gas thermal energy

The findings that hot gas mass scaling is steeper than self-similar, while gas temperature is shallower, combine to result in thermal energy scaling relations having slopes that, for high-mass clusters,

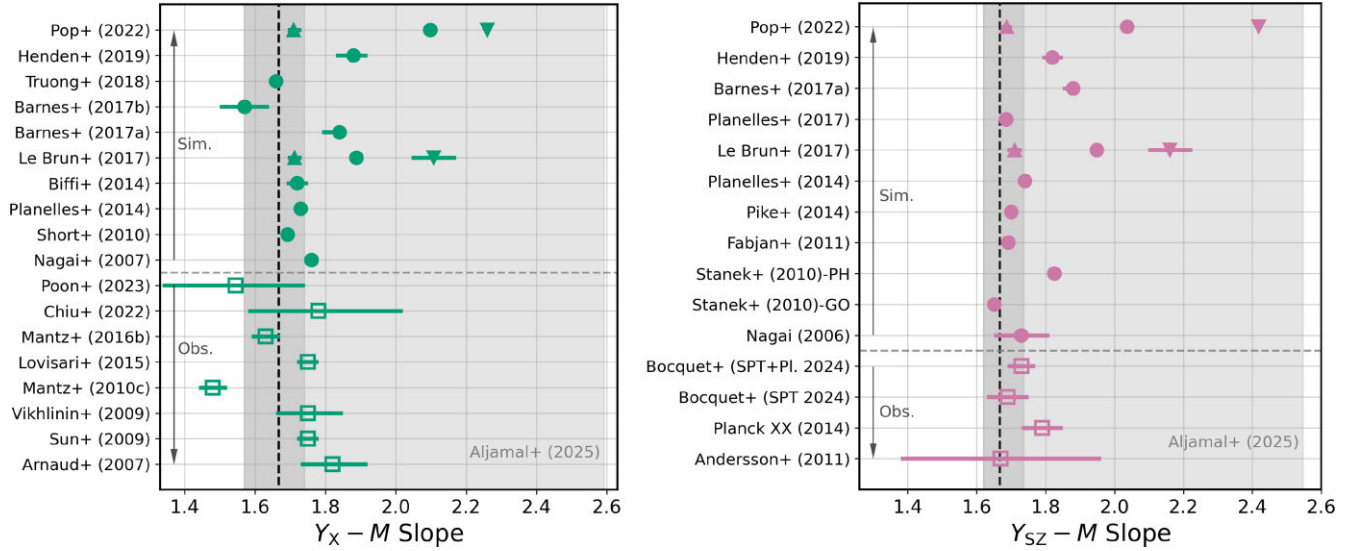
Figure 14. Hot gas thermal energy–mass scaling relation slopes from simulations and observations for Y_X (left) and Y_{SZ} (right). Vertical dashed lines represent the self-similar slope of 5/3. Format is the same as that used in Fig. 11.

properties using halo populations produced by two large-volume cosmological simulations. These simulations, containing thousands of haloes in the combined TNG300 and TNG-CLUSTER samples and close to a hundred thousand in FLAMINGO (L1_m8), enable detailed examination of the mass and redshift dependence of the mass–property scaling relations, the slope and intrinsic scatter of which set the limiting precision on inferred total mass. We focus on five aggregate gas properties: the hot gas mass M_{gas}, the spectroscopic-like temperature T_{sl}, the soft $[0.5 - 2.0]$ keV X-ray luminosity L_X, the X-ray derived thermal energy Y_X, and its thermal SZ equivalent, Y_{SZ}.

Using KLLR with a kernel resolving 0.2 dex in halo mass, we extract mass-dependent log-means, slopes and covariance of the above properties across $10^{13} \lesssim M_{500c}/M_\odot \lesssim 10^{15}$ at four redshifts, $z \in 0, 0.5, 1, 2$. We compute MPQ as the ratio of property's local scatter divided by its slope, which illuminates the *precision* of that property's use as a total mass proxy. The *accuracy* of each property involves understanding the absolute normalization of the scaling relation, a more difficult task (see item vii below).

Our main results are as follows:

(i) At all redshifts, slopes and standard deviations of hot gas properties are sensitive to total halo mass, leading to scale-dependent MPQ values displayed in Fig. 1. As halo mass increases, intrinsic property variance decreases while slopes converge toward self-similar expectations. These features, common to both simulation samples, reflect the decreasing importance of star formation and AGN feedback in shaping hot gas properties in the deepest gravitational potentials.

(ii) At $z = 0$, the thermal energy measures, Y_X and Y_{SZ}, as well as hot gas mass, M_{gas}, compete for best halo mass proxy. Property-conditioned mass scatter decreases from ~ 18 per cent at $M_{500c} = 10^{13} M_\odot$ to $4 - 5$ per cent at $M_{500c} = 10^{15} M_\odot$. With the exception of Y_X in TNG at high masses, the MPQ values and mass dependence are nearly identical similar in both simulation samples.

(iii) At cluster scales, $M_{500c} > 10^{14} M_\odot$, the slope and intrinsic scatter in M_{gas} and Y_{SZ} are nearly redshift independent, and log-mean normalizations of these properties scale self-similarly with redshift.

These properties, observable via their X-ray and millimetre-wave observable signatures, are thus ideal proxies for identifying massive haloes and characterizing the population in the near cosmic field.

(iv) Residuals about scale-dependent means in M_{gas} and Y_{SZ} are very consistent with a lognormal form. The MPQ estimate of the E14 model thus agrees well with directly measured mass variance for these properties.

(v) For galaxy groups, $M_{500c} < 10^{14} M_\odot$, the slopes and normalizations of all properties are strongly redshift dependent in a manner that, at fixed total mass, reflects a progressive heating and loss of hot gas over time. The scale-dependent scatter of gas properties is nearly redshift independent in TNG while FLAMINGO haloes show significantly lower scatter at high redshift for M_{gas} and Y_{SZ}.

