Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **495,** 686–704 (2020)
Advance Access publication 2020 April 28

doi:10.1093/mnras/staa1147

Stellar property statistics of massive haloes from cosmological hydrodynamics simulations: common kernel shapes

Dhayaa Anbajagane [1]★ August E. Evrard [1,2] Arya Farahi [3,4] David J. Barnes [5] Klaus Dolag,[6,7] Ian G. McCarthy [8] Dylan Nelson [6] and Annalisa Pillepich[9]

[1]*Department of Physics and Leinweber Center for Theoretical Physics, University of Michigan, Ann Arbor, MI 48109, USA*
[2]*Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA*
[3]*Michigan Institute for Data Science, University of Michigan, Ann Arbor, MI 48109, USA*
[4]*McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15312, USA*
[5]*Department of Physics, Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[6]*Max-Planck Institut für Astrophysik, Karl-Schwarzschild Str 1, D-85741 Garching, Germany*
[7]*University Observatory Munich, Scheinerstr 1, D-81679 München, Germany*
[8]*Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK*
[9]*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*

Accepted 2020 April 21. Received 2020 April 20; in original form 2020 February 5

ABSTRACT

We study stellar property statistics, including satellite galaxy occupation, of haloes in three cosmological hydrodynamics simulations: BAHAMAS + MACSIS, IllustrisTNG, and Magneticum Pathfinder. Applying localized linear regression, we extract halo mass-conditioned normalizations, slopes, and intrinsic covariance for (i) $N_{\rm sat}$, the number of stellar mass-thresholded satellite galaxies within radius $R_{\rm 200c}$ of the halo; (ii) $M_{\star,\rm tot}$, the total stellar mass within that radius, and (iii) $M_{\star,\rm BCG}$, the gravitationally bound stellar mass of the central galaxy within a 100 kpc radius. The parameters show differences across the simulations, in part from numerical resolution, but there is qualitative agreement for the $N_{\rm sat}$–$M_{\star,\rm BCG}$ correlation. Marginalizing over $M_{\rm halo}$, we find the $N_{\rm sat}$ kernel, $p(\ln N_{\rm sat} \mid M_{\rm halo}, z)$ to be consistently skewed left in all three simulations, with skewness parameter $\gamma = -0.91 \pm 0.02$, while the $M_{\star,\rm tot}$ kernel shape is closer to lognormal. The highest resolution simulations find $\gamma \simeq -0.8$ for the $z = 0$ shape of the $M_{\star,\rm BCG}$ kernel. We provide a Gaussian mixture fit to the low-redshift $N_{\rm sat}$ kernel as well as local linear regression parameters tabulated for $M_{\rm halo} > 10^{13.5}\,{\rm M}_\odot$ in all simulations.

Key words: methods: statistical – galaxies: clusters: general – galaxies: evolution – galaxies: haloes.

1 INTRODUCTION

Clusters of galaxies, and the underlying dark matter haloes that host these systems, are important to study as their population behaviour is sensitive to both the expansion history and the gravitational growth of large-scale structure in our Universe (Allen, Evrard & Mantz 2011).

When using clusters of galaxies for cosmology, the statistical relationship between an observable cluster property and the total mass of its host halo, which we call the mass–property relation (MPR), is a key model element (e.g. Vikhlinin et al. 2009; Mantz et al. 2010; Rozo et al. 2010; Zhang et al. 2011; de Haan et al. 2016; Pillepich et al. 2018c; Bocquet et al. 2019; Costanzi et al. 2019; Mulroy et al. 2019). Current and near-future experiments will expand cluster sample sizes into the tens of thousands (The Dark Energy Survey Collaboration 2005; Laureijs et al. 2011; Merloni et al. 2012; Pillepich, Porciani & Reiprich 2012; Predehl et al. 2014; Spergel et al. 2015; Ade et al. 2019), allowing for improved understanding of the MPR for hot gas and stellar properties. This understanding, along with careful modelling of survey selection, are crucial elements that empower studies of cosmic acceleration and of new physics using massive haloes.

When the observable property is a count of galaxies above some size (luminosity, stellar mass, etc.) threshold, the MPR is analogous to the halo occupation distribution (HOD, Berlind & Weinberg (HOD; Berlind & Weinberg 2002; Cooray & Sheth 2002; Hearin et al. 2013; Zentner, Hearin & van den Bosch 2014; Hearin et al. 2016; Zehavi et al. 2018). The stellar property statistics of dark matter haloes lie within the broad category of the 'galaxy–halo connection' which was reviewed recently by Wechsler & Tinker (2018). The focus of this work is the high mass population of haloes, each of which host multiple bright galaxies.

★ E-mail: dhayaa@umich.edu

© 2020 The Author(s)
Published by Oxford University Press on behalf of the Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

The statistics of the stellar component properties – galaxy occupation, central galaxy stellar mass, and total stellar mass of a halo – across a broad range of total halo mass and redshift is a fundamental outcome of the complex astrophysical processes that drive galaxy formation (see e.g. Pillepich et al. 2018b, for a detailed discussion).

A low-accretion rate mode of supermassive black hole (SMBH) feedback within the cores of large galaxies (Croton et al. 2006; De Lucia et al. 2006) is employed to solve the problem of excessive cooling and star formation at the centres of groups and clusters seen in early simulations (Katz & White 1993; Evrard, Summers & Davis 1994). Jet-driven, turbulent feedback from SMBH accretion appears to be an important regulator of the thermodynamic state of core gas (McNamara & Nulsen 2012; Voit et al. 2015) and its inclusion in cosmological simulations has significantly improved the fidelity of galaxy and hot gas properties within the population of high mass haloes (Gaspari, Ruszkowski & Oh 2013; Ragone-Figueroa et al. 2013; Hirschmann et al. 2014; Khandai et al. 2015; Rasia et al. 2015; Hahn et al. 2017; Kaviraj et al. 2017; McCarthy et al. 2017; Nelson et al. 2018; Pillepich et al. 2018b)

While hydrodynamic and magnetohydrodynamic methods have improved dramatically in terms of both numerical resolution and astrophysical treatments (Kravtsov & Borgani 2012; Vogelsberger et al. 2020), direct inter-comparisons of different numerical solutions reveal varying degrees of inconsistency (e.g. Scannapieco et al. 2012; Elahi et al. 2016). The aim of this work is to compare compressed statistical summaries of the aforementioned stellar properties for populations of massive haloes realized by independent state-of-the-art methods. We apply a local linear regression (LLR) approach first used by Farahi et al. (2018) to describe the statistics of hot gas and stellar mass, conditioned on total halo mass, for haloes realized by the BAHAMAS and MACSIS simulations. As long as halo properties are well behaved functions of mass and redshift, the LLR method is highly effective at compressing the full range of discrete population measurements into a small number of statistical parameters.

This work expands on Farahi et al. (2018) by (i) using multiple stellar properties associated with a halo and (ii) performing a verification test using results from multiple simulation teams.

We utilize three cosmological hydrodynamics simulations – a superset of BAHAMAS and MACSIS, the IllustrisTNG 300-1 run, and the Magneticum Pathfinder 500 Mpc volume – each of which contains >1000 haloes with $M_{200c} > 10^{13.5}\,\mathrm{M}_\odot$ at $z = 0$. Bootstrap resampling of each discrete population is used to estimate statistical uncertainties in scaling relation parameters.

The statistical power of these large halo samples is reflected by relatively small fractional errors in the recovered LLR parameters. It is worth keeping in mind that the bootstrap-estimated errors should conservatively be treated as lower bounds relative to what might be achieved by sampling an essentially infinite cosmic volume. Still, the different simulations frequently produce results in mild statistical tension with one another, in which case the range of values in a particular quantity of interest can be considered an estimate of the magnitude of that parameter's theoretical uncertainty as estimated by first-principle simulations.

While statistical tensions exist for many of the derived LLR parameter values, we also find areas of congruence, particularly in the fundamental forms of mass-conditioned property kernels. Congruent results offer a necessary step of *verification* (e.g. Salvadori 2019), meaning that halo populations with consistent stellar MPRs emerge from independent solutions of the equations governing the complex, non-linear system of large-scale structure. A *validation* step using observational data must be done using observable proxies for the intrinsic true properties we use here. Observational analysis with careful treatment of sample selection is emerging (Mantz et al. 2016; Bocquet et al. 2019; Farahi et al. 2019a; Mulroy et al. 2019) but we do not attempt detailed comparisons to observational samples in this paper.

We employ an M_{200c} spherical overdensity mass scale convention,[1] and define N_{sat} as the integer count of satellite galaxies with $M_{\star,sat} > 10^{10}\,\mathrm{M}_\odot$ lying within that radius.[2] The form of the conditional likelihood, $p(\ln N_{sat} \mid M_{halo}, z)$, a core component of HOD models, is a particular area of focus, and a key finding of our study is that all three simulations produce a consistent *shape* for both this kernel and that of the total stellar mass within R_{200c}.

We also find that, at fixed halo mass, satellite galaxy number is anticorrelated with the stellar mass of the central galaxy of a halo, as would be expected if central galaxies, which we refer to as the brightest central galaxy (BCG), grow primarily at late times by cannibalizing satellite galaxies (Tremaine & Richstone 1977; De Lucia & Blaizot 2007). Larger than average BCG stellar masses are also associated with magnitude gaps measured with respect to lower ranked galaxies (e.g. Golden-Marx & Miller 2018).

The structure of this paper is as follows. In Section 2, we describe the simulation samples while Section 3 describes the localized linear regression (LLR) method applied to generate summary statistics for each simulation's halo sample. We examine satellite galaxy scaling relations in Section 4, along with relevant aspects of scaling relations of total stellar mass and central galaxy stellar mass. Here, we also provide a two-component Gaussian mixture model (GMM) fit to $p(N_{sat} \mid M, z)$ that describes the consistent shape seen in all simulation ensembles at $z < 1$. The interrelationships of BCG stellar mass with other properties is explored in Section 5. We elaborate briefly on the future of such multisimulation comparison studies in Section 6, and summarize our findings in Section 7. Appendices provide complete results for all stellar properties (Appendix A) as well as the GMM and LLR parameter tables (Appendix B).

2 SIMULATIONS AND HALO POPULATIONS

We analyse three different simulations – BAHAMAS + MACSIS (BM), TNG300-1 from the IllustrisTNG project (TNG300), and the 500 Mpc high-resolution box 2 from Magneticum Pathfinder (MGTM) – with characteristics summarized in Table 1. We also include a $z = 0.12$ output of a higher resolution BAHAMAS run, a 143 Mpc volume (labelled B100), to examine resolution dependence of the derived statistics. Collectively, the simulations produce nearly 15 000 haloes with total mass, $M_{200c} > 10^{13.5}\,\mathrm{M}_\odot$, at $z = 0$.

The simulations use slightly different cosmological parameters, but all assume a flat geometry with matter density parameter in the range $\Omega_m \in [0.2726, 0.3175]$. The cosmic baryon fraction, $f_b \equiv \Omega_b/\Omega_m$, which varies from a low of 0.154 for BM to a high of 0.167 for MGTM, is important for setting the normalization of the stellar mass formed within haloes and so we make small adjustments to the satellite galaxy stellar mass cut-off described below.

The BAHAMAS simulations (McCarthy et al. 2017) use a version of the smoothed particle hydrodynamics (SPH) code GADGET (Springel 2005) to model a 596 Mpc periodic cube. The MACSIS

[1]The radius R_{200c} satisfies $3M(< R_{200c})/(4\pi R_{200c}^3) = 200\rho_{crit}(z)$, where $\rho_{crit}(z)$ is the critical density of the universe, and $M_{200c} \equiv M(<R_{200c})$.
[2]Slight adjustments are made to normalize the mean cosmic baryon fraction, as described in Section 2.2.

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

Table 1. Simulation characteristics and $z = 0$ halo population sizes. Empirical sources for tuning sub-grid parameters are given in the last column and consist of the galaxy stellar mass function (GSMF), supermassive black hole (SMBH) scaling, metallicity scaling (Metals), and cluster hot gas mass fraction $<R_{500c}$ (CL $f_{\rm gas}$). All assume a flat Λ cold dark matter cosmology, so $\Omega_\Lambda = 1 - \Omega_m$, and their cosmic mean baryon fraction is $f_b \equiv \Omega_b/\Omega_m$. The MGTM output is actually $z = 0.03$ and B100 (used only for testing resolution) is $z = 0.12$ while the other two are exactly $z = 0$. See text for references.

Simulation[a]	L (Mpc)[b]	Ω_m	f_b	σ_8	$\epsilon_{\rm DM}^{z=0}$ (kpc)	$m_\star$ ($\rm M_\odot$)[c]	$\log_{10}(M_{20})$[d]	$N_{\rm sam}$[e]	Calibration
BM	596	0.3175	0.154	0.834	5.96	1.2×10^9	15.6	9430	GSMF, CL $f_{\rm gas}$
B100	143	0.2793	0.166	0.821	2.86	1.4×10^8	14.1	96	GSMF, CL $f_{\rm gas}$
MGTM	500	0.2726	0.167	0.809	5.33	5.0×10^7	14.9	4207	SMBH, metals, CL $f_{\rm gas}$
TNG300	303	0.3089	0.157	0.8159	1.48	1.1×10^7	14.6	1146	See Pillepich et al. (2018a)

[a]See text for description of acronyms.
[b]Comoving simulation cube length except for MACSIS (subset of the BM data), which subsamples a 3.2 Gpc cubic volume.
[c]Initial stellar particle mass.
[d]Upper limit of LLR regression at $z = 0$, the 20th most massive halo mass, in $\rm M_\odot$.
[e]Number of haloes with total mass, $M_{200c} > 10^{13.5}\,\rm M_\odot$. The number above $10^{13.8}\,\rm M_\odot$ for BM is ≈4400.

ensemble (Barnes et al. 2017) comprises 390 'zoom-in' simulations of individual halo regions, drawn from a parent 3.2 Gpc N-body simulation, that completely resamples the most massive haloes in the large parent volume. The MACSIS resimulations use the same code base and have the same numerical resolution and astrophysical treatments as BAHAMAS but the resimulation technique enables LLR fits extending to $M_{200c} = 4 \times 10^{15}\,\rm M_\odot$ at $z = 0$. The BM sample is the superset of BAHAMAS and MACSIS populations. Conversely, the 140 Mpc Bahamas higher resolution simulation (B100) contains only 96 haloes with $M_{200c} > 10^{13.5}\,\rm M_\odot$ at $z = 0$, only a few dozen of which lie above $10^{14}\,\rm M_\odot$.

The MGTM 500 Mpc run (Hirschmann et al. 2014) uses a different fork of GADGET with an improved fluid solver (Donnert et al. 2013) to model the co-evolution of supermassive black holes and their host galaxies. While both BM and MGTM employ variants of the same base GADGET code, the detailed treatments of star formation, black hole seeding and growth, and feedback from these compact populations were developed independently.

IllustrisTNG (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018; Pillepich et al. 2018b; Springel et al. 2018) is a follow-up to the Illustris simulations (Vogelsberger et al. 2014) based on the deformable-mesh hydrodynamics solver, AREPO (Springel 2010). As detailed in Pillepich et al. (2018b), the stellar mass functions and stellar mass to halo mass relations are in reasonable agreement with observational and empirical constraints. At the same time, Springel et al. (2018) show that the observed large-scale spatial clustering of galaxies, and its dependence on mass and colour, is also reproduced. In this paper, we use the TNG300 run, the simulation of the TNG suite with the largest volume.

The last four columns of Table 1 highlight important similarities and differences among the simulations. The gravitational softening length used for dark matter and stars is similar in BM and MGTM (5–6 kpc) but both are a factor ∼3 larger than that of TNG300. TNG300 and MGTM employ star particles that are significantly less massive than that of BM. Because BM models a $10^{10}\,\rm M_\odot$ stellar mass galaxy with as few as nine star particles, we find below that the variance in satellite galaxy count is enhanced with respect to both B100 and the other two simulations.

