# The FABLE simulations: a feedback model for galaxies, groups, and clusters

Nicholas A. Henden,[1][*] Ewald Puchwein,[1,2] Sijing Shen[3] and Debora Sijacki[1,2]

[1] *Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[2] *Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[3] *Institute of Theoretical Astrophysics, University of Oslo, PO Box 1029 Blindern, NO-0315 Oslo, Norway*

## ABSTRACT

We present the Feedback Acting on Baryons in Large-scale Environments suite of cosmological hydrodynamical simulations of galaxies, groups, and clusters. The simulations use the AREPO moving-mesh code with a set of physical models for galaxy formation based on the successful Illustris simulation, but with updated active galactic nucleus (AGN) and supernovae feedback models. This allows us to simultaneously reproduce the observed redshift evolution of the galaxy stellar mass function together with the stellar and gas mass fractions of local groups and clusters across a wide range of halo masses. Focusing on the properties of groups and clusters, we find very good agreement with a range of observed scaling relations, including the X-ray luminosity–total mass and gas mass relations as well as the total mass–temperature and Sunyaev–Zel'dovich flux–mass relations. Careful comparison of our results with scaling relations based on X-ray hydrostatic masses as opposed to weak-lensing-derived masses reveals some discrepancies, which hint towards a non-negligible X-ray mass bias in observed samples. We further show that radial profiles of density, pressure, and temperature of the simulated intracluster medium are in very good agreement with observations, in particular for $r > 0.3 \, r_{500}$. In the innermost regions however we find too large entropy cores, which indicates that a more sophisticated modelling of the physics of AGN feedback may be required to accurately reproduce the observed populations of cool-core and non-cool-core clusters.

**Key words:** methods: numerical – galaxies: clusters: general – galaxies: clusters: intracluster medium – galaxies: groups: general – X-rays: galaxies: clusters.

## 1 INTRODUCTION

Galaxy clusters collapse from the rarest peaks in the matter density field and therefore provide large leverage to probe cosmic structure growth and cosmology. Ongoing and forthcoming surveys are set to greatly extend the number of known galaxy clusters and groups. In X-rays, the upcoming eROSITA telescope (Merloni et al. 2012; Pillepich, Porciani & Reiprich 2012) is expected to detect as many as one hundred thousand groups and clusters of galaxies out to $z \sim 1$, whilst observations of the Sunyaev–Zel'dovich (SZ) effect with SPT-3G (Benson et al. 2014) and ACTpol (Henderson et al. 2016) extend the search to higher redshift. Detailed follow-up observations will be possible at X-ray wavelengths with missions such as *XARM* (replacement for *Hitomi*; Takahashi et al. 2010, 2016) and *Athena* (Nandra et al. 2013) and at optical and near-infrared wavelengths with ground-based surveys such as the Dark Energy Survey (Dark

Energy Survey Collaboration 2005) and the Large Synoptic Survey Telescope (LSST Dark Energy Science Collaboration 2012), as well as future space-based missions such as Euclid (Laureijs et al. 2011).

This wealth of data holds great potential to provide stringent constraints on cluster physics and cosmological parameters, such as the amplitude and slope of the matter power spectrum and the densities of baryons, dark matter, and dark energy (see Allen, Evrard & Mantz 2011; Kravtsov & Borgani 2012; Planelles, Schleicher & Bykov 2015 for recent reviews). However, the cosmological interpretation of observed cluster data is crucially dependent on the degree to which we can connect cluster observables to theoretical predictions for a given cosmological model. Simulations of cosmic structure formation play an important role in this process. At the most basic level, they provide predictions for the abundance of clusters as a function of their mass for a particular cosmology (e.g. Cohn & White 2008; Tinker et al. 2008; Courtin et al. 2010; Crocce et al. 2010; Bhattacharya et al. 2011; Murray, Power & Robotham 2013; Watson et al. 2013). Yet as larger cluster surveys reduce the statistical uncertainty, we are increasingly limited by our incom-

[*] E-mail: n.henden@ast.cam.ac.uk

plete understanding of cluster physics and its impact on the relation between cluster observables (such as X-ray luminosity, temperature or SZ flux) and mass (e.g. Rudd, Zentner & Kravtsov 2008; Semboloni et al. 2011; van Daalen et al. 2011; Cui, Borgani & Murante 2014; Cusworth et al. 2014; Velliscig et al. 2014; Henson et al. 2017; Chisari et al. 2018).