(vi) Correlations among hot gas properties are generally positive, with some (e.g. $L_X - T_{sl}$) exhibiting complex scale-dependent behaviour. Combining the entire set of properties yields a minimum halo mass scatter of 3 per cent for high-mass clusters.

(vii) Mass-dependent means of hot gas properties at $z = 0$ can differ significantly between the two astrophysical treatments, challenging the role of such simulations in setting absolute expectations for observed group and cluster properties. However, good agreement is found for gas thermal energy, Y_{SZ}, for high-mass clusters. Respect for virial equilibrium may be at play, as gas masses and temperatures in the two simulations differ in directions that cancel when combined. Cosmological analysis using thermal SZ cluster samples may benefit from these findings by employing more informative, simulation-based priors on astrophysical scaling relation priors.

These theoretical expectations should be verified by other hydrodynamical simulations and their robustness to different galaxy formation and feedback models examined. We have verified that the results summarized above are robust to numerical resolution, but similar analysis on higher resolution simulations with different gas physics and AGN feedback models remains imperative. Scaling relation mean values are a particular area deserving

more careful attention in both simulated and observed cluster samples.

Validation studies should also be pursued. Extraction of population-level statistical features from multiwavelength observational studies is an essential opportunity for samples of galaxy groups and clusters already in hand, as well as those to be defined by *eROSITA*, *Euclid*, LSST[7], the Simons Observatory, and other next-generation facilities. As group and cluster samples with uniform multiwavelength coverage grow to encompass thousands of systems, scale-sensitive studies of *observable-conditioned* properties offer key pathways to precisely reveal the baryonic components of massive haloes and, by comparison to simulated group and cluster population expectations, to understand their astrophysical evolution.

DATA AVAILABILITY

The halo catalogues used in this work, the code used to analyse it, and KLLR analysis data, are available on EA's github: https://github.com/ealjamal/MPQScaling.

REFERENCES

Abbott T. M. C. et al., 2020, Phys. Rev. D, 102, 023509
Abdullah M. H., Klypin A., Wilson G., 2020, ApJ, 901, 90
Abril-Pla O. et al., 2023, PeerJ Comput. Sci., 9, e1516
Aguena M. et al., 2021, MNRAS, 502, 4435
Akino D. et al., 2022, PASJ, 74, 175
Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
Anbajagane D., Evrard A. E., Farahi A., Barnes D. J., Dolag K., McCarthy I. G., Nelson D., Pillepich A., 2020, MNRAS, 495, 686
Anbajagane D., Evrard A. E., Farahi A., 2022a, MNRAS, 509, 3441
Anbajagane D. et al., 2022b, MNRAS, 510, 2980
Andersson K. et al., 2011, ApJ, 738, 48
Andreon S., Wang J., Trinchieri G., Moretti A., Serra A. L., 2017, A&A, 606, A24

Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046
Arnaud M., Evrard A. E., 1999, MNRAS, 305, 631
Arnaud M., Pointecouteau E., Pratt G. W., 2007, A&A, 474, L37
Ayromlou M., Nelson D., Pillepich A., 2023, MNRAS, 524, 5391
Bahcall N. A., 1974, ApJ, 193, 529
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Barnes D. J., Kay S. T., Henson M. A., McCarthy I. G., Schaye J., Jenkins A., 2017a, MNRAS, 465, 213
Barnes D. J. et al., 2017b, MNRAS, 471, 1088
Battaglia N., Bond J. R., Pfrommer C., Sievers J. L., 2013, ApJ, 777, 123
Bhattacharya S., Di Matteo T., Kosowsky A., 2008, MNRAS, 389, 34
Biffi V., Sembolini F., De Petris M., Valdarnini R., Yepes G., Gottlöber S., 2014, MNRAS, 439, 588
Bocquet S. et al., 2019, ApJ, 878, 55
Bocquet S. et al., 2024, Phys. Rev. D, 110, 083510
Böhringer H. et al., 2007, A&A, 469, 363
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Borrow J., Borrisov A., 2020, J. Open Source Softw., 5, 2430
Borrow J., Schaller M., Bower R. G., Schaye J., 2022, MNRAS, 511, 2367
Braspenning J. et al., 2024, MNRAS, 533, 2656
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Chatzigiannakis D., Pillepich A., Simionescu A., Truong N., Nelson D., 2025, preprint (arXiv:2503.01983)
Chiu I. et al., 2018, MNRAS, 478, 3072
Chiu I. N. et al., 2022, A&A, 661, A11
Chiu I. N., Klein M., Mohr J., Bocquet S., 2023, MNRAS, 522, 1601
Cohn J. D., Battaglia N., 2020, MNRAS, 491, 1575
Collette A., 2013, Python and HDF5. O'Reilly, Sebastopol, CA
Costanzi M. et al., 2019, MNRAS, 488, 4779
Czakon N. G. et al., 2015, ApJ, 806, 18
da Silva A. C., Kay S. T., Liddle A. R., Thomas P. A., 2004, MNRAS, 348, 1401
de Andres D. et al., 2022, Nat. Astron., 6, 1325
de Haan T. et al., 2016, ApJ, 832, 95
Driver S. P. et al., 2022, MNRAS, 513, 439
Eckmiller H. J., Hudson D. S., Reiprich T. H., 2011, A&A, 535, A105
Elbers W., Frenk C. S., Jenkins A., Li B., Pascoli S., 2022, MNRAS, 516, 3821
Erickson B. M. S., Cunha C. E., Evrard A. E., 2011, Phys. Rev. D, 84, 103506
Evrard A. E., Henry J. P., 1991, ApJ, 383, 95
Evrard A. E., Metzler C. A., Navarro J. F., 1996, ApJ, 469, 494
Evrard A. E., Arnault P., Huterer D., Farahi A., 2014, MNRAS, 441, 3562 (E14)
Fabian A. C., Crawford C. S., Edge A. C., Mushotzky R. F., 1994, MNRAS, 267, 779
Fabjan D., Borgani S., Tornatore L., Saro A., Murante G., Dolag K., 2010, MNRAS, 401, 1670
Fabjan D., Borgani S., Rasia E., Bonafede A., Dolag K., Murante G., Tornatore L., 2011, MNRAS, 416, 801
Farahi A., Evrard A. E., McCarthy I., Barnes D. J., Kay S. T., 2018, MNRAS, 478, 2618
Farahi A. et al., 2019, Nat. Commun., 10, 2504
Farahi A., Ho M., Trac H., 2020, MNRAS, 493, 1361
Farahi A., Anbajagane D., Evrard A. E., 2022, ApJ, 931, 166
Ferland G. J. et al., 2017, Rev. Mex. Astron. Astrofis., 53, 385
Forouhar Moreno V. J., Helly J., McGibbon R., Schaye J., Schaller M., Han J., Kugel R., 2025, MNRAS, 543, 1339
Gaspari M., Brighenti F., Temi P., Ettori S., 2014, ApJ, 783, L10
Giodini S., Lovisari L., Pointecouteau E., Ettori S., Reiprich T. H., Hoekstra H., 2013, Space Sci. Rev., 177, 247
Gladders M. D., Yee H. K. C., Majumdar S., Barrientos L. F., Hoekstra H., Hall P. B., Infante L., 2007, ApJ, 655, 128
Gonzalez A. H. et al., 2019, ApJS, 240, 33
Gough B., 2009, GNU Scientific Library Reference Manual. Network Theory Ltd., Surrey, UK.
Green S. B., Ntampaka M., Nagai D., Lovisari L., Dolag K., Eckert D., ZuHone J. A., 2019, ApJ, 884, 33