Resolution convergence is a non-trivial issue for all hydrodynamical simulations of galaxy formation. For example, Pillepich et al. (2018a,b) analyse the rate of convergence in stellar mass contents of haloes in the TNG model, finding shifts of ∼40 per cent in galaxy stellar mass between the two highest levels of resolution of TNG at the high-mass end. The three simulations span a range of length, mass, and time resolution, so the differences in stellar mass statistics

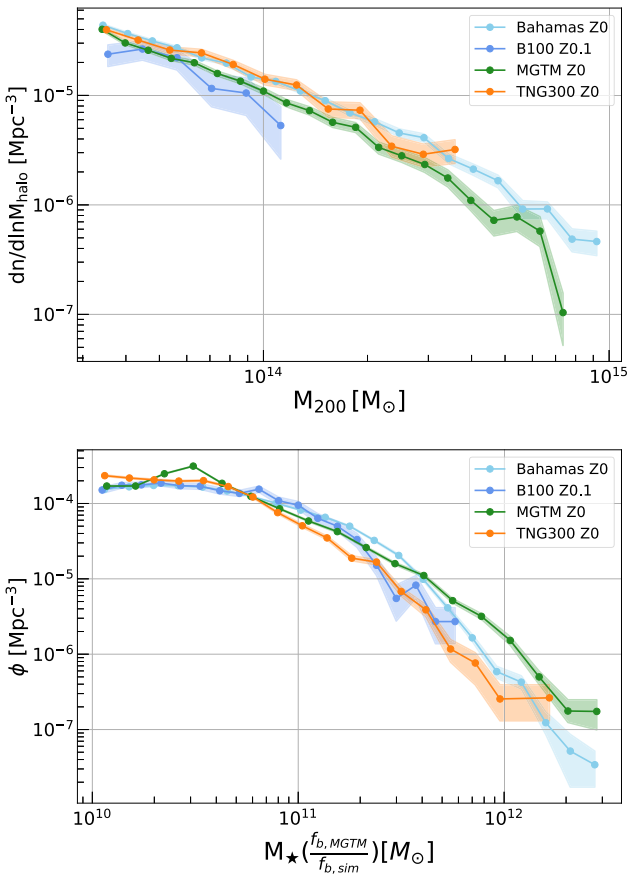

Figure 1. *Upper:* Halo mass function of all four simulations for $M_{\rm halo} > 10^{13.5}\,\rm M_\odot$ at $z = 0$ (BAHAMAS, TNG300), $z = 0.03$ (MGTM), and 0.12 (B100). *Lower:* Satellite galaxy stellar mass function within R_{200c} in haloes with $M_{200c} > 10^{13.5}\,\rm M_\odot$ at $z = 0$ (BAHAMAS, TNG300), $z = 0.06$ (MGTM), and 0.12 (B100). Stellar masses for BAHAMAS, B100, and TNG300 are adjusted to match the MGTM baryon fraction. Errorbars do not include cosmic variance.

represents a convolution of numerical (e.g. resolution) and physical model (e.g. forms of feedback) differences.

Given the modest range of cosmologies explored by the simulations in the ensemble, one expects similar dark matter halo mass functions at $z = 0$. The top panel of Fig. 1 shows that this is the case, with the high Ω_m cosmologies of BM and TNG300 shifted

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

Table 2. Property definitions.

Quantity	Definition
Halo centre	Minimum gravitational potential[a]
Halo total mass, M_{200c}	All species within R_{200c} sphere[b]
Halo stellar mass, $M_{\star,\rm tot}$	All stellar particles within R_{200c}
Galaxy centre	Most bound particle[c]
Galaxy stellar mass, $M_\star$	Gravitationally bound stellar particles

[a]FOF links of 0.2 (BM, TNG) and 0.16 (MGTM) mean separation.
[b]TNG halo masses use only FOF particle set.
[c]Of any species (BM, TNG) or collisionless only (MGTM).

above MGTM, which has comparatively lower values of both Ω_m and σ_8.

The satellite galaxy stellar mass function (S-GSMF) for groups and clusters, shown in the bottom panel of Fig. 1, is derived from the full collection of satellite galaxies within R_{200c} of our target halo population with $M_{200c} > 10^{13.5}\,{\rm M}_\odot$. The S-GSMF space density is computed here using the full simulation volume (not the volume occupied by haloes), and stellar masses have been normalized to the baryon fraction of MGTM.

There is fairly good agreement in the S-GSMF among the simulations for galaxies with $10^{10} < M_\star/{\rm M}_\odot < 10^{11}$ but the populations diverge somewhat at high masses. Note that the BAHAMAS simulation, despite the higher stellar particle mass, shows no sign of incompleteness down to the $10^{10}\,{\rm M}_\odot$ stellar mass limit. At $10^{12}\,{\rm M}_\odot$, the MGTM simulation produces nearly seven times as many galaxies per unit volume as the TNG300 solution. Given that MGMT has the lowest space density of high mass haloes (top panel of Fig. 1), we anticipate that the HOD normalization of MGTM is higher than those of the other two simulations. The differences in the outcome of the $z \sim 0$ GSMF are a consequence of different underlying choices and implementations of stellar and black hole feedback leading to different effective outcomes in the regulation and quenching of star formation.

2.1 Halo finding and catalogues

Our analysis methods employ catalogue-level products derived independently by each simulation team. The methods operate to identify common halo and stellar properties listed in Table 2.

The identification of the parent population of haloes is done in a similar fashion across the four simulations, with a percolation (friends-of-friends) step followed by an identification of locally bound sub-structures using the SUBFIND algorithm (Springel et al. 2001; Dolag et al. 2009). Halo centres are identified as the location of the most-bound particle and M_{200c} enclosed masses are derived using all particle species.

Stellar properties of galaxies within sub-haloes are then derived. The common use of SUBFIND means that the galaxy stellar masses are well aligned across the simulations, employing common definitions of the centre as well as the local binding energy condition.

Our study is based on halo catalogues defined by a total mass threshold, $M_{200c} > 10^{13.5}\,{\rm M}_\odot$, and we examine snapshots at $z = 0$, 0.5, 1, 1.5, and 2. Due to their larger volumes, BM and MGTM offer samples of more than 4000 haloes at $z = 0$ while the higher resolution TNG300 simulation yields 1130.

2.2 Stellar property vector

We use a vector of stellar properties listed in Table 3. The satellite galaxy count, $N_{\rm sat}$, is the number of galaxies within R_{200c} having individual stellar mass $M_\star > (\frac{f_{\rm b,\,Sim}}{f_{\rm b,\,MGTM}}) \times 10^{10}\,{\rm M}_\odot$. We apply the linear correction to account for global baryon fraction differences across the simulations (see Table 1) and arbitrarily normalize to the MGTM value. The mass limits for BM and TNG300 are thus ∼8 per cent smaller than the MGTM value.

Table 3. Property components of LLR regression vector, **S**.

Symbol	Quantity
$N_{\rm sat}$	Count of non-central galaxies within R_{200c}[a]
$M_{\star,\rm tot}$	Total stellar mass within R_{200c} ($\rm M_\odot$)
$M_{\star,\rm BCG}$	Central galaxy stellar mass within 100 kpc ($\rm M_\odot$)

[a]Stellar mass-limited, $M_\star > (\frac{f_{b,\,\rm Sim}}{f_{b,\,\rm MGTM}}) \times 10^{10}\,{\rm M}_\odot$.

This stellar mass threshold is chosen to balance resolution considerations with discreteness in $N_{\rm sat}$ counts. At $10^{10}\,{\rm M}_\odot$, we are working with galaxies resolved by 9 (BM), 73 (B100), 200 (MGTM), and 910 star particles (TNG300). Lowering the threshold would decrease the minimum number of stellar particles in our selected galaxies leading to unacceptably small values for BM. Raising the threshold instead leaves us with more haloes exhibiting either no satellites or just 1 or 2 of them, which complicates our calculations using $\ln N_{\rm sat}$. Our stellar mass cut-off of $\sim 10^{10}\,{\rm M}_\odot$ lies significantly below the $\sim 10^{10.8}\,{\rm M}_\odot$ knee of the observed stellar mass function (Moustakas et al. 2013).

The rest of the stellar property vector consists of $M_{\star,\rm tot}$, the total stellar mass within R_{200c} as well as $M_{\star,\rm BCG}$, the central galaxy's bound stellar mass within a fixed sphere of 100 kpc physical radius.

3 LOCAL LINEAR REGRESSION METHOD

From dimensional arguments one can infer that integral properties of massive haloes, such as aggregate stellar mass or global X-ray temperature measured at some redshift will, in the mean, scale as power laws with total system mass (Kaiser 1986; Bryan & Norman 1998). Due to variations in formation history and dynamical state, any individual halo will be offset from the population mean, and this intrinsic dispersion is often assumed to be lognormal in form (Evrard et al. 2014).

Linear regression of the simple least-squares variety has been a canonical method used to characterize cluster scaling laws, but its utility is limited by the fact that it reduces full population statistics for a given property down to three numbers: a slope, normalization, and variance/standard deviation. With large halo samples extracted from cosmological simulations, we can perform a more sensitive analysis using *localized* linear regression (Farahi et al. 2018). The LLR method generates mass-conditioned estimates of the slope, normalization, and property covariance, where the term 'mass-conditioned' implies that we are determining these parameters given a certain halo mass and redshift.

3.1 Mass-conditioned parameters and normalized residuals

Using natural logarithms of the properties, $\boldsymbol{s} = \ln \boldsymbol{S}$, the population mean of the log of property, S_a, at a fixed redshift scales with halo mass M as

$$\langle s_a \mid M, z\rangle \;=\; \pi_a(M_c, z) + \alpha_a(M_c, z)\ln(M/M_c), \qquad (1)$$

where M_c is a mass scale of interest, and the log-linear relation has slope, $\alpha_a(M_c, z)$, and intercept, $\pi_a(M_c, z)$, that, in general, depend on both redshift and the chosen halo mass scale. The subscript a

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

denotes the property under consideration and, at mass M_c, the modal and median value of S_a is $e^{\pi_a(M,z)}$.

The mass dependence of the fit parameters is found by applying a mass-dependent weight factor centred on the chosen mass scale. Letting $M_c \rightarrow M$ for simplicity, we minimize the weighted square error

$$\epsilon_a^2(M,z) = \sum_{i=1}^{n} w_i^2 \, (s_{a,i} - \alpha_a(M,z)\mu_i - \pi_a(M,z))^2, \quad (2)$$

where $\mu_i \equiv \ln(M_{\rm halo,i}/M)$, the sum i is over all haloes, and w_i is the mass-dependent Gaussian weight factor

$$w_i = \frac{1}{\sqrt{2\pi}\sigma_{\rm LLR}} \exp\left\{-\frac{\mu_i^2}{2\sigma_{\rm LLR}^2}\right\}, \quad (3)$$

with $\sigma_{\rm LLR}$ the width of the log-mass filter. For ideal mass localization, we want $\sigma_{\rm LLR}$ to be very small, but the finite sample sizes from the simulations prevent us from using too small a value. We use $\sigma_{\rm LLR} = 0.46$, equivalent to 0.2 dex in halo mass. Our results are relatively insensitive to this choice, but choosing too small a value leads to noisy features at high mass where the population density is sparse.

As the central filter mass scale, M, is varied in fixed logarithmic steps, we estimate the local slope, $\widetilde{\alpha}_a(M,z)$, intercept, $\widetilde{\pi}_a(M,z)$, and scatter, $\widetilde{\sigma}_a(M,z)$, parameters by minimizing the locally weighted square error (equation 2). The residual deviation in property a for a specific halo of mass M_i is then defined as

$$\delta s_{a,i} \equiv s_{a,i} - \widetilde{\pi}_a(M_i,z), \quad (4)$$

where the second term is determined by linear interpolation of the values sampled uniformly in the log of halo mass.

These residuals are combined to form the halo mass-conditioned property covariance

$$\mathrm{COV}(s_a,s_b) = A\sum_{i=1}^{n} w_i \, \delta s_{a,i} \, \delta s_{b,i}, \quad (5)$$

with normalizing pre-factor

$$A = \sum_{i=1}^{n} w_i \Big/ \left[\left(\sum_{i=1}^{n} w_i\right)^2 - \sum_{i=1}^{n} w_i^2\right]. \quad (6)$$

The corresponding property pair correlation coefficient is

$$r_{a,b} = \frac{\mathrm{COV}(s_a,s_b)}{\sigma_a\sigma_b}, \quad (7)$$

where $\sigma_a = \sqrt{\mathrm{COV}(s_a,s_a)}$ is the intrinsic scatter in property a at fixed halo mass, and similarly for property b.

Finally, we focus attention below on the *normalized residuals* in logarithmic properties, defined as

$$\widehat{\delta} s_{a,i} \equiv \frac{\delta s_{a,i}}{\widetilde{\sigma}_a} = \frac{s_{a,i} - \widetilde{\pi}_a(M_i,z)}{\widetilde{\sigma}_a} \quad (8)$$

4 LLR SCALING AND KERNEL SHAPES OF STELLAR PROPERTIES

We begin by presenting $z = 0$ scaling behaviours of the three stellar properties with $M_{\rm halo}$, finding qualitative agreement in many respects but also discrepancies in the details. Marginalizing over halo mass, we obtain estimates of the kernel shapes and find consistent support for $p(\ln N_{\rm sat}|M,z)$ to be negatively skewed while the kernel of total stellar mass is much closer to lognormal. We then examine off-diagonal elements of the mass-conditioned property covariance as motivation for exploring secondary selection by $M_{\star,\rm BCG}$, presented in Section 5.

We provide tables of LLR fit parameters for these properties in Appendix A. While our figures present parameter values in the range $10^{13.5}\,{\rm M}_\odot < M_{\rm halo} < M_{20}$, where M_{20} is the mass of the 20th most massive halo, the LLR method computes these parameters using *all* available haloes

4.1 Present-epoch stellar property scaling relations

Fig. 2 shows the $z = 0$ scalings of satellite galaxy counts, $N_{\rm sat}$, with halo mass for the four simulation populations. Solid lines show the LLR mean behaviours, which are inter-compared in the bottom right panel. As anticipated from the space densities of Fig. 1, the MGTM simulation has a higher normalization compared to the other two simulations. Numerical resolution is an important factor; the B100 simulation has 58 per cent more galaxies per halo compared to the BM model, whose dark matter particle mass is ∼9 times larger than that of B100. A shift of similar order of magnitude is found for the TNG300 suite. TNG300-1, the highest resolution run, has 86 per cent more galaxies per halo than TNG300-2, whose dark matter particle mass is eight times that of the TNG300-1 run. The quoted values are the mean shifts for haloes with $M_{\rm halo} > 10^{13.8}\,{\rm M}_\odot$.

The local slope and scatter for $N_{\rm sat}$ as a function of halo mass are presented in Fig. 3, with shaded regions showing 68 per cent confidence intervals from bootstrap resampling. The local slopes lie close to the simplest self-similar expectation of unity, with values in the range 0.90–1.05.

Above a halo mass of $\sim 10^{14}\,{\rm M}_\odot$, where the mean occupation numbers are ten or larger, the scatter in $\ln N_{\rm sat}$ declines with mass in a manner that roughly follows Poisson expectations, $\langle N_{\rm sat} \mid M,z\rangle^{-1/2}$, shown as dashed lines. The scatter in the BM model is significantly larger than this, but the low stellar mass resolution of this simulation appears to be adding extra variance. Evidence for this is given by the B100 simulation result. The scatter in $\ln N_{\rm sat}$ near $10^{14}\,{\rm M}_\odot$ drops from 44 per cent (BM) to 30 per cent (B100), a factor of 2 decline in variance that brings B100 much closer to the scatter values seen in TNG300 and MGTM populations. We confirm that a similar increase in variance occurs for the lower resolution TNG300-2 simulation. At $10^{14}\,{\rm M}_\odot$, the scatter in $\ln N_{\rm sat}$ is 48 per cent in the TNG300-2 run, compared to 37 per cent for TNG300-1.