Cosmological hydrodynamical simulations can aid in understanding and constraining these relations, as well as explaining potential systematic biases in observations. As an example, the most common method for measuring cluster masses is by analysing X-ray data under the assumption that the X-ray emitting gas is in hydrostatic equilibrium with the gravitational potential of the cluster. This method has been reported to systematically underestimate the true halo mass (e.g. Miralda-Escude & Babul 1995; Wu & Fang 1997; Mahdavi et al. 2008; Richard et al. 2010; Mahdavi et al. 2013; Foëx, Böhringer & Chon 2017); however, there is still no consensus on the exact magnitude of the effect or indeed whether such a bias exists at all (e.g. Zhang et al. 2010; Gruen et al. 2014; Applegate et al. 2016; Planck Collaboration XXIV 2016; Rines et al. 2016). Furthermore, as cluster surveys expand in size and depth, it will be vital to understand potential selection biases associated with a given observable. For example, X-ray surveys may be biased towards cool-core clusters with strong central X-ray emission, the prevalence and evolution of which remains uncertain (e.g. Andrade-Santos et al. 2017; Rossetti et al. 2017). Cosmological hydrodynamical simulations are in a unique position to quantify such potential biases and therefore facilitate the use of clusters as precise cosmological probes.

The hydrodynamical modelling of cluster formation has progressed considerably in recent years, largely due to increases in computing power and improvements in the treatment of physical processes that act below the resolution scale of the simulation, commonly referred to as 'sub-grid' models. In particular, the feedback from active galactic nuclei (AGNs) has proven critical in explaining a range of cluster observables (e.g. Sijacki et al. 2007; Puchwein, Sijacki & Springel 2008; Fabjan et al. 2010; McCarthy et al. 2010, 2011; Puchwein et al. 2010; Gaspari et al. 2014). These improvements have facilitated hydrodynamical simulations aimed at reproducing various observational properties of clusters, such as the cool-core/non-cool-core dichotomy (e.g. Rasia et al. 2015; Hahn et al. 2017) and the X-ray and SZ scaling relations (e.g. Pike et al. 2014; Planelles et al. 2014), including their redshift evolution (e.g. Fabjan et al. 2011; Le Brun et al. 2017; Truong et al. 2018). A small number of groups have also produced hydrodynamical simulations of statistical samples of massive haloes useful for cluster cosmology (Le Brun et al. 2014; Dolag, Komatsu & Sunyaev 2016; Barnes et al. 2017b; McCarthy et al. 2017). One of the limitations of these works however is that they lack the numerical resolution needed to resolve the detailed structure of the cluster galaxies. Given that galaxies are often used as observational tracers of large-scale structure, it is important to work towards reproducing the properties of the galaxy populations of clusters in addition to their global stellar, gas, and halo properties.

As progress has been made in the modelling of cluster formation, so too has there been a marked increase in the realism of simulated field galaxy populations. A number of groups have produced high-resolution cosmological hydrodynamical simulations of galaxy formation capable of reproducing a range of key observations of galaxies, such as their sizes, morphologies, passive fractions, and the build-up of their stellar mass. Notable examples include Illustris (Vogelsberger et al. 2014), EAGLE (Schaye et al. 2015), Horizon-AGN (Dubois et al. 2014), and MassiveBlack-II

(Khandai et al. 2015). The computational expense associated with the high-resolution and complex sub-grid models of these simulations limits their total volume. As a result, few, if any, galaxy clusters exist within the simulation volume and it is difficult to assess the ability of the galaxy formation model to reproduce observations of massive collapsed structures.