[7]Legacy Survey of Space and Time.

Gupta N., Saro A., Mohr J. J., Dolag K., Liu J., 2017, MNRAS, 469, 3069
Hahn O., Rampf C., Uhlemann C., 2021, MNRAS, 503, 426
Harris C. R. et al., 2020, Nature, 585, 357
Henden N. A., Puchwein E., Sijacki D., 2019, MNRAS, 489, 2439
Ho M., Ntampaka M., Rau M. M., Chen M., Lansberry A., Ruehle F., Trac H., 2022, Nat. Astron., 6, 936
Ho M., Rau M. M., Ntampaka M., Farahi A., Trac H., Póczos B., 2019, ApJ, 887, 25
Ho M., Soltis J., Farahi A., Nagai D., Evrard A., Ntampaka M., 2023, MNRAS, 524, 3289
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Ider Chitham J. et al., 2020, MNRAS, 499, 4768
Jansen F. et al., 2001, A&A, 365, L1
Kaiser N., 1986, MNRAS, 222, 323
Kay S. T., Braspenning J., Chluba J., Helly J. C., Kugel R., Schaller M., Schaye J., 2024, MNRAS, 534, 251
Kodi Ramanah D., Wojtak R., Ansari Z., Gall C., Hjorth J., 2020, MNRAS, 499, 1985
Kodi Ramanah D., Wojtak R., Arendse N., 2021, MNRAS, 501, 4080
Kravtsov A. V., Borgani S., 2012, ARA&A, 50, 353
Kravtsov A. V., Vikhlinin A., Nagai D., 2006, ApJ, 650, 128
Krippendorf S. et al., 2024, A&A, 682, A132
Kugel R. et al., 2023, MNRAS, 526, 6103
Kugel R., Schaye J., Schaller M., McCarthy I. G., Braspenning J., Helly J. C., Forouhar Moreno V. J., McGibbon R. J., 2024, MNRAS, 534, 2378
Kugel R., Schaye J., Schaller M., Forouhar Moreno V. J., McGibbon R. J., 2025, MNRAS, 537, 2179
Lau E. T., Kravtsov A. V., Nagai D., 2009, ApJ, 705, 1129
Le Brun A. M. C., McCarthy I. G., Schaye J., Ponman T. J., 2017, MNRAS, 466, 4442
Lee E. et al., 2022, MNRAS, 517, 5303
Lehle K., Nelson D., Pillepich A., Truong N., Rohr E., 2024, A&A, 687, A129
Lin Y.-T., Stanford S. A., Eisenhardt P. R. M., Vikhlinin A., Maughan B. J., Kravtsov A., 2012, ApJ, 745, L3
Lovisari L. et al., 2020, ApJ, 892, 102
Lovisari L., Maughan B. J., 2022, in Bambi C., Sangangelo A., eds, Handbook of X-ray and Gamma-ray Astrophysics. Springer Nature, Singapore, p. 65
Lovisari L., Reiprich T. H., Schellenberger G., 2015, A&A, 573, A118
Mantz A. B., 2019, MNRAS, 485, 4863
Mantz A. B. et al., 2015, MNRAS, 446, 2205
Mantz A. B., Allen S. W., Morris R. G., Schmidt R. W., 2016a, MNRAS, 456, 4020
Mantz A. B. et al., 2016b, MNRAS, 463, 3582
Mantz A. B. et al., 2022, MNRAS, 510, 131
Mantz A., Allen S. W., Rapetti D., Ebeling H., 2010a, MNRAS, 406, 1759
Mantz A., Allen S. W., Ebeling H., Rapetti D., Drlica-Wagner A., 2010b, MNRAS, 406, 1773
Marinacci F. et al., 2018, MNRAS, 480, 5113
Markevitch M., 1998, ApJ, 504, 27
Maturi M. et al., 2023, A&A, 678, A145
Maughan B. J., Giles P. A., Randall S. W., Jones C., Forman W. R., 2012, MNRAS, 421, 1583
Mazzotta P., Rasia E., Borgani S., Moscardini L., Dolag K., Tormen G., 2004a, preprint(astro-ph/0412536)
Mazzotta P., Rasia E., Moscardini L., Tormen G., 2004b, MNRAS, 354, 10
McCarthy I. G. et al., 2010, MNRAS, 406, 822
McCarthy I. G., Schaye J., Bower R. G., Ponman T. J., Booth C. M., Dalla Vecchia C., Springel V., 2011, MNRAS, 412, 1965
McGibbon R., Helly J. C., Schaye J., Schaller M., Vandenbroucke B., 2025, J. Open Source Softw., 10, 8252
McKay T. A. et al., 2001, preprint(astro-ph/0108013)
McKay T. A. et al., 2002, ApJ, 571, L85
Mckinney W., 2011, Python High Performance Science Computer.
Mehrtens N. et al., 2012, MNRAS, 423, 1024
Miyatake H. et al., 2022, Phys. Rev. D, 106, 083520