For the sake of economy, we show the raw data for the scalings of central and total stellar mass measures ($M_{\star,\rm BCG}$ and $M_{\star,\rm tot}$) with halo mass at $z = 0$ in Appendix A. The mean stellar mass fraction within R_{200c} of a $10^{14}\,{\rm M}_\odot$ halo ranges from a high of 0.025 in MGTM to a low of 0.013 in TNG300, with Bahamas intermediate. These population values lie within the range of individual cluster stellar mass fractions inferred from small observational samples [see fig. 3 from Tremmel et al. (2019), and also work from Pillepich et al. (2018b) for additional comparisons]. The stellar mass fractions are fairly stable with redshift in all models (see tables in Appendix B) but its absolute value is again sensitive to resolution. Mean stellar masses in the higher resolution B100 run are elevated by ∼25 per cent relative to those of BM.

The slope and scatter of $M_{\star,\rm tot}$ and $M_{\star,\rm BCG}$ derived from the simulations are compared in Fig. 4. For the $M_{\star,\rm tot}$–$M_{\rm halo}$ relation, all simulations show sub-linear scaling, but approach self-similarity ($\alpha = 1$) at high halo mass, a result already published for the BM simulation (Farahi et al. 2018).

An observational study of 21 nearby clusters using an R_{500c} scale finds a somewhat shallower slope for the total stellar mass scaling, $\alpha = 0.6 \pm 0.1$ (Kravtsov, Vikhlinin & Meshcheryakov 2018). The

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

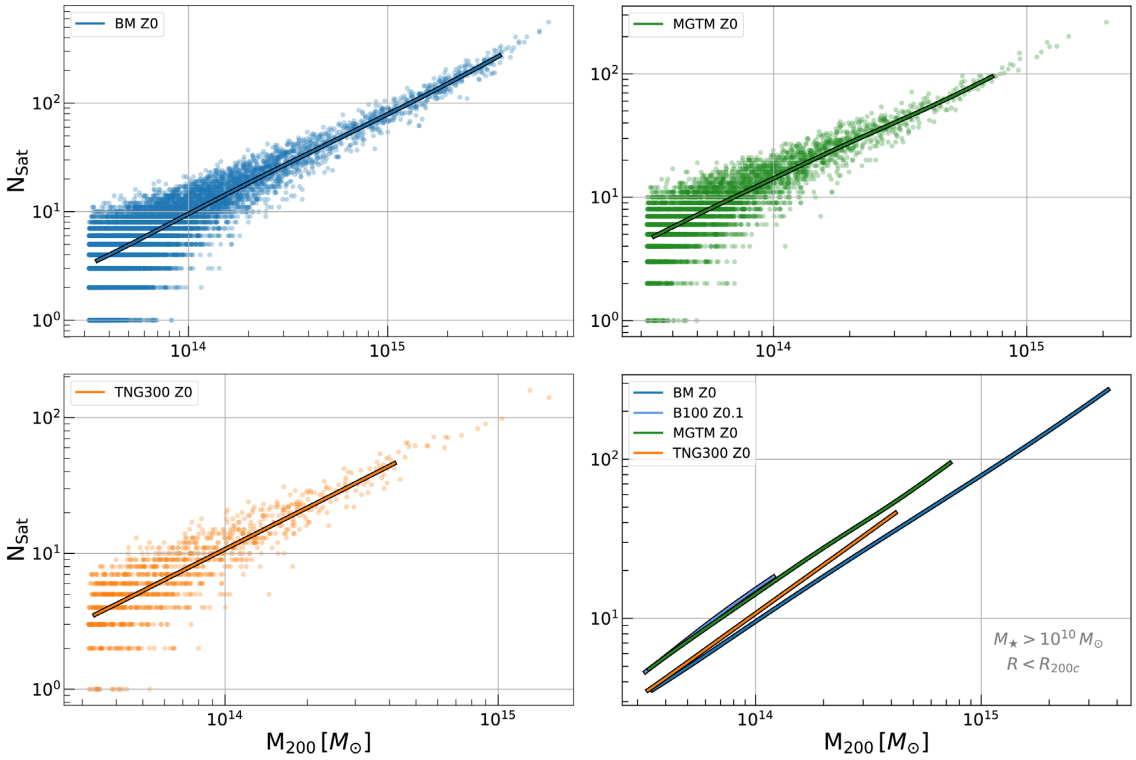

Figure 2. Number of stellar mass limited satellite galaxies, N_{sat}, as a function of total halo mass in the halo populations of each simulation at $z = 0$. Lines show mean LLR fits, computed using all available haloes, from the lower halo mass limit of $10^{13.5}$ $M_\odot$ to an upper limit determined by the 20th most massive halo of each sample. The lower right panel compares mean behaviours and includes results from B100, the higher resolution run of the BAHAMAS simulation from the BM sample.

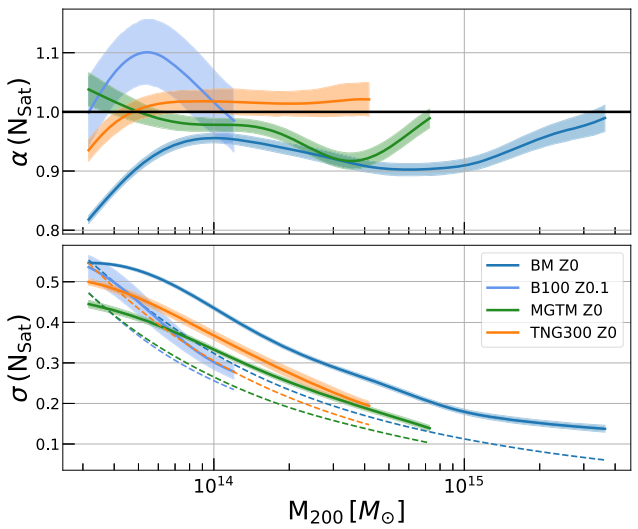

Figure 3. LLR slope, α (top), and natural logarithmic scatter, σ (bottom), of the N_{sat} scaling with halo mass for the $z = 0$ populations shown in Fig. 2. Dashed lines show the Poisson-expected fractional scatter, $\langle N_{sat} \mid M, z\rangle^{-1/2}$ obtained from the mean satellite number as a function of halo mass. B100, the high-resolution 140 Mpc BAHAMAS run, exhibits smaller scatter over the mass range accessible within that volume, reducing the tension with the other simulation outcomes.

tension in slope may reflect physical or numerical deficiencies in the simulations or it may reflect systematic differences in the two quantities being compared – true quantities in simulations versus those inferred from multiband photometry and other observations. Future work using synthetic observations to analyse the simulation expectations directly in the space of survey observables, including intracluster light, is needed to explore this discrepancy in more detail.

For the central galaxy scaling, all simulations display slopes that are sub-linear, lying in the range [0.45, 0.65], a range that encompasses values derived from the UniverseMachine semi-analytic models (Bradshaw et al. 2020).

Empirical studies of the BCG stellar mass slope find contradictory results. Some are in agreement with our results (Golden-Marx & Miller 2018) while some are shallower (Zhang et al. 2016; Kravtsov et al. 2018; Mulroy et al. 2019). The seemingly inconsistent results can be mitigated by redshift evolution of the slope (Golden-Marx & Miller 2019). We note, as quantified by Pillepich et al. (2018b), that both the slope and scatter of the stellar mass-halo mass (SMHM) relations depend sensitively on the operational definition of stellar mass. For example, different choices of aperture for the $M_{\star,\mathrm{BCG}}$ calculation results in slopes varying from ∼0.5 to 0.75 in TNG300, and also in TNG100, the 100 Mpc box simulation from the TNG suite (see table 4 of Pillepich et al. 2018b).

The scatter in the BCG stellar mass relation lies in the range 0.3−0.4 for all simulations over the entire mass range. This range encompasses the empirical value of 0.39 ± 0.07 found by Kravtsov et al. (2018) but it is lower than the value of 0.5 derived from the UniverseMachine analysis of Bradshaw et al. (2020).

Consistency with Bradshaw et al. (2020) is found for the scatter in total stellar mass. The simulations also find that the scatter in $M_{\star,\mathrm{tot}}$ is below 10 per cent for haloes above $\sim 10^{14.3}$ $M_\odot$. This supports the observational finding that total K-band luminosity ($\propto M_{\star,\mathrm{tot}}$) is a tight mass proxy for high-mass clusters (Mulroy et al. 2014, 2019; Chiu et al. 2016) selected by X-ray flux.

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

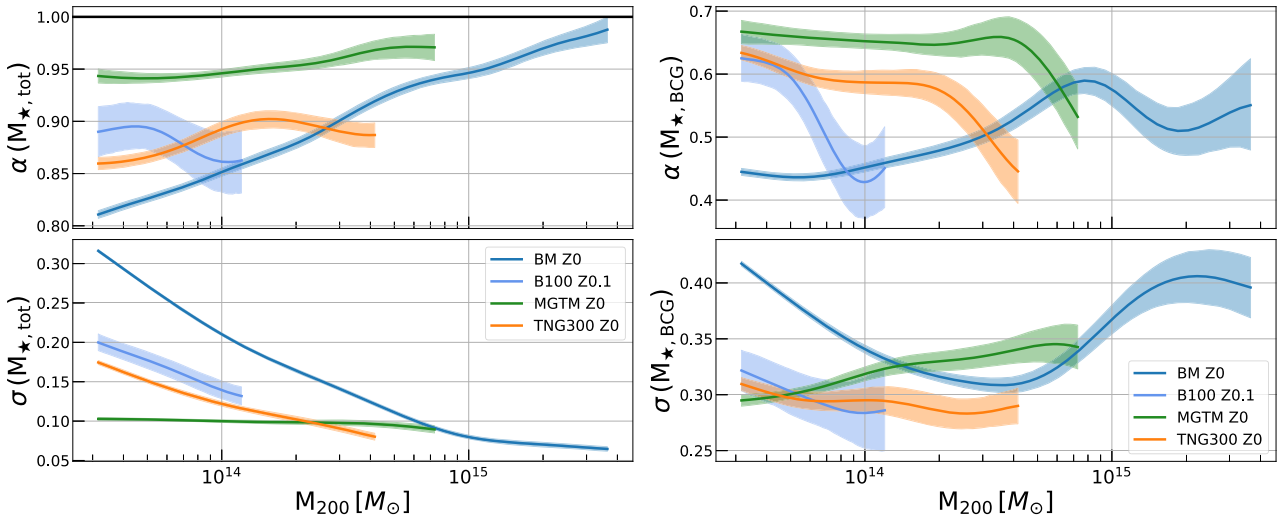

Figure 4. LLR parameters (slope, α, and natural log scatter σ) for the $M_{\star,\rm tot}$–$M_{\rm halo}$ (left) and $M_{\star,\rm BCG}$- – $M_{\rm halo}$ (right) scaling relations for all simulations at $z = 0$. The underlying data for $M_{\star,\rm tot}$ and $M_{\star,\rm BCG}$ in each simulation are shown in Figs A1 and A2, respectively.

We note that the MGTM simulation finds no trend of scatter with halo mass while the other two simulations, along with the UniverseMachine study, find that the scatter in total stellar mass monotonically decreases with increasing halo mass. In MGTM, the constant scatter in total stellar mass arises from compensating effects of central and satellite galaxy contributions. At low halo masses, the scatter in the central galaxy stellar mass is lower than at high masses, but the scatter in satellite count is higher, leaving the scatter in total stellar mass nearly constant.

4.2 Kernel shapes of normalized residuals

Forward modelling counts of massive haloes as a function of an observable property, s, is sensitive to the assumed shape of the conditional kernel, $p(s \mid M, z)$ (e.g. Shaw, Holder & Dudley 2010; Erickson, Cunha & Evrard 2011; Costanzi et al. 2019). Moving beyond the mild tensions in stellar property variance seen above, we seek here to test whether consistent kernel shapes emerge from the different simulation treatments.

While the complex coupling of physical and numerical factors makes it difficult to prove that kernel shapes should be preserved under transformations in the simulation control space, we take a purely empirical approach and simply ask whether consistent forms emerge. While lack of consistency in kernel shapes would be troubling, we caution that consistency is a necessary, but not sufficient, condition that the simulations have converged to the true form.

4.2.1 Satellite galaxy count kernel

At halo masses near our cut-off value of $10^{13.5}\,{\rm M}_\odot$, the mean occupation number is less than 10 and the fractional scatter is larger than ∼0.4. The odds of encountering a halo with $N_{\rm sat} = 0$ is therefore non-negligible. The BM simulation has 72 such satellite-empty haloes, a higher frequency than MGTM and TNG300, both of which have 5. For BM, we therefore introduce a cut-off of $10^{13.8}\,{\rm M}_\odot$ when we extract the normalized residual kernel. Above this modified mass, we find only two haloes with zero satellites. The skewness

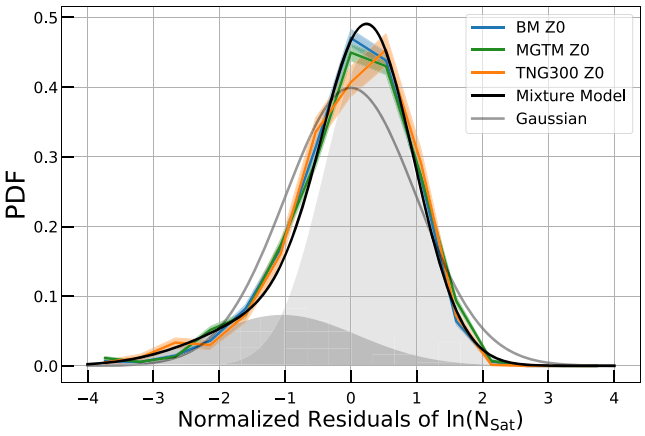

Figure 5. Mean mass-conditioned kernel – PDF of the normalized residuals, equation (8) – for $\ln N_{\rm sat}$ at $z = 0$ using a lower halo mass limit of $10^{13.8}\,{\rm M}_\odot$ for BM (to avoid discreteness features from the lowest $N_{\rm sat}$ haloes, see text) and $10^{13.5}\,{\rm M}_\odot$ for the other two simulations. A Gaussian mixture model (bold black line, with two components shaded) provides a much better fit than a single Gaussian (grey line). GMM fit parameters for the combined $z \leq 1$ samples across all simulations are given in Table 4 and individual sample values in Table B1.

and other statistics are not strongly affected by this choice.[3] In all simulations, the small number of empty haloes are removed before producing LLR statistics.

The PDF of the normalized, mass-conditioned residuals, equation (8), in $\ln N_{\rm sat}$, defines the kernel shown for the $z = 0$ halo populations in Fig. 5. Confidence bands are constructed from 1000 bootstrap realizations of the samples, through computing 1σ confidence intervals for the probability in each normalized residual bin.

[3] An alternative approach where we smooth the occupation count by adding random deviates in the range ±0.5 to $N_{\rm sat}$ produces a similar effect of removing the discreteness feature.

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

The three simulation populations exhibit very similar kernel shapes evident in their overlapping contours. The standard normal, $G(x, 1)$,[4] shown as the grey line, is not a good description of this left-leaning distribution, which has a normalized skewness value of $\gamma \equiv \mathrm{E}[(\frac{x-\mu}{\sigma})^3] = -0.9$.

A natural extension, one that enables efficient calculation with the analytical framework of Evrard et al. (2014), is a two-component Gaussian mixture

$$\Pr(x) = f_1\, G(x - \mu_1, \sigma_1) + (1 - f_1)\, G(x - \mu_2, \sigma_2), \tag{9}$$

where f_1 is the weight of a Gaussian with mean, μ_1, and standard deviation, σ_1, and $1 - f_1$ the weight of a second Gaussian with mean, μ_2, and standard deviation, σ_2.