Mulroy S. L. et al., 2019, MNRAS, 484, 60
Nagai D., 2006, ApJ, 650, 538
Nagai D., Kravtsov A. V., Vikhlinin A., 2007, ApJ, 668, 1
Naiman J. P. et al., 2018, MNRAS, 477, 1206
Navarro J. F., Frenk C. S., White S. D. M., 1995, MNRAS, 275, 720
Nelson D. et al., 2018, MNRAS, 475, 624
Nelson D. et al., 2019, MNRAS, 490, 3234
Nelson D., Pillepich A., Ayromlou M., Lee W., Lehle K., Rohr E., Truong N., 2024, A&A, 686, A157
Nord B., Stanek R., Rasia E., Evrard A. E., 2008, MNRAS, 383, L10
Norton C. E., Adams F. C., Evrard A. E., 2024, MNRAS, 531, 1685
Novicki M. C., Sornig M., Henry J. P., 2002, AJ, 124, 2413
Ntampaka M., Trac H., Sutherland D. J., Fromenteau S., Póczos B., Schneider J., 2016, ApJ, 831, 135
Ntampaka M. et al., 2019, ApJ, 876, 82
Ostriker J. P., Peebles P. J. E., Yahil A., 1974, ApJ, 193, L1
Pacaud F. et al., 2018, A&A, 620, A10
Particle Data Group et al., 2020, Prog. Theor. Exp. Phys., 2020, 083C01
Pedregosa F. et al., 2011, J. Mach. Learn. Res., 12, 2825
Pierre M. et al., 2016, A&A, 592, A1
Pike S. R., Kay S. T., Newton R. D. A., Thomas P. A., Jenkins A., 2014, MNRAS, 445, 1774
Pillepich A. et al., 2018a, MNRAS, 473, 4077
Pillepich A. et al., 2018b, MNRAS, 475, 648
Planck Collaboration XX, 2014, A&A, 571, A20
Planck Collaboration XXIV, 2016, A&A, 594, A24
Planelles S., Borgani S., Fabjan D., Killedar M., Murante G., Granato G. L., Ragone-Figueroa C., Dolag K., 2014, MNRAS, 438, 195
Planelles S. et al., 2017, MNRAS, 467, 3827
Poon H., Okabe N., Fukazawa Y., Akino D., Yang C., 2023, MNRAS, 520, 6001
Pop A.-R. et al., 2022, preprint (arXiv:2205.11528)
Pratt G. W., Croston J. H., Arnaud M., Böhringer H., 2009, A&A, 498, 361
Puchwein E., Sijacki D., Springel V., 2008, ApJ, 687, L53
Rasia E., Mazzotta P., Borgani S., Moscardini L., Dolag K., Tormen G., Diaferio A., Murante G., 2005, ApJ, 618, L1
Rasia E. et al., 2014, ApJ, 791, 96
Reichert A., Böhringer H., Fassbender R., Mühlegger M., 2011, A&A, 535, A4
Rohr E., Pillepich A., Nelson D., Ayromlou M., Zinger E., 2024, A&A, 686, A86
Rozo E. et al., 2010, ApJ, 708, 645
Rykoff E. S. et al., 2014, ApJ, 785, 104
Schaller M. et al., 2024, MNRAS, 530, 2378
Schaye J. et al., 2023, MNRAS, 526, 4978
Schellenberger G., Reiprich T. H., Lovisari L., Nevalainen J., David L., 2015, A&A, 575, A30
Sehgal N. et al., 2011, ApJ, 732, 44
Sereno M., Ettori S., 2015, MNRAS, 450, 3675
Sereno M. et al., 2020, MNRAS, 492, 4528
Shaw L. D., Holder G. P., Bode P., 2008, ApJ, 686, 206
Shaw L. D., Holder G. P., Dudley J., 2010, ApJ, 716, 281
Short C. J., Thomas P. A., Young O. E., Pearce F. R., Jenkins A., Muanwong O., 2010, MNRAS, 408, 2213
Silk J., Mamon G. A., 2012, Res. Astron. Astrophys., 12, 917
Springel V. et al., 2018, MNRAS, 475, 676
Staffehl M., Nelson D., Ayromlou M., Rohr E., Pillepich A., 2025, preprint (arXiv:2503.01960)
Stanek R., Rasia E., Evrard A. E., Pearce F., Gazzola L., 2010, ApJ, 715, 1508
Sunyaev R. A., Zeldovich Y. B., 1972, Comment. Astrophys. Space Phys., 4, 173
Truong N. et al., 2018, MNRAS, 474, 4089
Vikhlinin A., 2006, ApJ, 640, 710
Vikhlinin A. et al., 2009, ApJ, 692, 1033
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Voit G. M., 2005, Rev. Mod. Phys., 77, 207
Wadekar D. et al., 2023a, Proc. Natl. Acad. Sci., 120, e2202074120
Wadekar D. et al., 2023b, MNRAS, 522, 2628

Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, Phys. Rep., 530, 87

Weinberger R. et al., 2017, MNRAS, 465, 3291

Weisskopf M. C., Tananbaum H. D., Van Speybroeck L. P., O'Dell S. L., 2000, in Truemper J. E., Aschenbach B., eds, Proc. SPIE Conf. Ser. Vol. 4012, X-Ray Optics, Instruments, and Missions III. SPIE, Bellingham, p. 2

Wen Z. L., Han J. L., 2022, MNRAS, 513, 3946

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

White S. D. M., Efstathiou G., Frenk C. S., 1993, MNRAS, 262, 1023

Wu H.-Y., Evrard A. E., Hahn O., Martizzi D., Teyssier R., Wechsler R. H., 2015, MNRAS, 452, 1982

Yan Z., Mead A. J., Van Waerbeke L., Hinshaw G., McCarthy I. G., 2020, MNRAS, 499, 3445

APPENDIX A: KERNEL LOCALIZED LINEAR REGRESSION

Simple linear regression using the least-squares algorithm has been a staple approach to cluster scaling relations, providing a normalization, slope, and variance defining the relationship between two integral log-space properties. However, such an approach is too inflexible as it attempts to summarize a population of thousands of haloes by only the three aforementioned quantities. Furthermore, as mentioned in Section 2.4, the steeply falling nature of the HMF means that parameters recovered via linear regression are strongly biased by the behaviour of low-mass haloes. KLLR (A. Farahi et al. 2022) allows for a more sensitive analysis of the scaling relations by conditioning the normalization, slope, and variance on halo mass.

A1 Scaling relations using KLLR

We follow the treatment in E14 and A. Farahi et al. (2018) by considering a set of halo properties, S, associated with haloes with total mass M_{500c}. Under the assumption of lognormal distribution around the scaling relation, we define $s = \ln S/S_{\text{fid}}$ and $\mu = \ln(M_{500c}/M_{\text{fid}})$ where S_{fid} and M_{fid} are chosen fiducial scales. We write the mean scaling relation of property a at a fixed redshift z as:

$$\langle s_a \mid \mu, z \rangle = \pi_a(\mu, z) + \alpha_a(\mu, z)\mu, \tag{A1}$$

where $\pi_a(\mu, z)$ and $\alpha_a(\mu, z)$ are the mass- and redshift-dependent normalization and slope for the property of interest, respectively. At a given redshift, we estimate the local normalization and slope around some chosen mass scale μ_c by minimizing the weighted residual sum of squares :

$$\epsilon_a^2(\mu_c) = \sum_{i=1}^{n} w^2(\mu_i - \mu_c)[s_{a,i} - \pi_a(\mu_c) - \alpha_a(\mu_c)(\mu_i - \mu_c)]^2, \tag{A2}$$

where the sum i is over all haloes. This procedure is carried out for 20 halo mass centres, μ_c, equally spaced in μ while recording the normalization, slope, and intrinsic variance for each. Weights are Gaussian,

$$w(\mu_i - \mu_c) = \frac{1}{\sqrt{2\pi}\sigma_{\text{KLLR}}} \exp\left[-\frac{(\mu_i - \mu_c)^2}{2\sigma_{\text{KLLR}}^2} \right], \tag{A3}$$

where we choose $\sigma_{\text{KLLR}} = 0.2 \ln(10)$. The kernel width controls the size of the local features that the algorithm tracks, but like any hyperparameter it also controls the degree of overfitting to the noise in the data. Thus, a small kernel width would be ideal for more local analysis of the scaling behaviour but that would result in a noisy fit, and if the kernel width is too large, any local features would be washed out and we would approach a simple linear regression fit.

A2 Scale-dependent scatter and covariance

To calculate the mass scatter implied by a single property, we only need to use the slope and intrinsic scatter for that property, but to calculate the mass scatter implied by multiple properties, we also need to calculate the pairwise covariance between the properties (Section 5). Using the local slope and normalization, we can calculate the covariance between s_a and s_b by using the unbiased estimator (B. Gough 2009):

$$C_{a,b} = A \sum_{i=1}^{n} w_i \delta s_{a,i} \delta s_{b,i}, \tag{A4}$$

where the residual for the ith data point is given by $\delta s_{a,i} = s_{a,i} - \alpha_a \mu_i - \pi_a$, the pre-factor is

$$A = \frac{\sum_{i=1}^{n} w_i}{\left(\sum_{i=1}^{n} w_i \right)^2 - \sum_{i=1}^{n} w_i^2}, \tag{A5}$$

and the correlation coefficient normalizes this covariance to be in the range $[-1, 1]$

$$r_{a,b} = \frac{C_{a,b}}{\sqrt{C_{aa}}\sqrt{C_{bb}}}. \tag{A6}$$

The scale-dependent scatter is given by the square root of the diagonal entries of the covariance:

$$\sigma_a = \sqrt{C_{a,a}} = \sqrt{A \sum_{i=1}^{n} w_i \delta s_{a,i} \delta s_{a,i}}. \tag{A7}$$

For the mass-conditioned residual kernels, we focus on the *normalized residuals*,

$$\widehat{\delta s}_{a,i} \equiv \frac{\delta s_{a,i}}{\sigma_a} = \frac{s_{a,i} - \alpha_a \mu_i - \pi_a}{\sigma_a}. \tag{A8}$$

As discussed in the text, we employ a narrower kernel width of 0.1 dex to compute the distribution of residuals.

APPENDIX B: CORE-EXCISED LUMINOSITY

Because cluster cores contain high-density gas with complex thermal structure, excess X-ray emission can often be produced within this region, typically taken to be radii interior to $0.15R_{500c}$. Here, we illustrate how core excision affects population statistics examined for non-core excised quantities in the main text. The core-excised luminosity, $L_{X,ce}$, is calculated in the same way as L_X (see Section 2.1), but with all gas cells within $0.15R_{500c}$ being excluded.