The GMM result, shown as the bold black line with individual components as grey, shaded regions, provides a good fit to the reduced population statistics of all three simulations. Performing a Bayesian Information Criterion (BIC) test confirms that a two-component model is the most optimal at replicating this distribution. Increasing the number of components in the GMM is not supported by the BIC criteria and adds no significant improvement in fitting the residuals. We considered an alternative fit using an Edgeworth expansion (Shaw et al. 2010) but the mixture model is preferred because, unlike the Edgeworth expansion, it guarantees a positive-definite probability distribution.

The GMM fit parameters and skewness measures are presented in Section 4.4, where we explore the dependence of this shape on redshift. The dominant component of the GMM, representing 80 per cent of the halo population, is centred near 0.3 and has normalized variance of $(0.68)^2$, smaller than the complete population by a factor of 2. The remaining one-fifth of the population is centred near -1 with variance $(1.13)^2$. Given the important role of variance in cluster cosmological applications (e.g. Allen et al. 2011), empirical methods to separate these populations could yield significant benefits.

4.2.2 *Total and central stellar mass kernels*

Unlike the skewed residuals in satellite galaxy count, the residuals in total stellar mass within R_{200c}, shown in the top panel of Fig. 6, are much closer to Gaussian in shape. The skewness is consistent with zero for TNG300 and MGTM, and the BM value of $\gamma = -0.24 \pm 0.02$ at $z = 0$ value is much smaller than the -0.9 value displayed by the satellite galaxy counts. The Gaussianity of the residuals in $\ln M_{\star,\mathrm{tot}}$ for TNG300 and MGTM confirm the same result found previously for the BM sample only by Farahi et al. (2018).

When a system is subject to many random multiplicative factors, the central limit theorem argues for a kernel shape that is lognormal, or Gaussian in log-space (Adams & Fatuzzo 1996). We postulate that the formation of individual stellar particles in cosmological simulations is dictated by such multiplicative factors, while the aggregated effort to form many individual star particles into a single galaxy entails fewer effective degrees of freedom and so can deviate more from log-normality.

The kernel shape for central galaxy stellar mass, shown in the bottom panel of Fig. 6, is also negatively skewed. The shapes for TNG300 and MGTM show good agreement, with $\gamma \simeq -0.8$, while BM tends closer to log-normality, with $\gamma \simeq -0.3$. Unfortunately, the shape of the kernel from the higher resolution B100 run is not well defined because of the much smaller sample of haloes available in that simulation. Given the low stellar particle resolution of BM, we tentatively promote the -0.8 value as more likely, but defer more careful analysis to future simulations with higher resolution.

[4]Using notation $G(x - \mu, \sigma)\,\mathrm{d}x = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(x-\mu)^2}{2\sigma^2})\,\mathrm{d}x$.

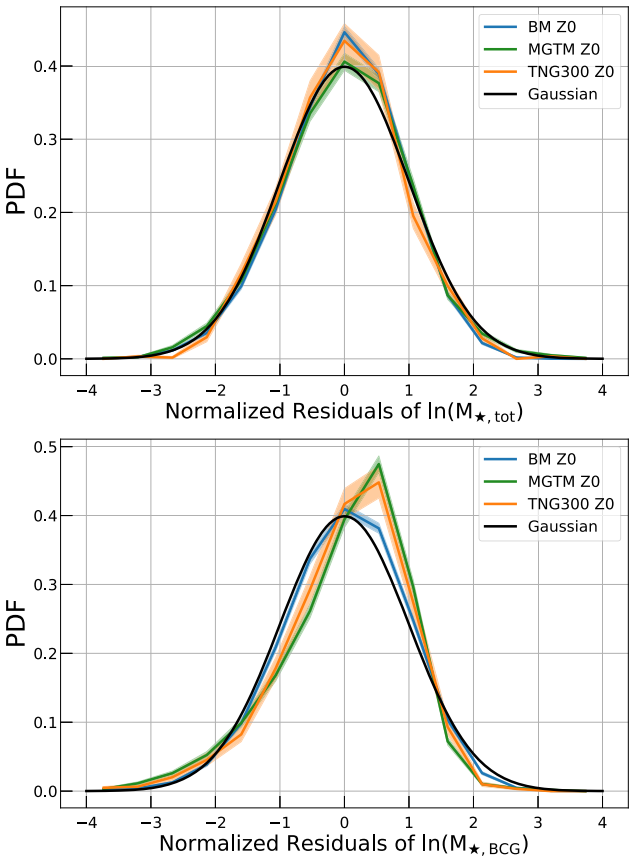

Figure 6. Mean mass-conditioned kernels for $\ln M_{\star,\mathrm{tot}}$ (top) and $\ln M_{\star,\mathrm{BCG}}$ (bottom) at $z = 0$ for haloes with mass above $10^{13.5}\,\mathrm{M}_\odot$ in all simulations. The solid black line shows the standard normal distribution.

4.3 Mass-conditioned correlations in stellar properties

There are physical reasons to anticipate correlations between elements of the stellar property vector conditioned on total halo mass. The tidal disruption and accretion of satellite galaxy stellar material on to a halo's central galaxy is a well-known mechanism for producing an anticorrelation between N_{sat} and $M_{\star,\mathrm{BCG}}$.

It has long been postulated that central galaxies in groups and clusters increase their stellar mass via mergers (e.g. Tremaine & Richstone 1977). Recent high-resolution hydrodynamical simulations (Cui et al. 2014; Bahé et al. 2017; Bahé et al. 2019; Tremmel et al. 2019) and semi-analytic models (Croton et al. 2006; De Lucia & Blaizot 2007; Bradshaw et al. 2020) confirm the growth of centrals at the expense of satellites.

The Hydrangea suite of simulations (Bahé et al. 2019) find that, across the mass range we study here, the majority of galaxies accreted at $z < 2$ do not survive to the present. While some caution is required, as even the highest resolution studies may not yet be converged (van den Bosch & Ogiya 2018, but see appendix of Bahé et al. 2019), observational studies also support the growth of central galaxies and intracluster light over time (Zhang et al.

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

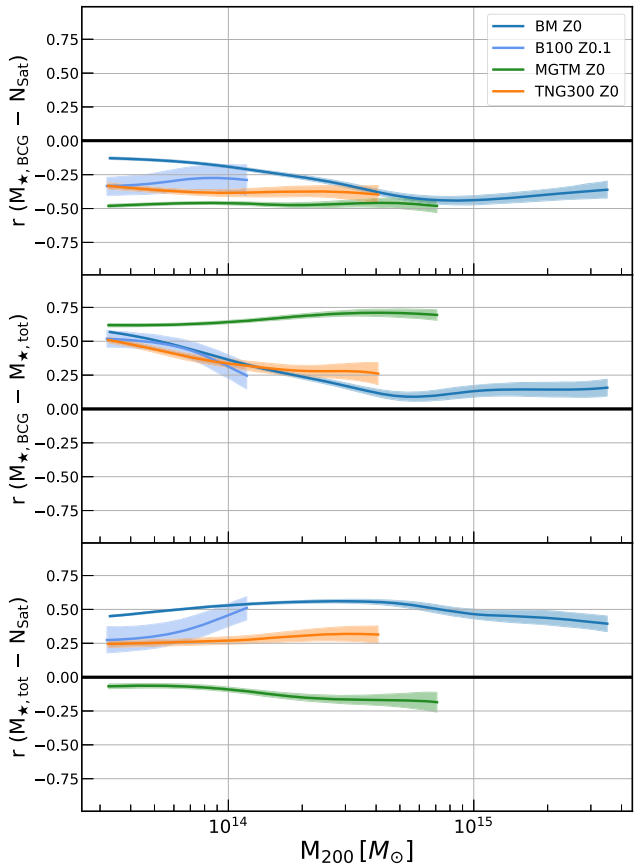

Figure 7. Correlation coefficients, at fixed $M_{\rm halo}$, of $N_{\rm sat}$ and $M_{\star,\rm BCG}$ (top), $M_{\star,\rm tot}$ and $M_{\star,\rm BCG}$ (middle), and $M_{\star,\rm tot}$ and $N_{\rm sat}$ (bottom).

2016; Tang et al. 2018; Zhang et al. 2019). In the Illustris and TNG300 simulations, galaxies with stellar mass greater than a few $10^{11}\,{\rm M}_\odot$ are mostly made of *ex situ* stars accumulated through the merging and accretion of material from other galaxies (see Rodriguez-Gomez et al. 2016; Pillepich et al. 2018b, and references therein).

As a result of this dynamical processing, all simulations display mildly anticorrelated behaviour between $N_{\rm sat}$ and $M_{\star,\rm BCG}$ at $z = 0$. The top panel of Fig. 7 shows that correlation coefficient, $r \simeq -0.4$, is nearly independent of halo mass in all realizations. The smaller values of the BM correlation at low halo masses are driven by the larger variance in $\ln N_{\rm sat}$ seen in that run. The high-resolution B100 model, with lower scatter in $N_{\rm sat}$ compared to BM, yields a larger correlation coefficient consistent with the values seen in the TNG300 and MGTM.

The correlation coefficients of the remaining pairs are sensitive to the central galaxy stellar mass statistics, particularly the normalization and scatter, and the MGTM central galaxy population is extreme in both measures. As a result, the mass-conditioned correlation coefficients of $M_{\star,\rm tot}$ and $M_{\star,\rm BCG}$, as well as $N_{\rm sat}$ and $M_{\star,\rm tot}$ (middle and lower panels of Fig. 7, respectively) show behaviours for which MGTM differs from the others.

For the pairing of $M_{\star,\rm tot}$ and $M_{\star,\rm BCG}$ at fixed halo mass, stronger correlation is seen in MGTM because that simulation has a very low scatter in total stellar mass, making the role of central galaxy variations more prominent. The central galaxies in MGTM also contribute the largest fraction of total stellar mass. At $10^{14}\,{\rm M}_\odot$, the stellar mass fraction of the central galaxy, $\frac{M_{\star,\rm BCG}}{M_{\star,\rm tot}}$, in the simulations are 0.42 (MGTM), 0.40 (TNG300), 0.35 (B100), and 0.27 (BM).

One would reasonably anticipate a positive correlation between $N_{\rm sat}$ and $M_{\star,\rm tot}$ at fixed $M_{\rm halo}$, as haloes with more satellite galaxies should also have a larger total stellar mass. The bottom panel of Fig. 7 shows that the BM, B100, and TNG300 simulations follow that expectation, albeit with somewhat different magnitudes between 0.25 and 0.5. The MGTM simulation, however, exhibits a weak *anticorrelation* between these two properties. This counter-intuitive result is explained by the non-Gaussian scatter in the full space of residuals that we examine next.

4.3.1 Non-Gaussian features in residual space

Except for total stellar mass, the one-dimensional kernels in Figs 5 and 6 display non-Gaussian features that we speculate could be related to merger tree entropy, which also displays a skew normal form (Obreschkow et al. 2020). To expand the view into the full 3D space of residuals, Fig. 8 shows normalized residuals in $N_{\rm sat}$ and $M_{\star,\rm tot}$ for each halo coloured by its $M_{\star,\rm BCG}$ residuals.

It is evident by eye that the three simulations exhibit somewhat different forms. The BM and TNG300 residuals show positive correlation in $N_{\rm sat}$ and $M_{\star,\rm tot}$ with shapes that are approximately elliptical. In particular, all haloes with extreme low satellite galaxy counts (given a halo mass) also have low total stellar mass. The lower right quadrant, corresponding to haloes with low numbers of satellites but high total stellar mass, is relatively vacant.

The MGTM residuals, in contrast, include a few outlying points in this lower-right quandrant, and it is these systems that drive the weak anticorrelation between $N_{\rm sat}$ and $M_{\star,\rm tot}$ seen for this simulation in the lower panel of Fig. 7.

Note that the applied point colours change in the same manner in all the simulations, with low to high $M_{\star,\rm BCG}$ residuals running from the top-left to the bottom-right. This pattern reflects the rough agreement of the correlations involving $M_{\star,\rm BCG}$ shown in the top two panels of Fig. 7.

4.4 Redshift behaviour and low-z GMM $N_{\rm sat}$ fit

Returning to the issue of kernel shapes, we find that the skewness in the $N_{\rm sat}$ and $M_{\star,\rm tot}$ kernel shapes varies little with redshifts $z \leq 1$. Fig. 9 shows these values at discrete redshifts for simulation samples with at least 300 haloes above $10^{13.5}\,{\rm M}_\odot$.

As noted previously, the BCG stellar mass (shown at $z = 0$ only) is skew negative with $\gamma \simeq -0.8$ in both MGTM and TNG300, while the value for BM is much smaller, -0.3. For B100 we find $\gamma = -0.71 \pm 0.26$, where the large error reflects the small sample size of <100 haloes.

The skewness in total stellar mass varies across the simulations, but is limited to the range $[-0.3, 0.3]$. The TNG300 results are consistent with zero at all redshifts while zero values are found by MGTM and BM at some redshifts.

In contrast, the skewness in $\ln N_{\rm sat}$ is both consistent across simulations and persistent in redshift. Having verified robustness of the $N_{\rm sat}$ kernel shape, we combine the $z = 0$, 0.5, and 1 redshift samples from all three simulations into a superset ensemble of more than 26 000 haloes. Parameters from this superset, given in Table 4, are precisely constrained by this large halo ensemble, with statistical uncertainties of a few per cent in most parameters.

Beyond skewness, there is also good agreement in the GMM parameters of the $N_{\rm sat}$ kernel, shown for the three different

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

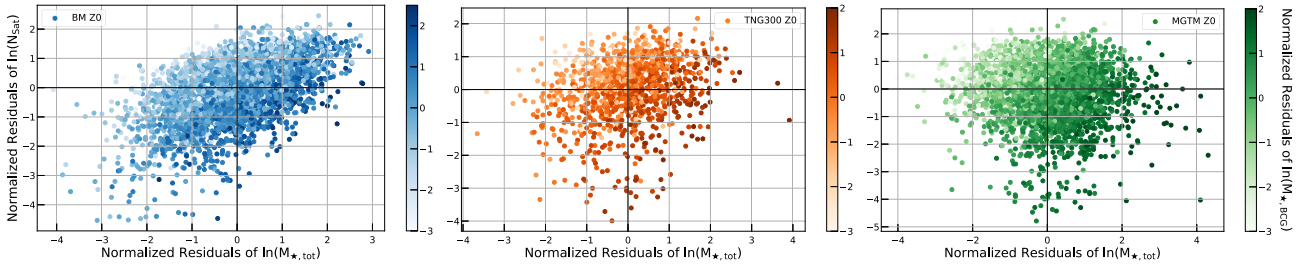

Figure 8. Normalized residuals of $N_{\rm sat}$ and $M_{\star,\rm tot}$ coloured by the $M_{\star,\rm BCG}$ residuals for all haloes at $z = 0$. The BM sample uses a halo mass threshold of $10^{13.8}\,{\rm M}_\odot$ while the others use $10^{13.5}\,{\rm M}_\odot$.

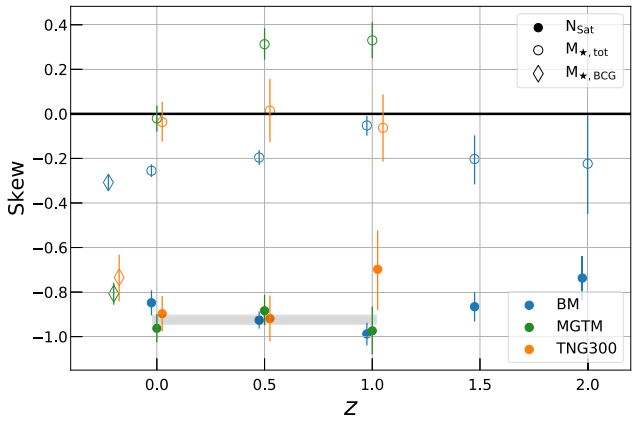

Figure 9. Skewness of the normalized, mass-conditioned residuals in $N_{\rm sat}$ (solid circles) and $M_{\star,\rm tot}$ (open circles) for samples with at least 300 haloes above $10^{13.5}\,{\rm M}_\odot$ at each redshift. The $M_{\star,\rm BCG}$ skewness (open diamond) is shown only at $z = 0$, offset by -0.05 in redshift to improve readability. The grey band is a 68 per cent interval for the $N_{\rm sat}$ skewness obtained using the combined halo populations of all three simulations at $z = 0$, 0.5, and 1.