Fig. B1 present the MPQ for both measures in both simulated halo samples at $z = 0$. In TNG, we observe no change in the MPQ for group-scale haloes with $M_{500c} < 10^{14}\,\mathrm{M}_\odot$. However, for more massive haloes, the $L_{X,ce}$ MPQ is nearly mass-independent exhibiting a mass scatter of ~ 20 per cent, significantly smaller than that of L_X. Conversely, in the FLAMINGO sample, the MPQ of $L_{X,ce}$ for low-mass groups is slightly smaller than that of L_X, while the two measures are nearly identical for larger haloes. We also measure the luminosity MPQs in FLAMINGO with the inclusion of gas cells recently heated by AGN feedback. We observe that these cells introduce an added mass scatter of < 5 per cent throughout the mass range for L_X, while for $M_{500c} > 10^{14}\,\mathrm{M}_\odot$, there is little to no added scatter for $L_{X,ce}$. The added scatter can be attributed to the increased heating of dense gas which results in extremely luminous gas cells – most of which located near the core. These cells contribute to the added

intrinsic scatter of L_X and $L_{X,ce}$ without affecting their MPR slopes (see Fig. B2).
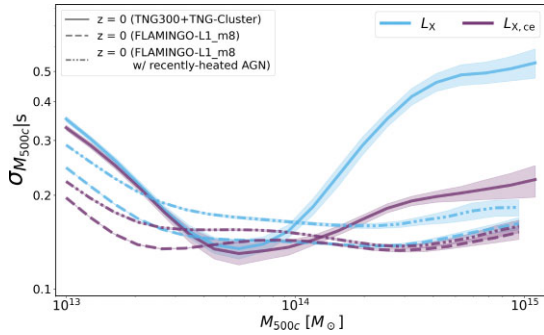

Figure B1. MPQ for X-ray luminosities without (L_X, cyan lines) and with ($L_{X,ce}$, purple) core-excision at $z = 0$ for TNG300 + TNG-CLUSTER (solid line) and FLAMINGO-L1_m8 excluding cells recently heated by AGN feedback (dashed line) and including those cells (dash–dotted line).

In Fig. B2, we display KLLR slope and scatter values for L_X–M_{500c} (left) and $L_{X,ce}$–M_{500c} (right) relations at $z = 0$. We observe modest changes in slopes after core excision, with a slight increase in slope for haloes with $M_{500c} < 10^{13.5}$ M$_\odot$. At these mass scales, the intrinsic scatter is also unaffected, implying that cool core activity is not important in low-mass groups.

At cluster scales, core excision reduces the scatter by nearly a factor of three for TNG haloes, whereas there is no difference in scatter between L_X and $L_{X,ce}$ at this scale in the FLAMINGO sample. The differing behaviours indicate that the gas phase structure of cores in TNG-CLUSTER and high-mass TNG300 haloes contain very luminous gas cells near the core, which are likely part of the abundant cool cores.

APPENDIX C: MASS-WEIGHTED TEMPERATURE

While the spectroscopic-like temperature is sensitive to gas phase structure, the simpler mass-weighted temperature is not. Here, we present behaviours for T_{mw}.

Fig. C1 compares the MPQs of T_{sl} and T_{mw} at $z = 0$. At the group scale, the mass scatter of T_{mw} is modestly reduced compared to

Figure B2. Scale-dependent MPR slope (top) and intrinsic scatter (bottom) of the L_X–M_{500c} (without core excision, left) and $L_{X,ce}$–M_{500c} (with core excision, right) at $z = 0$ in TNG300 + TNG-CLUSTER (solid line), in FLAMINGO-L1_m8 excluding recently heated AGN cells (dashed line) and including those cells (dotted line). The inclusion of recently heated AGN cells does not affect the slope, but increases the scatter of L_X and $L_{X,ce}$, specifically at the group scale.

that of T_{sl} for TNG haloes, while remaining nearly unchanged for FLAMINGO. However, at the cluster scale, T_{mw} has reduced scatter in both simulated samples, with the reduction being nearly a factor of two for TNG-CLUSTER and high-mass TNG300 haloes.

The effect of including gas that is recently heated by AGN feedback is minimal for T_{mw}, while for T_{sl}, there is an added mass scatter of ~ 17 per cent at the group scale that decreases to < 1 per cent at the cluster scale. The recently heated AGN gas cells boosts the temperature of dense gas cells which increases the intrinsic scatter in the T_{sl}–M_{500c} relation, while slightly decreasing the slope, leading to a larger mass scatter.

Paralleling the time-dependent MPRs shown in Figs 4 and 5, Fig. C2 presents the slope and intrinsic scatter of the T_{mw}–M_{500c} scaling relation over time. The scale-dependence of the slope at $z = 0$ is similar to that of T_{sl}, with minimum slope values in both simulations occurring near $10^{13.8}$ M$_\odot$. In both simulations, the sensitivity to redshift is more modest in T_{mw} at nearly all mass scales.

Intrinsic scatter values are substantially smaller for T_{mw} than T_{sl}, and there is a much less pronounced scale-dependence, particularly at $z = 0$. For FLAMINGO, the scatter in T_{mw} is larger than that of T_{sl} for haloes with mass $10^{13.7}$ M$_\odot$ $< M_{500c} < 10^{14.2}$ which, when combined with smaller slopes, results in the halo mass scatter of T_{mw} shown in Fig. C1 being slightly larger for this mass range. In contrast, for more massive haloes, the scatter in T_{mw} is smaller by approximately 25 per cent which, when combined with a slightly larger slope, yields a smaller mass scatter.
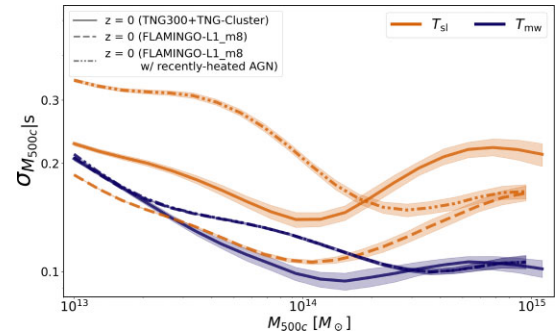

Figure C1. MPQs of core-excised spectroscopic-like temperature T_{sl} (orange) and mass-weighted average temperature T_{mw} (dark blue) at $z = 0$ for TNG300 + TNG-CLUSTER (solid line) and FLAMINGO-L1_m8 excluding cells recently heated by AGN feedback (dashed line) and including those cells (dash–dotted line). For T_{mw}, there is no difference in MPQ when adding the gas that is recently heated by AGN feedback.
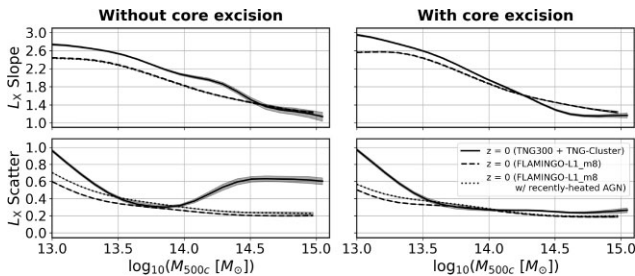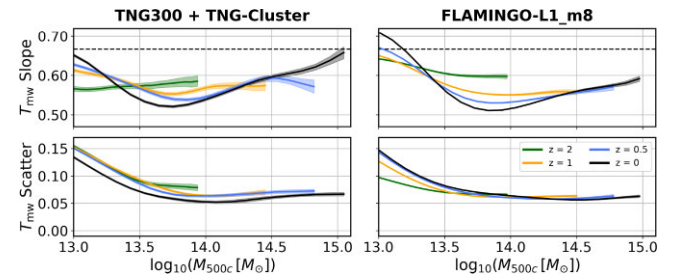

Figure C2. Scale-dependent MPR slopes (top) and intrinsic scatter values (bottom) of the mass-weighted temperature in TNG300 + TNG-CLUSTER (left) and FLAMINGO-L1_m8 (right) samples at redshifts indicated in the legend.
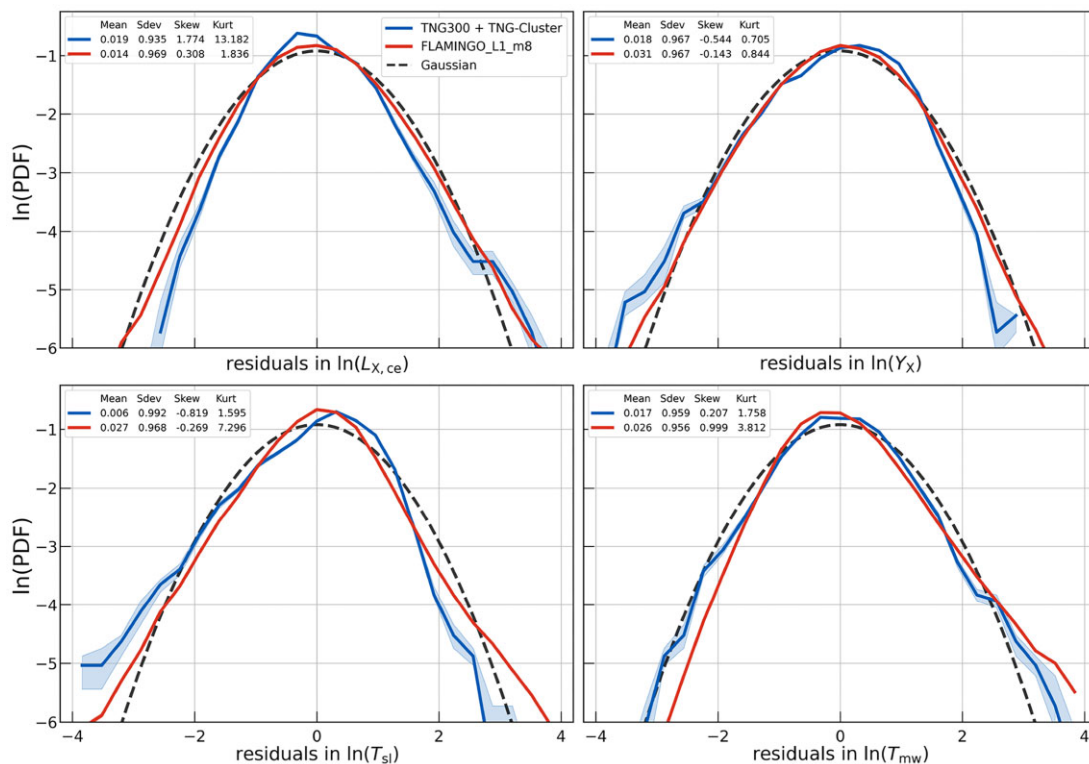
Figure D1. Mass-conditioned-likelihood forms at $z = 0$ of core-excised X-ray luminosity $L_{X,ce}$ (top left), X-ray thermal energy Y_X (top right), core-excised spectroscopic-like temperature T_{sl} (bottom left), and mass-weighted average temperature T_{mw} (bottom right).

APPENDIX D: ADDITIONAL LIKELIHOOD SHAPES

Expanding upon the three property kernels shown in Fig. 7, we show in Fig. D1 mass-conditioned likelihoods of four additional properties at $z = 0$: core-excised luminosity $L_{X,ce}$; X-ray thermal energy Y_X; core-excised spectroscopic-like temperature T_{sl}; and mass-weighted average temperature T_{mw}.