Table 4. Satellite galaxy kernel skewness, γ, and GMM fit parameters for the combined $z = 0$, 0.5, 1 halo populations of all three simulations above a halo mass of $10^{13.5}\,{\rm M}_\odot$ ($10^{13.8}\,{\rm M}_\odot$ for BM at $z = 0$ only), with uncertainties from bootstrap (skewness) and MCMC posterior sampling (GMM parameters). The superset contains a total of 26 332 haloes.

Parameter	Value
γ	-0.91 ± 0.02
f_1	0.79 ± 0.01
μ_1	0.28 ± 0.01
μ_2	-1.04 ± 0.05
σ_1	0.68 ± 0.01
σ_2	1.13 ± 0.02

simulations as a function of redshift in Fig. 10. The superset sample values of Table 4, shown as the grey bands in Fig. 10, indicate that the halo population consists of an 80 per cent majority with mean 0.28 ± 0.01 and dispersion 0.68 ± 0.01 along with a wider, left-leaning minority having mean -1.04 ± 0.05 and scatter 1.13 ± 0.02. In the next section, we use importance sampling to trace how these components map to different distributions in central galaxy stellar mass.

Due to the steepness of the cosmic mass function, the shape of the $N_{\rm sat}$ kernel is heavily weighted by haloes near the cut-off mass scale of $10^{13.5}\,{\rm M}_\odot$. In Fig. 11, we show how the skewness runs

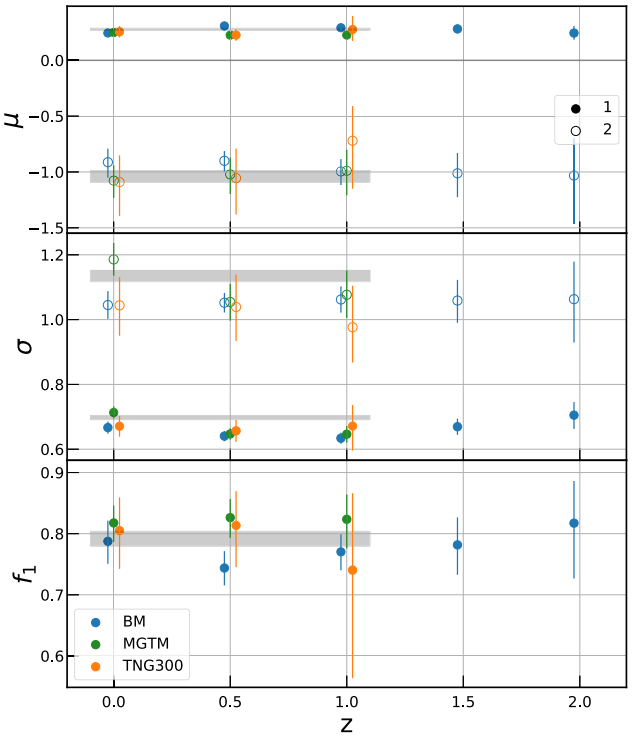

Figure 10. Gaussian mixture model parameters for the $N_{\rm sat}$ kernel, equation (9), with solid and open circles giving the first and second components, respectively. Values for each simulation are given in Table B1. Grey bands are 68 per cent confidence bands derived from the superset of all simulation populations at $z = 0$, 0.5 and 1 (see Table 4).

with applied cut-off mass, up to a limit for each sample at which the number of haloes falls below 300. While somewhat arbitrary, the 68 per cent bootstrap uncertainties for smaller samples become large and the results uninformative.

At $10^{14}\,{\rm M}_\odot$, all simulations show that the shape is somewhat less skewed, with $\gamma \sim -0.75$. At higher masses, we rely solely on the massive MACSIS sample, which displays asymptotic behaviour to $\sim -0.65 \pm 0.15$. We leave it to future work with larger simulation ensembles to address this question in more detail.

5 TOWARDS SURVEY VALIDATION: SECONDARY SELECTION EFFECTS

In this section, we explore how secondary selection in BCG stellar mass affects the statistical properties of satellite count and total stellar mass. For the TNG300 simulation, we also explore secondary selection in halo formation epoch.

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

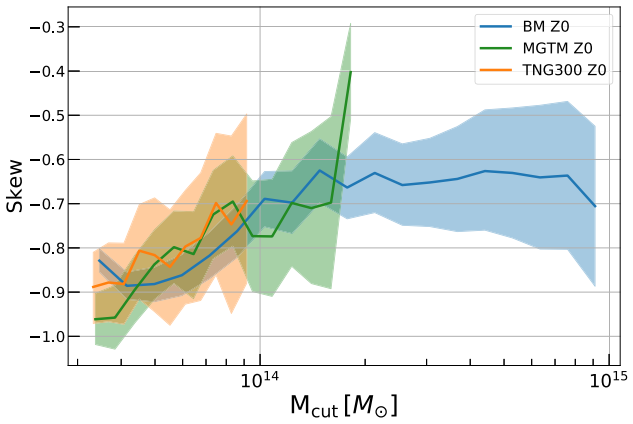

Figure 11. Skew in the $\ln N_{sat}$ kernel as a function of lower mass limit, M_{cut}. The maximum M_{cut} value is set by a requirement that samples contain at least 300 haloes.

Application of secondary selection to cluster surveys requires a statistical model relating cluster properties, especially those involved in selection, to our primary selection variable. The consistent patterns exhibited by these simulations are testable with current surveys when selection and projection effects are properly included. A preparatory step towards sample modelling could be to use synthetic *Chandra* and XMM observations of the simulated halo ensembles (Biffi et al. 2012; Le Brun et al. 2014; Koulouridis et al. 2018; ZuHone et al. 2018) to explore expectations for cluster samples selected by core-excised X-ray flux (Mantz et al. 2018).

The ultimate aim is to *validate* these expectations in observed cluster samples with high-quality, uniform optical properties, such as SDSS (York et al. 2000), DES (The Dark Energy Survey Collaboration 2005) and, in the future, LSST (Ivezić et al. 2019) and Euclid (Laureijs et al. 2011; Racca et al. 2016). Such validation will require an observable mass proxy, such as weak lensing mass or hot gas mass, that itself is likely to correlate with the stellar properties under consideration (Wu et al. 2015; Farahi et al. 2018, 2019b).

5.1 Secondary selection on $M_{\star,BCG}$

The correlation structure in the top two panels of Fig. 7 provides a lever arm for secondary selection in BCG stellar mass. We first explore this structure using the mixture model in satellite galaxy count, N_{sat}.

The top panel of Fig. 12 displays the residual correlations of $M_{\star,BCG}$ and N_{sat} for the $z = 0$ TNG300 halo population (results are similar for the other two simulations). This panel is merely a rotated version of the middle panel in Fig. 8. An anticorrelation is apparent, with non-Gaussian tails in both directions. As shown below, the tail to low central galaxy stellar mass is associated with late-forming systems.

The lower panel of Fig. 12 illustrates the utility of the Gaussian mixture model for N_{sat} to stratify the halo population in a property correlated with it. The two shaded regions shown in the lower panel of Fig. 12 are built from importance sampling the two components of the mixture whose parameters are given in Table 4. Haloes associated with the minority component, the broad tail displaced to lower N_{sat}, possess central galaxies with higher $M_{\star,BCG}$ values (mean of 0.4 and width of 0.7) shown by the darker shaded component (consistent with the darker shaded component of Fig. 5). This region overlaps with the dominant, lighter shaded component that has a

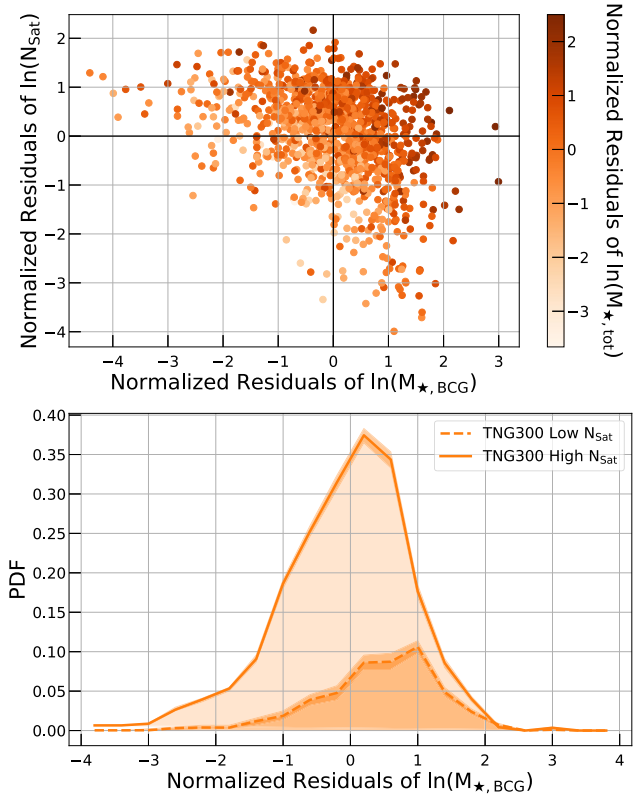

Figure 12. Mixture model demonstration for the TNG300 halo population at $z = 0$. *Upper:* Distribution of normalized residuals in the $\ln M_{\star,BCG} - \ln N_{sat}$ plane, coloured by the residual in $\ln M_{\star,tot}$. *Lower:* Importance-sampling of the two GMM components in satellite galaxy count produces offset kernels in $\ln M_{\star,BCG}$. The darker shading at high $M_{\star,BCG}$ corresponds to the darker shaded GMM component at low N_{sat} in Fig. 5, and vice versa for the light shading, as anticipated by the residuals above.

mean of −0.09 and standard deviation 0.99. Both components are skewed negative with similar values of $\gamma = -0.75$, reflecting the non-Gaussian structure of the residuals in both components.

This structure implies that sub-samples of haloes with lower than average (given their halo mass) central galaxy stellar masses, those below roughly -2σ, are comprised almost exclusively of the dominant component in N_{sat}. Haloes with higher than average central galaxy stellar masses, in contrast, consist of an N_{sat} mixture in which the minority component is enhanced but not dominant.

We can expand on this result by treating $M_{\star,BCG}$ as a secondary selection variable (total halo mass being the primary selection). We divide each simulation sample into two subsets based on whether a halo's central galaxy stellar mass lies above or below the LLR mean expectation at that halo mass, meaning we are selecting secondarily on the sign of the residual (equation 4). Note that, due to the non-Gaussian shape of the $M_{\star,BCG}$ kernel, this does not split the population into sub-samples of equal size.

Fig. 13 shows cumulative distribution functions (CDFs) for N_{sat} (top) and $M_{\star,tot}$ (bottom) for the full population (dotted lines) as well as the high and low $M_{\star,BCG}$ halo subsets. Because the scatter in N_{sat} for the BM simulations is spuriously enhanced by its coarse stellar mass resolution (see Figs 3 and 7), we do not show that model in the top panel.

Secondary selection by $M_{\star,BCG}$ generates fairly dramatic shifts in the CDFs of both $\ln N_{sat}$ and $\ln M_{\star,tot}$. Haloes with lower than

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

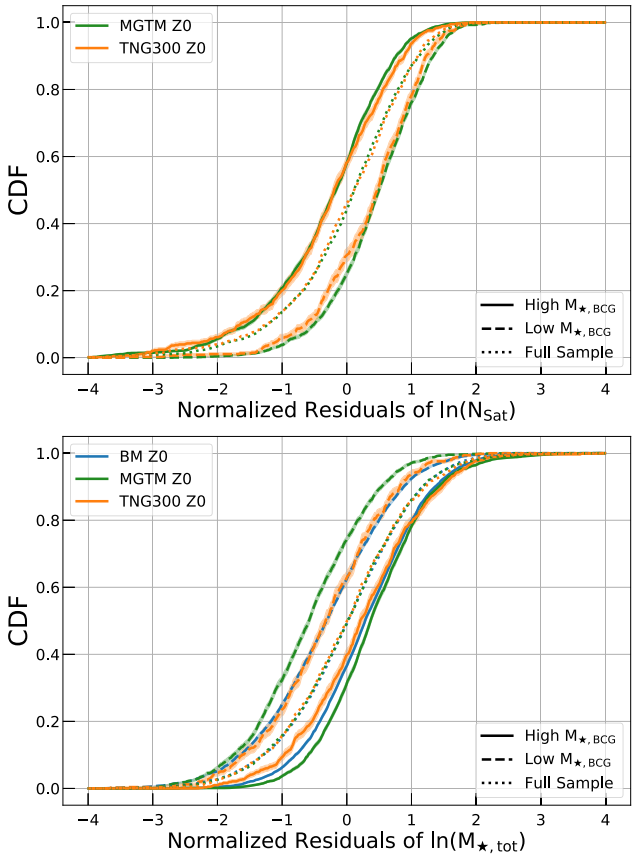

Figure 13. Cumulative distribution functions for residuals in $\ln N_{\rm sat}$ (upper) and $\ln M_{\star,\rm tot}$ (lower) conditioned on relative $M_{\star,\rm BCG}$ and total halo mass. Solid lines show haloes with higher than LLR-averaged $M_{\star,\rm BCG}$ while dashed lines show the opposite. Dotted lines show the behaviour of the full population, unconditioned on $M_{\star,\rm BCG}$. The BM sample is omitted from the top plot (see text).

Table 5. Kernel offsets, the shifts in normalized cumulants shown as the dashed and solid lines in Figs 13 and 14, for sub-samples split by $M_{\star,\rm BCG}$ and (for TNG300 only) $z_{\rm form}$.

Sample (selection)	$N_{\rm sat}$		$M_{\star,\rm tot}$		$M_{\star,\rm BCG}$	
	rms	max	rms	max	rms	max
BM ($M_{\star,\rm BCG}$)	(0.30)	(0.39)	0.63	0.86	–	–
MGTM ($M_{\star,\rm BCG}$)	0.72	1.03	0.98	1.16	–	–
TNG300 ($M_{\star,\rm BCG}$)	0.63	1.02	0.57	0.65	–	–
TNG300 ($z_{\rm form}$)	0.41	0.59	0.18	0.39	0.79	1.22

average BCG stellar masses tend to have both higher numbers of satellite galaxies as well as lower total stellar masses. Table 5 lists root mean square and maximum values of the CDF offsets, $|x_2 - x_1|$, in normalized $N_{\rm sat}$ or $M_{\star,\rm tot}$ deviation, where x is the cumulant location at which the integrated probability takes some fixed value, $\mathrm{CDF}(x_2) = \mathrm{CDF}(x_1) = \text{constant}$. To minimize discreteness effects in the rare event tails of these distributions, values in the table are limited to CDF values in the range (0.1, 0.9).

The rms values, in Table 5 give us the same information as the correlations, as evidenced by the simulation ordering of the cumulant shifts reflecting the ordering of the $M_{\star,\rm BCG}$–$N_{\rm sat}$ and $M_{\star,\rm BCG}$–$M_{\star,\rm tot}$ correlations in Fig. 7. For residual CDFs in both $\ln N_{\rm sat}$ and $\ln M_{\star,\rm tot}$, MGTM shows the largest rms offset, and its correlations also have the largest magnitude. For the residual CDFs of $\ln M_{\star,\rm tot}$, BM and TNG300 show similar deviations, since their correlations are in agreement for a large part of the halo mass range.

5.2 Secondary selection on formation epoch, $z_{\rm form}$

The mass-conditioned covariance among stellar properties and non-Gaussian kernel shapes in $N_{\rm sat}$ and $M_{\star,\rm BCG}$ are related to the formation histories sampled by these discrete halo populations. For example, Bradshaw et al. (2020) use the semi-analytic UniverseMachine model (Behroozi et al. 2019) to demonstrate that relative BCG stellar mass is correlated with the age of a halo, while total stellar mass is nearly independent of age.

We examine this behaviour for TNG300, the simulation for which data to derive formation time estimates are publicly available. Analysing the merger tree of each $z = 0$ halo, we define the formation redshift, $z_{\rm form}$, as the epoch at which the total mass of a halo falls to half of its final value. After LLR fitting $z_{\rm form}$ versus halo mass, we condition the residuals of $\ln N_{\rm sat}$, $\ln M_{\star,\rm tot}$, and $\ln M_{\star,\rm BCG}$ on the sign of the $z_{\rm form}$ residuals.

In Fig. 14, we present the residual CDFs for the two subpopulations in all three stellar properties. For the upper two panels, showing $N_{\rm sat}$ and $M_{\star,\rm tot}$, we compare to secondary conditioning using $M_{\star,\rm BCG}$, presented above.

The $N_{\rm sat}$ CDF (top panel) shows similar deviations when conditioned on either $M_{\star,\rm BCG}$ or $z_{\rm form}$. Namely, haloes of a younger age (i.e. with lower than average $z_{\rm form}$) and with lower central galaxy stellar mass are surrounded by a larger number of satellite galaxies. Our results agree with those of Bose et al. (2019), who find a split in the $N_{\rm sat}$ scaling relations when conditioning on $z_{\rm form}$ for a TNG sample spanning a wider range in halo mass.

In contrast, the $M_{\star,\rm tot}$ CDF (middle panel) shows differences between the two secondary selection variables. The positive correlation with $M_{\star,\rm BCG}$ produces a shift of 0.57 in the $M_{\star,\rm tot}$ CDF split by central galaxy stellar mass. A much weaker correlation with $z_{\rm form}$ yields a smaller shift of 0.16 in the $M_{\star,\rm tot}$ CDF. The latter finding is in qualitative agreement with Bradshaw et al. (2020), who find no offset in the SMHM relation when conditioning on $z_{\rm form}$.

The central galaxy stellar mass (bottom panel of Fig. 14) is most sensitive to formation history, with an 0.8σ rms shift in the CDF. The tail below -1σ in $M_{\star,\rm BCG}$ is almost exclusively late-forming haloes. These shifts are again in qualitative agreement with Bradshaw et al. (2020), who find a difference of 0.2 dex in the $M_{\star,\rm BCG}$–$M_{\rm halo}$ relation between the top and bottom 20 per cent ranked haloes in $z_{\rm form}$.

Fig. 15 provides a view of how formation redshift maps on to the space of residuals in total and BCG stellar masses in the TNG300 population. Loci of constant formation time are oriented roughly along the diagonal, with early-forming systems tending to have brighter than average BCGs overall but with total stellar masses that span the full range above and below the mean.

Other hydrodynamic simulation studies have also studied the satellite galaxy HOD conditioned on $z_{\rm form}$, finding that younger haloes have preferentially more satellites (Artale et al. 2018; Bose et al. 2019). This feature they explain by older haloes losing more satellites to mergers with central galaxies, an interpretation supported by Matthee et al. (2017) and Bradshaw et al. (2020).

Bradshaw et al. (2020) also find that the *ex situ* stellar mass, material obtained through mergers with satellites, correlates strongly with $z_{\rm form}$, with central galaxies in older haloes containing more *ex*

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

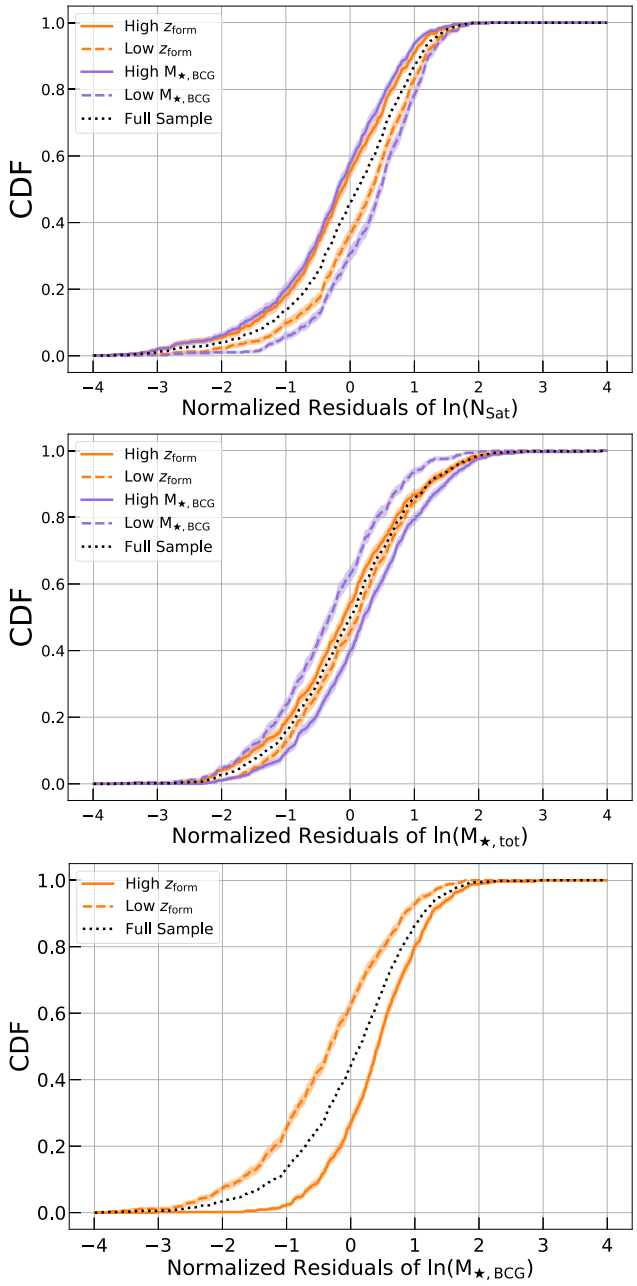

Figure 14. For the TNG300 $z = 0$ sample, cumulative distribution functions for satellite galaxy counts (top) and total stellar mass (middle) conditioned on halo formation time, $z_{\rm form}$ (orange) and $M_{\star,\rm BCG}$ (purple). The bottom panel shows the residual CDFs of the central galaxy stellar mass conditioned on $z_{\rm form}$.

situ stellar mass. Conversely, the *in situ* stellar mass shows no $z_{\rm form}$ dependence.

While halo formation times are not directly observable, other properties, such as the magnitude gap, can serve as a proxy. Farahi, Ho & Trac (2020) use the same TNG300 simulation as us to study the secondary conditioning of multiple halo properties – including the three we investigate here – on the magnitude gap. Their results are qualitatively in agreement with ours, showing magnitude gap-dependent separations in the $N_{\rm sat}$- and $M_{\star,\rm BCG}$–$M_{\rm halo}$ scaling relations, and little to no striation in the $M_{\star,\rm tot}$–$M_{\rm halo}$ scaling relation.

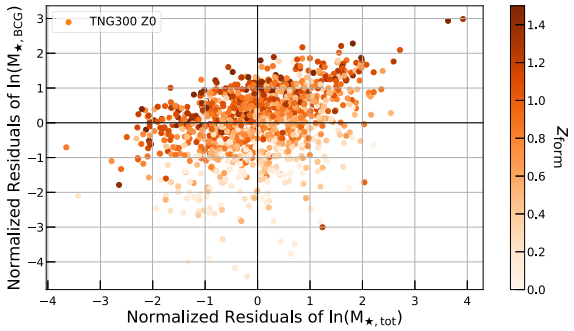

Figure 15. Normalized residuals in the $M_{\star,\rm tot}$–$M_{\star,\rm BCG}$ plane coloured by $z_{\rm form}$ for the TNG300 $z = 0$ halo sample.

6 DISCUSSION

6.1 Property extensions and scale dependence of halo population statistics

The capacity of cosmological hydrodynamical simulations has expanded to the point where multiple simulation methods can produce halo populations containing many millions of objects across the full range of resolved scales, from individual galaxies to rich clusters of galaxies. Future work can investigate the degree of self-similarity of our findings by considering a wider range in halo mass. The sensitivity of a halo's properties to its merger tree entropy may yield new insights (Obreschkow et al. 2020).

In general, integrated physical properties connect to halo mass and redshift in a way that combines secular scaling of mean behaviour with stochastic variations for individual objects. We intend to expand our study to include more features, such as hot gas masses, X-ray temperatures and luminosities, and galaxy velocities. For example, Farahi et al. (2018) find that the joint kernel of hot gas mass and total stellar mass is well described by a simple two-dimensional Gaussian with moderate anticorrelation between these mass fractions. The anticorrelation reflects the fact that, compared to the $10^{12}\,{\rm M}_{\odot}$ haloes hosting individual bright galaxies that lose considerable fraction of their baryons (Bregman et al. 2018), the cosmic mix of baryons is more nearly retained within the deep gravitational potential wells of clusters. Verifying this and other multiphase signatures in TNG and MGTM solutions remains to be done.

Our study is focused on verifying stellar property statistics for massive halo populations derived from simulations that employ multiple cosmological hydrodynamics methods. However, other approaches, particularly semi-analytic methods (SAMs) that assign galaxy properties to N-body haloes and sub-haloes in a manner constrained by empirical data, are also capable of producing population-level expectations. The qualitative agreement we find with Bradshaw et al. (2020) on property correlations, and other features such as running of $N_{\rm sat}$ scatter with mass, should be investigated more carefully to gain insights into the robustness of our findings and the limitations of both SAM and full hydro methods.

An open question is whether SAM methods produce kernel shapes for satellite galaxy counts and central galaxy stellar mass that are negatively skewed to the degree seen in the three cosmological simulations studied here. It is worth noting that, compared to cosmological hydrodynamics simulations, SAM populations require significantly less computational time and energy to produce, meaning that sensitive kernel shape measurements may be possible

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

using large populations even at mass scales above $10^{15}\,M_\odot$. Direct comparison of SAM and hydro population statistics would be a preliminary step in this direction.

6.2 Implications for optical cluster survey likelihoods

The intrinsic satellite galaxy count is a core ingredient of models that use optical cluster counts in richness and redshift to constrain cosmology (Rozo et al. 2010; Chiu et al. 2019; Costanzi et al. 2019). However, due primarily to line-of-sight projection, the optical richness measured for a cluster is a noisy measure of the $1 + N_{\rm sat}$ occupation intrinsic to haloes (e.g. Cohn et al. 2007), an expectation confirmed by spectroscopic follow-up studies of SDSS redMaPPer clusters (Sohn et al. 2018).

Cluster richness can be modelled as a sum of partial contributions from haloes (Farahi et al. 2016), so future work may lead to likelihoods employing a mixture of mixtures, one mixture representing contributions of multiple haloes to the richness λ, the other mixture expressing intrinsic richness at a given halo mass (the one we study here).

We note that secondary selection to identify the narrow $N_{\rm sat}$ component is a potential boon to cosmological studies. Shifts in mean halo mass associated with property selection scale as the variance (Evrard et al. 2014), so the factor two smaller variance in satellite galaxy count at fixed halo mass for the narrow component (Table 4) could potentially be exploited to more accurately calibrate mean masses via stacked weak lensing analysis (e.g. McClintock et al. 2019).

6.3 Cyberinfrastructure improvements

Our study has been greatly enabled by the availability of full data releases of the IllustrisTNG simulations (Nelson et al. 2019), and a partial public release of Magneticum Pathfinder (Ragagnin et al. 2017). The democratization of the data, through the availability of uniform, catalogue-level simulation products to the public, is key to permitting more in-depth analyses of halo population statistics derived from multiple cosmological simulations.

Looking even further, reproducible computational science benefits from having open community access to the specific simulation methods, both production and analysis code bases, used in simulation studies (e.g. Stodden et al. 2016). The scale of simulation data volumes makes it difficult to move high-resolution data to a central location, but a future in which distributed, containerized analysis environments (Raddick et al. 2019) operate using improved discoverability standards (Languignon et al. 2017) could greatly simplify and empower verification studies of the type we perform here.

7 SUMMARY

Motivated by the need to verify the highly detailed solutions produced by complex cosmological simulations, we perform a statistical study of the stellar and satellite galaxy contents within halo populations produced by three large cosmological hydrodynamics simulations. These include simulated samples from BAHAMAS and MACSIS, a volume from the Magneticum suite, and TNG300 from the IllustrisTNG Project. We focus here on features that describe the galaxy–halo connection – the stellar-mass limited satellite galaxy occupation, $N_{\rm sat}$, the stellar mass of the central galaxy, $M_{\star,\rm BCG}$, and the total stellar mass within the halo, $M_{\star,\rm tot}$ – in haloes with total mass exceeding $10^{13.5}\,M_\odot$.

By applying a local linear regression method, we summarize the form of the mass-conditioned kernel, $\Pr(\ln \boldsymbol{S} \mid M_{\rm halo}, z)$, where $\boldsymbol{S}$ is the set of aforementioned stellar properties. We provide local regression fit parameters for these properties – normalizations, slopes, and covariance – sampled over roughly two orders of magnitude in halo mass at $z = 0$. These samples contain from 1000 to 9000 objects, and this large scale allows us, by marginalizing over halo mass, to analyse the mean shape of the conditional kernel for each stellar property. Our analysis effectively compresses a large amount of raw output data into a highly compact analytic form useful for modelling statistical likelihoods.

Our main results are summarized in the following points.

(i) We verify common kernel shapes for the satellite count, $N_{\rm sat}$, and total stellar mass, $M_{\star,\rm tot}$, where the former is skewed negatively, with skewness parameter $\gamma = -0.91 \pm 0.02$, and the latter being much closer to Gaussian ($|\gamma| < 0.3$). For the population of haloes above $10^{13.5}$ at $z \leq 1$, we provide a two-component Gaussian mixture fit to the $N_{\rm sat}$ kernel (Table 4), but note the shape is mildly sensitive to this halo mass threshold. The $z = 0$ kernel shape of $M_{\star,\rm BCG}$ is also negatively skewed, with MGTM and TNG300 finding skewness -0.8.

(ii) While the values of halo mass-conditioned regression parameters (slope, normalization, and covariance) are often in mild tension among the three simulations, we find areas of qualitative agreement: (i) the scatter in $\ln N_{\rm sat}$ depends on halo mass and is slightly super-Poissonion (Fig. 3); (ii) the scaling of total stellar mass is consistently sub-linear and the fractional scatter in $M_{\star,\rm tot}$ is below 10 per cent for haloes above $2 \times 10^{14}\,M_\odot$ (Fig. 4); (iii) the mass-conditioned residuals in $N_{\rm sat}$ and $M_{\star,\rm BCG}$ are anticorrelated while those in $M_{\star,\rm BCG}$ and $M_{\star,\rm tot}$ are positively correlated (Fig. 7).

(iii) All simulations find that haloes with bigger than average central galaxies have both fewer than average satellite galaxies and larger than average total stellar mass. The former is consistent with a picture in which earlier-forming systems cannibalize satellite galaxies to the benefit of the BCG.

(iv) For TNG300, the mass-conditioned formation redshift correlates strongly with $M_{\star,\rm BCG}$ and $N_{\rm sat}$ but weakly with $M_{\star,\rm tot}$.

(v) The structure of the mass-conditioned stellar property residuals is more similar in BM and TNG300 than MGTM (Fig. 8).

The low-order statistical measures of our study are empirically testable through careful analysis of scaling behaviour in large cluster samples. Sample selection, mis-centring, projection, and the evolution of galaxy colours are all confounding factors to be addressed in such studies.