The shapes are approximately Gaussian but deviations are apparent. Compared to the case of L_X in Fig. 7, the distribution of $L_{X,ce}$ residuals is similar in form, with a tail to high values particularly pronounces in TNG. Residuals in both T_{sl} and Y_X are mildly skewed negative in both simulated samples, and the FLAMINGO distribution of T_{sl} has tails at both low and high values. The T_{mw} distribution has a tail to high values and is skew positive in both simulation samples.

APPENDIX E: SCALING RELATION ABSOLUTE MEANS

In Fig. 6 of the main text, the log-mean behaviours of the hot gas properties are shown as deviations relative to self-similar scaling in both halo mass and redshift. In Fig. E1, we present these scaling relations in a manner that retains the expected self-similar redshift scaling while removing the rescaling with halo mass scale. The $z = 0$ normalization differences listed in Table 4 are now apparent by close reading of this figure.

At the cluster scale, $M_{500c} > 10^{14}\,M_\odot$, hot gas mass M_{gas} and thermal energy Y_{SZ} adhere remarkably closely to self-similar evolution. Small deviations are seen for other measures, with T_{sl} and Y_X being slightly lower, and L_X slightly higher, at larger redshifts relative to the $z = 0$ self-similarly scaled values. These drifts reflect the cumulative heating of the ICM plasma, primarily from AGN feedback, over time.

The effects of such feedback are more apparent at the group scale. Low-mass haloes at high redshift have more gas, are cooler, and emit more X-ray luminosity than those at $z = 0$. Over time, moderate star formation in these haloes removes hot gas and feedback from growing SMBHs raises ambient gas temperatures in these systems. These effects conspire to reduce X-ray emission at late times. As noted in the main text, gas thermal energy measures adhere more closely to self-similar redshift evolution due to opposing trends in gas mass (declining) and temperature (rising) over time.
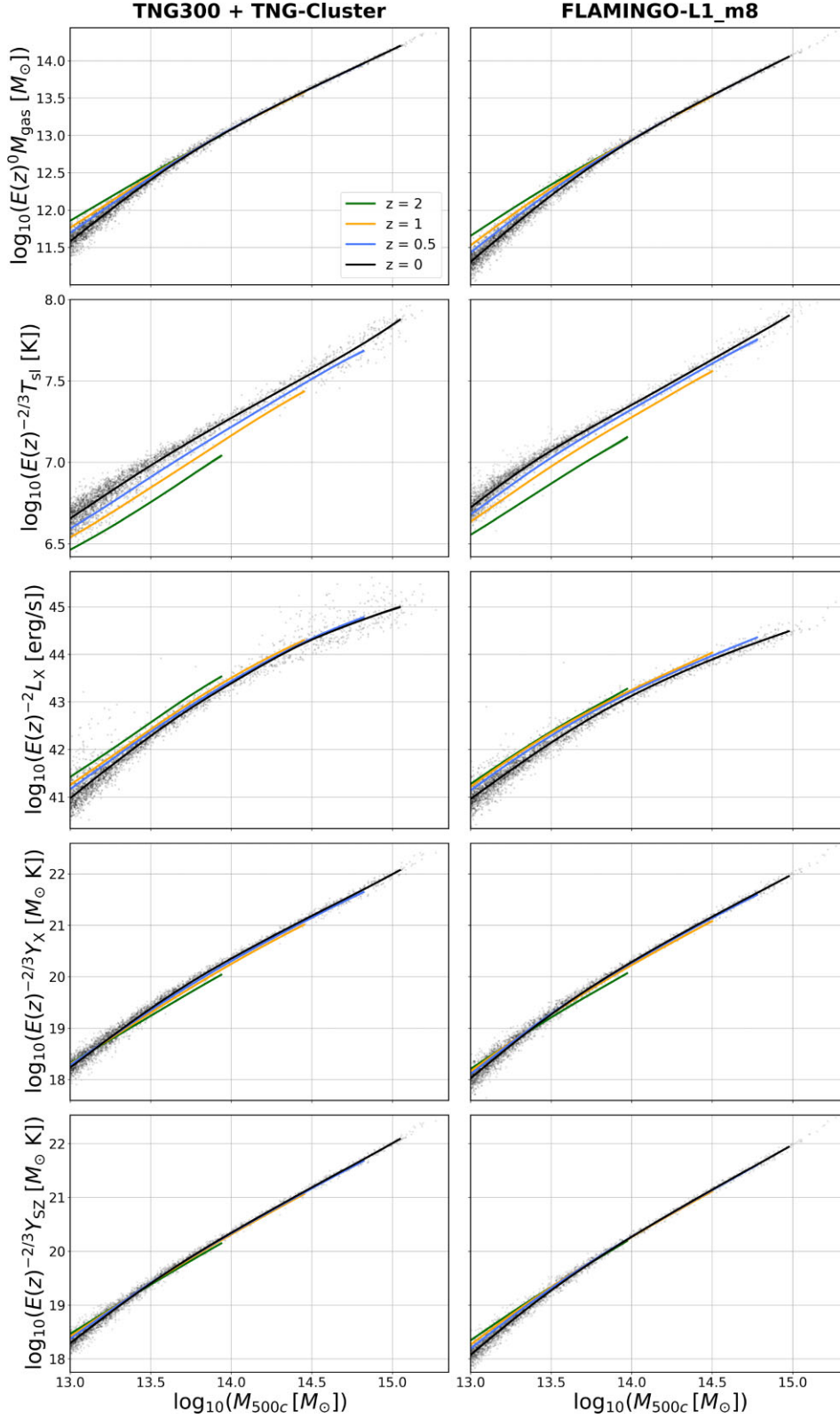
Figure E1. Logarithmic-mean gas property values as a function of halo mass, including self-similar redshift scaling for each property, using the same format as that of Fig. 6. The resampled bootstrap 1σ uncertainties are smaller than the width of the lines. We plot the individual halo properties at $z = 0$, omitting other redshifts for clarity, while subsampling FLAMINGO haloes to produce a comparable sample size to that of TNG in each bin.

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