With nearly one million haloes above $10^{14}\,M_\odot$ anticipated on the full sky (Allen et al. 2011), the demand for improved statistical representations linking cluster observable properties to those of the underlying halo population will only grow in the era of deep all-sky surveys across millimetre, optical-IR, and X-ray wavelengths.

ACKNOWLEDGEMENTS

AEE acknowledges support from NASA Chandra grants G08-19107B and TM9-20006X. AF is supported by a Michigan Institute for Data Science Fellowship. DJB acknowledges support through NSF grant AST-1814053. KD acknowledges support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2094 – 390783311. This research was supported in part through computational resources and services provided by Advanced Research Computing (ARC) at the University of Michigan, Ann Arbor, and

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

also the Max Planck Computing and Data Facility (MPCDF) of the Max Planck Society.

The TNG simulations were run with compute time granted by the Gauss Centre for Supercomputing (GCS) under Large-Scale Projects GCS-ILLU and GCS-DWAR on the GCS share of the supercomputer Hazel Hen at the High Performance Computing Center Stuttgart (HLRS).

The *Magneticum* simulations were performed at the Leibniz-Rechenzentrum with CPU time assigned to the Project 'pr86re'.

All analysis in this work was enabled greatly by the following software: PANDAS (McKinney 2011), NUMPY (van der Walt, Colbert & Varoquaux 2011), SCIPY (Virtanen et al. 2020), SCI-KIT LEARN (Pedregosa et al. 2011), and MATPLOTLIB (Hunter 2007).

REFERENCES

Adams F. C., Fatuzzo M., 1996, ApJ, 464, 256
Ade P. et al., 2019, J. Cosmol. Astropart. Phys., 2019, 056
Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
Artale M. C., Zehavi I., Contreras S., Norberg P., 2018, MNRAS, 480, 3978
Bahé Y. M. et al., 2017, MNRAS, 470, 4186
Bahé Y. M. et al., 2019, MNRAS, 485, 2287
Barnes D. J., Kay S. T., Henson M. A., McCarthy I. G., Schaye J., Jenkins A., 2017, MNRAS, 465, 213
Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, MNRAS, 488, 3143
Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587
Biffi V., Dolag K., Böhringer H., Lemson G., 2012, MNRAS, 420, 3545
Bocquet S. et al., 2019, ApJ, 878, 55
Bose S., Eisenstein D. J., Hernquist L., Pillepich A., Nelson D., Marinacci F., Springel V., Vogelsberger M., 2019, MNRAS, 490, 5693
Bradshaw C., Leauthaud A., Hearin A., Huang S., Behroozi P., 2020, MNRAS, 493, 337
Bregman J. N., Anderson M. E., Miller M. J., Hodges-Kluck E., Dai X., Li J.-T., Li Y., Qu Z., 2018, ApJ, 862, 3
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Chiu I. et al., 2016, MNRAS, 458, 379
Chiu I.-N., Umetsu K., Murata R., Medezinski E., Oguri M., 2019, preprint (arXiv:1909.02042)
Cohn J. D., Evrard A. E., White M., Croton D., Ellingson E., 2007, MNRAS, 382, 1738
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Costanzi M. et al., 2019, MNRAS, 488, 4779
Croton D. J. et al., 2006, MNRAS, 365, 11
Cui W. et al., 2014, MNRAS, 437, 816
de Haan T. et al., 2016, ApJ, 832, 95
De Lucia G., Blaizot J., 2007, MNRAS, 375, 2
De Lucia G., Springel V., White S. D. M., Croton D., Kauffmann G., 2006, MNRAS, 366, 499
Dolag K., Borgani S., Murante G., Springel V., 2009, MNRAS, 399, 497
Donnert J., Dolag K., Brunetti G., Cassano R., 2013, MNRAS, 429, 3564
Elahi P. J. et al., 2016, MNRAS, 458, 1096
Erickson B. M. S., Cunha C. E., Evrard A. E., 2011, Phys. Rev. D, 84, 103506
Evrard A. E., Summers F. J., Davis M., 1994, ApJ, 422, 11
Evrard A. E., Arnault P., Huterer D., Farahi A., 2014, MNRAS, 441, 3562
Farahi A., Evrard A. E., Rozo E., Rykoff E. S., Wechsler R. H., 2016, MNRAS, 460, 3900
Farahi A., Evrard A. E., McCarthy I., Barnes D. J., Kay S. T., 2018, MNRAS, 478, 2618
Farahi A. et al., 2019a, MNRAS, 490, 2299
Farahi A. et al., 2019b, Nat. Commun., 10, 2504
Farahi A., Ho M., Trac H., 2020, MNRAS, 493, 1361
Gaspari M., Ruszkowski M., Oh S. P., 2013, MNRAS, 432, 3401
Golden-Marx J. B., Miller C. J., 2018, ApJ, 860, 2
Golden-Marx J. B., Miller C. J., 2019, ApJ, 878, 14
Hahn O., Martizzi D., Wu H.-Y., Evrard A. E., Teyssier R., Wechsler R. H., 2017, MNRAS, 470, 166
Hearin A. P., Zentner A. R., Berlind A. A., Newman J. A., 2013, MNRAS, 433, 659
Hearin A. P., Zentner A. R., van den Bosch F. C., Campbell D., Tollerud E., 2016, MNRAS, 460, 2552
Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, MNRAS, 442, 2304
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Ivezić Ž. et al., 2019, ApJ, 873, 111
Kaiser N., 1986, MNRAS, 222, 323
Katz N., White S. D. M., 1993, ApJ, 412, 455
Kaviraj S. et al., 2017, MNRAS, 467, 4739
Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M.-S., 2015, MNRAS, 450, 1349
Koulouridis E. et al., 2018, A&A, 620, A4
Kravtsov A. V., Borgani S., 2012, ARA&A, 50, 353
Kravtsov A. V., Vikhlinin A. A., Meshcheryakov A. V., 2018, Astron. Lett., 44, 8
Languignon D., Le Petit F., Rodrigo C., Lemson G., Molinaro M., Wozniak H., 2017, Technical report, Simulation Data Access Layer Version 1.0
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Le Brun A. M. C., McCarthy I. G., Schaye J., Ponman T. J., 2014, MNRAS, 441, 1270
Mantz A., Allen S. W., Rapetti D., Ebeling H., 2010, MNRAS, 406, 1759
Mantz A. B. et al., 2016, MNRAS, 463, 3582
Mantz A. B., Allen S. W., Morris R. G., von der Linden A., 2018, MNRAS, 473, 3072
Marinacci F. et al., 2018, MNRAS, 480, 5113
Matthee J., Schaye J., Crain R. A., Schaller M., Bower R., Theuns T., 2017, MNRAS, 465, 2381
McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, MNRAS, 465, 2936
McClintock T. et al., 2019, MNRAS, 482, 1352
McKinney W., 2011, Python for High Performance and Scientific Computing, 14, 56
McNamara B. R., Nulsen P. E. J., 2012, New J. Phys., 14, 055023
Merloni A. et al., 2012, preprint (arXiv:1209.3114)
Moustakas J. et al., 2013, ApJ, 767, 50
Mulroy S. L. et al., 2014, MNRAS, 443, 3309
Mulroy S. L. et al., 2019, MNRAS, 484, 60
Naiman J. P. et al., 2018, MNRAS, 477, 1206
Nelson D. et al., 2018, MNRAS, 475, 624
Nelson D. et al., 2019, Comput. Astrophys. Cosmol., 6, 2
Obreschkow D., Elahi P. J., Lagos C. d. P., Poulton R. J. J., Ludlow A. D., 2020, MNRAS, 493, 4551
Pedregosa F. et al., 2011, J. Mach. Learn. Res., 12, 2825
Pillepich A., Porciani C., Reiprich T. H., 2012, MNRAS, 422, 44
Pillepich A. et al., 2018a, MNRAS, 473, 4077
Pillepich A. et al., 2018b, MNRAS, 475, 648
Pillepich A., Reiprich T. H., Porciani C., Borm K., Merloni A., 2018c, MNRAS, 481, 613
Predehl P. et al., 2014, eROSITA on SRG, SPIE Conference Series, p. 91441T
Racca G. D. et al., 2016, The Euclid Mission Design, SPIE Conference Series. p. 99040O
Raddick M. J., Kim J. W., Lemson G., Medvedev D., Taghizadeh-Popp M., 2019, SciServerCompute: Bring Analysis Close to the Data, Astronomical Society of the Pacific Conference Series. p. 749
Ragagnin A., Dolag K., Biffi V., Cadolle Bel M., Hammer N. J., Krukau A., Petkova M., Steinborn D., 2017, Astron. Comput., 20, 52
Ragone-Figueroa C., Granato G. L., Murante G., Borgani S., Cui W., 2013, MNRAS, 436, 1750
Rasia E. et al., 2015, ApJ, 813, L17
Rodriguez-Gomez V. et al., 2016, MNRAS, 458, 2371
Rozo E. et al., 2010, ApJ, 708, 645
Salvadori S., 2019, Uncertainty Quantification in CFD: The Matrix of Knowledge. Springer International Publishing, Cham, p. 33

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

Scannapieco C. et al., 2012, MNRAS, 423, 1726
Shaw L. D., Holder G. P., Dudley J., 2010, ApJ, 716, 281
Sohn J., Geller M. J., Rines K. J., Hwang H. S., Utsumi Y., Diaferio A., 2018, ApJ, 856, 172
Spergel D. et al., 2015, preprint (arXiv:1503.03757)
Springel V., 2005, MNRAS, 364, 1105
Springel V., 2010, MNRAS, 401, 791
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726
Springel V. et al., 2018, MNRAS, 475, 676
Stodden V. et al., 2016, Science, 354, 1240
Tang L., Lin W., Cui W., Kang X., Wang Y., Contini E., Yu Y., 2018, ApJ, 859, 85
The Dark Energy Survey Collaboration, 2005, preprint (astro-ph/0510346)
Tremaine S. D., Richstone D. O., 1977, ApJ, 212, 311
Tremmel M. et al., 2019, MNRAS, 483, 3336
van den Bosch F. C., Ogiya G., 2018, MNRAS, 475, 4066
van der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
Vikhlinin A. et al., 2009, ApJ, 692, 1060
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Vogelsberger M. et al., 2014, Nature, 509, 177
Vogelsberger M., Marinacci F., Torrey P., Puchwein E., 2020, Nat. Rev. Phys., 2, 42
Voit G. M., Donahue M., Bryan G. L., McDonald M., 2015, Nature, 519, 203
Wechsler R. H., Tinker J. L., 2018, ARA&A, 56, 435
Wu H.-Y., Evrard A. E., Hahn O., Martizzi D., Teyssier R., Wechsler R. H., 2015, MNRAS, 452, 1982
York D. G. et al., 2000, AJ, 120, 1579
Zehavi I., Contreras S., Padilla N., Smith N. J., Baugh C. M., Norberg P., 2018, ApJ, 853, 84
Zentner A. R., Hearin A. P., van den Bosch F. C., 2014, MNRAS, 443, 3044
Zhang Y. Y., Andernach H., Caretta C. A., Reiprich T. H., Böhringer H., Puchwein E., Sijacki D., Girardi M., 2011, A&A, 526, A105
Zhang Y. et al., 2016, ApJ, 816, 98
Zhang Y. et al., 2019, ApJ, 874, 165
ZuHone J. A., Kowalik K., Öhman E., Lau E., Nagai D., 2018, ApJS, 234, 4

APPENDIX A: LLR FITS TO BCG AND TOTAL STELLAR MASS

Fig. A1 shows the scaling of total stellar mass within R_{200c} with halo mass at $z = 0$, including a panel for comparisons of the normalizations. Note that we do not account for differences in the cosmic baryon fraction, $f_b = \Omega_b/\Omega_m$, of the simulations here. Accounting for these differences brings the normalizations into slightly better agreement.

The scaling relations for $M_{\star,\mathrm{BCG}}$–M_{halo}, shown in Fig. A2, reveal that the central galaxies in MGTM are a factor ∼2–3 more massive than those in TNG300 and BM.

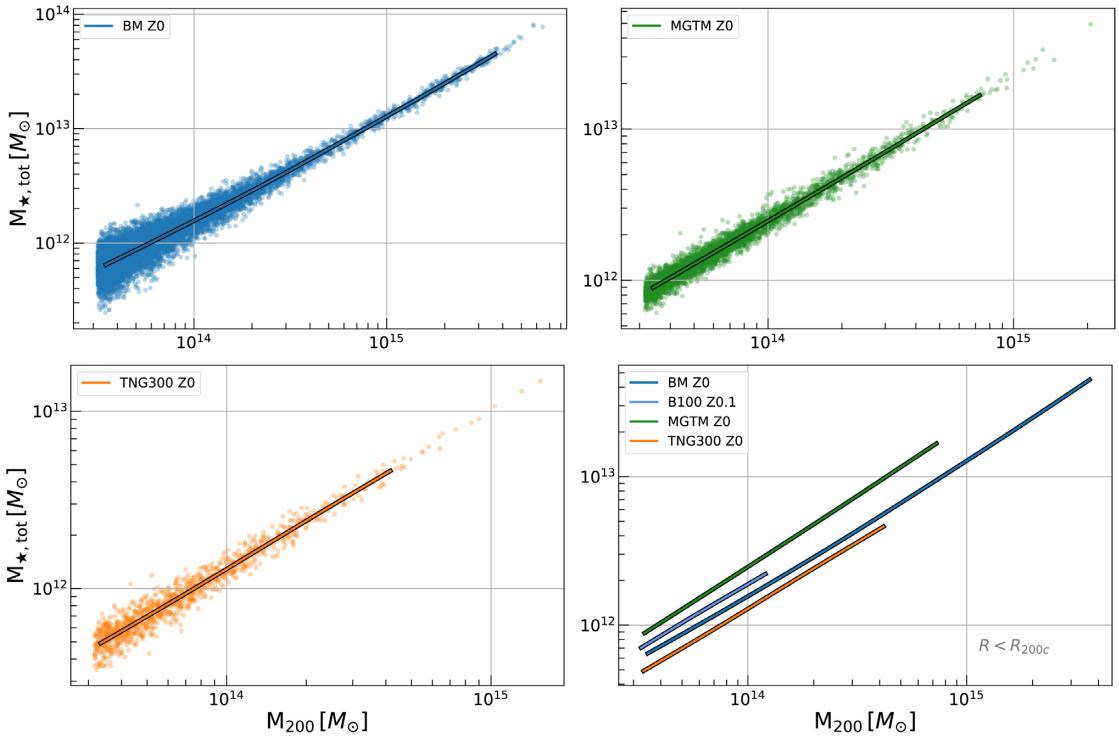

Figure A1. Scaling of total stellar mass within R_{200c} at $z = 0$ in the three simulations. Format is identical to Fig. 2.

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

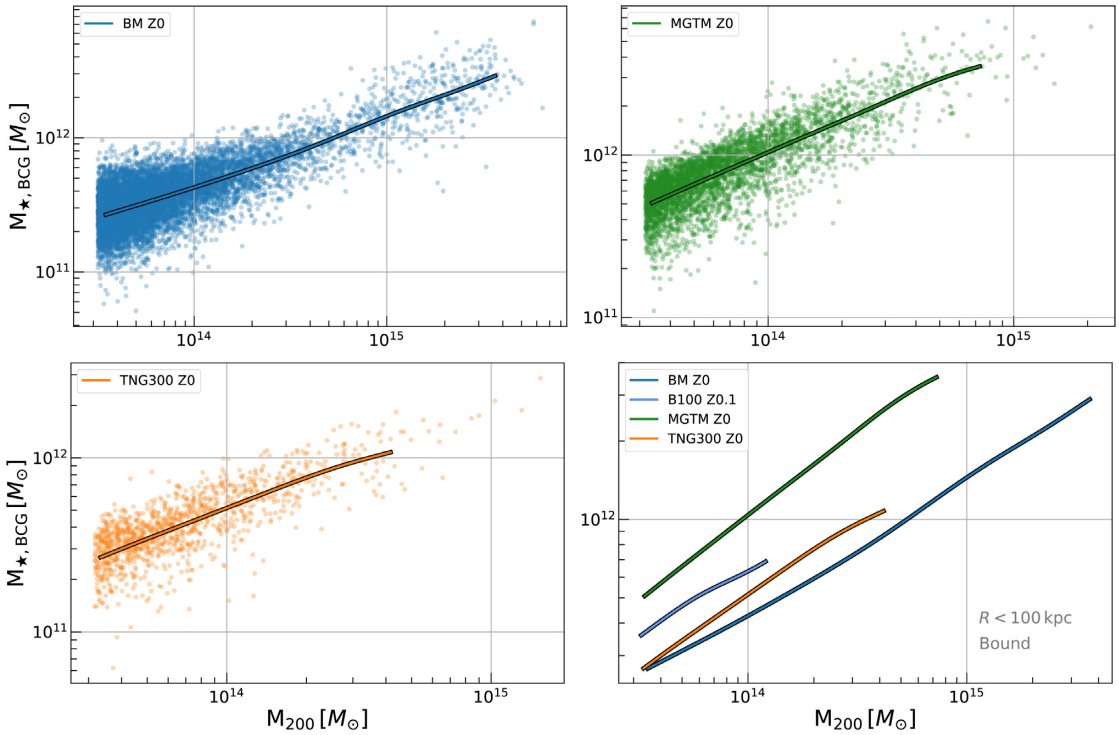

Figure A2. Scaling of central galaxy stellar mass within R_{200c} at $z = 0$ in the three simulations. Format is identical to Fig. 2.

APPENDIX B: GMM, SKEWNESS, AND LLR FIT PARAMETERS

We provide individual simulation values of the N_{sat} skewness and GMM fit parameters in Table B1. We also provide the LLR fit parameters for massive halos at z = 0 for BM (Table B2), B100 (Table B3), MGTM (Table B4), and TNG300 (Table B5).

The remaining tables provide $z = 0$ LLR fit parameters as a function of total halo mass, sampled in 0.1 dex intervals in the log of halo mass.

Table B1. Parameters of the N_{sat} Gaussian mixture model (Fig. 10), and the kernel skewness (Fig. 9) for all simulations having samples of >300 haloes with $M_{halo} > 10^{13.5}$ $M_\odot$ at the available redshifts.

z	μ_1	μ_2	σ_1	σ_2	f_1	$\gamma(N_{sat})$	$\gamma(M_{\star,tot})$	$\gamma(M_{\star,BCG})$
				BM				
0[a]	0.25 ± 0.03	-0.91 ± 0.12	0.67 ± 0.02	1.05 ± 0.04	0.79 ± 0.03	-0.85 ± 0.05	-0.26 ± 0.02	-0.31 ± 0.03
0.5	0.31 ± 0.02	-0.90 ± 0.08	0.64 ± 0.01	1.05 ± 0.03	0.74 ± 0.03	-0.93 ± 0.04	-0.19 ± 0.03	-0.39 ± 0.03
1	0.29 ± 0.02	-0.99 ± 0.1	0.63 ± 0.02	1.06 ± 0.04	0.77 ± 0.027	-0.99 ± 0.05	-0.05 ± 0.04	-0.34 ± 0.04
1.5	0.28 ± 0.03	-1.01 ± 0.18	0.67 ± 0.02	1.06 ± 0.06	0.78 ± 0.04	-0.87 ± 0.06	-0.20 ± 0.10	-0.39 ± 0.09
2	0.24 ± 0.06	-1.03 ± 0.35	0.70 ± 0.39	1.06 ± 0.11	0.81 ± 0.07	-0.73 ± 0.09	-0.23 ± 0.21	-0.56 ± 0.14
				MGTM				
0	0.25 ± 0.02	-1.08 ± 0.14	0.71 ± 0.02	1.19 ± 0.05	0.82 ± 0.03	-0.96 ± 0.05	-0.020 ± 0.05	-0.80 ± 0.05
0.5	0.23 ± 0.02	-1.02 ± 0.15	0.65 ± 0.02	1.05 ± 0.05	0.83 ± 0.03	-0.89 ± 0.07	0.31 ± 0.06	N/A[b]
1	0.23 ± 0.03	-0.99 ± 0.18	0.65 ± 0.02	1.08 ± 0.07	0.82 ± 0.04	-0.98 ± 0.10	0.33 ± 0.06	N/A[b]
				TNG300				
0	0.25 ± 0.05	-1.09 ± 0.25	0.67 ± 0.03	1.04 ± 0.08	0.81 ± 0.05	-0.90 ± 0.07	0.034 ± 0.06	-0.73 ± 0.09
0.5	0.23 ± 0.05	-1.05 ± 0.3	0.66 ± 0.03	1.04 ± 0.09	0.81 ± 0.06	-0.92 ± 0.08	0.016 ± 0.13	-0.57 ± 0.08
1	0.28 ± 0.11	-0.72 ± 0.35	0.67 ± 0.07	0.98 ± 0.11	0.74 ± 0.14	-0.70 ± 0.18	-0.06 ± 0.14	-0.59 ± 0.15

[a]For GMM parameters and N_{sat} kernel skewness, the sample uses $M_{halo} > 10^{13.8}$ $M_\odot$. For the other skewness computations, it is $M_{halo} > 10^{13.5}$ $M_\odot$

[b]We do not have $M_{\star,BCG}$ values for MGTM at $z > 0$.

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

Table B2. LLR Fits for BM at $z = 0$ for $N_{\rm sat}$, $M_{\star,\rm tot}$, and $M_{\star,\rm BCG}$. We show the decimal normalization ($\pi_{10} = \log_{10} e^{\pi}$), slope ($\alpha$), and scatter ($\sigma$, in ln terms) for each property, along with the correlation coefficients of property pairs (r).

	BM, $z = 0$											
	$N_{\rm sat}$			$M_{\star,\rm tot}$			$M_{\star,\rm BCG}$			r		
$\log_{10} M_{\rm Halo}$	π_{10}	α	σ	π_{10}	α	σ	π_{10}	α	σ	$N_{\rm Sat}$–$M_{\star,\rm tot}$	$N_{\rm Sat}$–$M_{\star,\rm BCG}$	$M_{\star,\rm tot}$–$M_{\star,\rm BCG}$
13.5	0.509	0.888	0.552	11.774	0.826	0.292	11.409	0.436	0.400	0.410	−0.120	0.630
13.6	0.600	0.911	0.543	11.857	0.828	0.279	11.452	0.434	0.390	0.439	−0.129	0.582
13.7	0.693	0.931	0.526	11.940	0.831	0.264	11.495	0.433	0.379	0.468	−0.139	0.530
13.8	0.788	0.947	0.502	12.023	0.836	0.247	11.539	0.436	0.366	0.493	−0.152	0.476
13.9	0.884	0.955	0.470	12.108	0.843	0.228	11.584	0.443	0.353	0.513	−0.168	0.421
14.0	0.980	0.957	0.435	12.194	0.852	0.210	11.629	0.451	0.340	0.529	−0.189	0.368
14.1	1.075	0.953	0.398	12.280	0.860	0.193	11.676	0.461	0.330	0.540	−0.214	0.319
14.2	1.169	0.946	0.364	12.368	0.869	0.177	11.723	0.469	0.322	0.548	−0.239	0.277
14.3	1.262	0.937	0.333	12.456	0.878	0.164	11.772	0.478	0.316	0.554	−0.266	0.240
14.4	1.354	0.928	0.307	12.546	0.890	0.151	11.821	0.489	0.311	0.559	−0.295	0.203
14.5	1.446	0.917	0.284	12.637	0.903	0.138	11.873	0.505	0.308	0.560	−0.331	0.163
14.6	1.536	0.908	0.263	12.730	0.916	0.124	11.927	0.526	0.308	0.555	−0.371	0.123
14.7	1.626	0.903	0.241	12.823	0.927	0.110	11.984	0.553	0.314	0.544	−0.406	0.096
14.8	1.716	0.902	0.218	12.917	0.936	0.098	12.043	0.579	0.327	0.522	−0.429	0.091
14.9	1.806	0.904	0.196	13.012	0.941	0.087	12.103	0.590	0.345	0.493	−0.438	0.107
15.0	1.897	0.909	0.179	13.106	0.946	0.079	12.161	0.576	0.366	0.469	−0.437	0.127
15.1	1.989	0.919	0.167	13.201	0.952	0.074	12.216	0.544	0.385	0.456	−0.427	0.138
15.2	2.082	0.936	0.158	13.297	0.960	0.072	12.268	0.516	0.398	0.449	−0.412	0.141
15.3	2.178	0.954	0.151	13.394	0.969	0.069	12.318	0.509	0.404	0.439	−0.395	0.141
15.4	2.275	0.967	0.145	13.492	0.977	0.067	12.371	0.522	0.404	0.424	−0.379	0.142
15.5	2.373	0.979	0.140	13.591	0.983	0.065	12.426	0.542	0.400	0.406	−0.365	0.148

Table B3. LLR Fits for B100 at $z = 0.12$ for $N_{\rm sat}$, $M_{\star,\rm tot}$, and $M_{\star,\rm BCG}$. We show the decimal normalization ($\pi_{10} = \log_{10} e^{\pi}$), slope ($\alpha$), and scatter ($\sigma$, in ln terms) for each property, along with the correlation coefficients of property pairs (r).

	B100, $z = 0.12$											
	$N_{\rm sat}$			$M_{\star,\rm tot}$			$M_{\star,\rm BCG}$			r		
$\log_{10} M_{\rm Halo}$	π_{10}	α	σ	π_{10}	α	σ	π_{10}	α	σ	$N_{\rm Sat}$–$M_{\star,\rm tot}$	$N_{\rm Sat}$–$M_{\star,\rm BCG}$	$M_{\star,\rm tot}$–$M_{\star,\rm BCG}$
13.5	0.646	1.105	0.516	11.836	0.899	0.180	11.545	0.628	0.314	0.269	−0.367	0.507
13.6	0.758	1.128	0.483	11.926	0.903	0.178	11.608	0.625	0.308	0.281	−0.349	0.499
13.7	0.871	1.127	0.439	12.016	0.897	0.172	11.667	0.592	0.301	0.300	−0.325	0.485
13.8	0.980	1.100	0.389	12.104	0.883	0.163	11.718	0.526	0.292	0.334	−0.298	0.455
13.9	1.084	1.059	0.341	12.190	0.868	0.151	11.759	0.455	0.285	0.385	−0.279	0.400
14.0	1.184	1.018	0.302	12.275	0.861	0.139	11.800	0.428	0.282	0.452	−0.281	0.320

Table B4. LLR Fits for MGTM at $z = 0.03$ for $N_{\rm sat}$, $M_{\star,\rm tot}$, and $M_{\star,\rm BCG}$. We show the decimal normalization ($\pi_{10} = \log_{10} e^{\pi}$), slope ($\alpha$), and scatter ($\sigma$, in ln terms) for each property, along with the correlation coefficients of property pairs (r).

	MGTM, $z = 0.03$											
	$N_{\rm sat}$			$M_{\star,\rm tot}$			$M_{\star,\rm BCG}$			r		
$\log_{10} M_{\rm Halo}$	π_{10}	α	σ	π_{10}	α	σ	π_{10}	α	σ	$N_{\rm Sat}$–$M_{\star,\rm tot}$	$N_{\rm Sat}$–$M_{\star,\rm BCG}$	$M_{\star,\rm tot}$–$M_{\star,\rm BCG}$
13.5	0.652	1.038	0.445	11.920	0.943	0.102	11.687	0.668	0.294	−0.068	−0.479	0.618
13.6	0.756	1.018	0.429	12.014	0.941	0.102	11.754	0.664	0.297	−0.066	−0.473	0.618
13.7	0.858	1.000	0.410	12.109	0.941	0.102	11.820	0.660	0.300	−0.065	−0.467	0.619
13.8	0.957	0.987	0.387	12.203	0.941	0.101	11.886	0.657	0.305	−0.067	−0.461	0.623
13.9	1.054	0.979	0.360	12.297	0.943	0.100	11.951	0.654	0.311	−0.075	−0.458	0.630
14.0	1.152	0.977	0.332	12.392	0.946	0.099	12.016	0.652	0.318	−0.090	−0.459	0.640
14.1	1.250	0.976	0.304	12.487	0.948	0.099	12.081	0.650	0.323	−0.109	−0.465	0.654
14.2	1.347	0.970	0.277	12.582	0.951	0.098	12.146	0.647	0.327	−0.130	−0.471	0.669
14.3	1.441	0.956	0.253	12.678	0.953	0.098	12.210	0.645	0.329	−0.148	−0.473	0.684
14.4	1.533	0.935	0.230	12.774	0.956	0.098	12.275	0.649	0.331	−0.161	−0.471	0.698
14.5	1.623	0.919	0.209	12.870	0.961	0.097	12.342	0.657	0.334	−0.167	−0.464	0.707
14.6	1.715	0.919	0.189	12.968	0.967	0.096	12.408	0.654	0.339	−0.169	−0.459	0.710
14.7	1.811	0.939	0.170	13.065	0.971	0.094	12.467	0.625	0.342	−0.170	−0.462	0.707
14.8	1.912	0.970	0.151	13.162	0.971	0.091	12.518	0.572	0.343	−0.175	−0.473	0.700

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020

Table B5. LLR fits for TNG300 at $z = 0$ for $N_{\rm sat}$, $M_{\star,\rm tot}$, and $M_{\star,\rm BCG}$. We show the decimal normalization ($\pi_{10} = \log_{10} e^{\pi}$), slope ($\alpha$), and scatter ($\sigma$, in ln terms) for each property, along with the correlation coefficients of property pairs (r).

	TNG300, $z = 0$											
	$N_{\rm sat}$			$M_{\star,\rm tot}$			$M_{\star,\rm BCG}$			r		
$\log_{10} M_{\rm Halo}$	π_{10}	α	σ	π_{10}	α	σ	π_{10}	α	σ	$N_{\rm Sat}$–$M_{\star,\rm tot}$	$N_{\rm Sat}$–$M_{\star,\rm BCG}$	$M_{\star,\rm tot}$–$M_{\star,\rm BCG}$
13.5	0.520	1.025	0.493	11.670	0.863	0.163	11.414	0.622	0.300	0.222	− 0.348	0.526
13.6	0.623	1.025	0.477	11.757	0.865	0.155	11.476	0.609	0.297	0.234	− 0.355	0.492
13.7	0.725	1.023	0.455	11.844	0.869	0.147	11.536	0.596	0.294	0.245	− 0.364	0.449
13.8	0.827	1.021	0.429	11.931	0.874	0.138	11.594	0.588	0.293	0.255	− 0.374	0.404
13.9	0.929	1.020	0.399	12.020	0.883	0.130	11.653	0.585	0.294	0.262	− 0.379	0.367
14.0	1.031	1.018	0.367	12.110	0.892	0.122	11.711	0.585	0.295	0.269	− 0.381	0.339
14.1	1.133	1.016	0.336	12.200	0.899	0.114	11.770	0.586	0.294	0.278	− 0.378	0.316
14.2	1.234	1.015	0.305	12.291	0.902	0.108	11.829	0.585	0.290	0.291	− 0.373	0.296
14.3	1.335	1.014	0.276	12.381	0.899	0.102	11.885	0.575	0.285	0.305	− 0.369	0.283
14.4	1.437	1.016	0.248	12.470	0.894	0.096	11.938	0.546	0.283	0.314	− 0.370	0.278
14.5	1.539	1.019	0.221	12.558	0.889	0.088	11.983	0.499	0.285	0.317	− 0.378	0.273
14.6	1.642	1.022	0.199	12.646	0.886	0.081	12.024	0.451	0.290	0.314	− 0.392	0.257

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

Downloaded from https://academic.oup.com/mnras/article-abstract/495/1/686/5826180 by Liverpool John Moores University user on 02 June 2020