Recently, the C-EAGLE (Bahé et al. 2017; Barnes et al. 2017c) and IllustrisTNG (Pillepich et al. 2018b) projects have combined these two approaches and produced high-resolution simulations that resolve the full dynamic range from galaxies to massive clusters. The C-EAGLE project consists of 'zoom-in' simulations of 30 galaxy clusters in the mass range of $M_{200} = 10^{14}$–$10^{15.4}\,\mathrm{M_\odot}$[1] simulated with the same galaxy formation model as the EAGLE simulation and at the same resolution. The C-EAGLE clusters are a good match to a number of cluster observables, including the satellite stellar mass function, the total stellar and metal content and the scaling of X-ray spectroscopic temperature and SZ flux with total mass. Conversely, the clusters are too gas rich and possess too high central temperatures. Their results imply that improved agreement with observations will require revision of their model for AGN feedback (see discussion in Barnes et al. 2017c). The IllustrisTNG project consists of three large, uniformly sampled cosmological volumes of approximately 50, 100, and 300 Mpc on a side simulated with magnetohydrodynamics (Pillepich et al. 2018a). The largest volume contains numerous low-mass clusters, with a few objects reaching masses $M_{200} \sim 10^{15}\,\mathrm{M_\odot}$ at $z = 0$. IllustrisTNG is a follow-up project of the Illustris simulation (Vogelsberger et al. 2013, 2014; Genel et al. 2014) and improves upon Illustris by extending the explorable mass range of simulated haloes and revising the astrophysical modelling to address some of the shortcomings of Illustris. The Illustris simulation was successful in reproducing a broad range of observations of galaxy populations at various redshifts, including the observed range of galaxy morphologies and the evolution of galaxy specific star formation rates. However, various tensions with observations remained, especially on the scale of galaxy groups and (low-mass) clusters. In particular, the AGN feedback in Illustris ejected too much gas from low-redshift, massive haloes ($M_{500} \approx 10^{13}$–$10^{14}\,\mathrm{M_\odot}$; Genel et al. 2014). IllustrisTNG introduce a new model for AGN feedback which is able to suppress star formation in massive haloes without removing too much gas (Weinberger et al. 2017a). Combined with changes to other aspects of the modelling, such as modified galactic winds and the introduction of magnetic fields (Pillepich et al. 2018a), this has allowed IllustrisTNG to reproduce a range of observables properties of galaxies, groups, and clusters.

The Feedback Acting on Baryons in Large-scale Environments (FABLE) project presented here shares a similar motivation to IllustrisTNG. We have worked independently to build a suite of simulations based upon the framework of the successful Illustris project but improve upon the agreement with observations on scales larger than galaxies, e.g. with constraints on the gas content of massive haloes. By using zoom-in simulations we also extend the comparison with observations to massive clusters, which were not present in the original Illustris simulation. Like IllustrisTNG, we have updated the Illustris models for galactic winds and AGN feedback. We find that our changes improve the realism of groups and clusters

---

[1]Throughout this paper, spherical-overdensity masses and radii use the critical density of the Universe as a reference point. Hence, $M_{200}$ is the total mass inside a sphere of radius $r_{200}$ within which the average density is 200 times the critical density of the Universe.

significantly. The FABLE simulations consist of a uniformly sampled cosmological volume about 60 Mpc on a side and a series of zoom-in simulations of groups and clusters approximately uniformly spaced in logarithmic halo mass. In this paper, we focus on the $z = 0$ properties of FABLE groups and clusters in comparison to observations and demonstrate good agreement in a number of key areas.

The outline of this paper is as follows. Section 2.1 describes the characteristics of the FABLE simulations and Section 3 describes our methods for comparing the simulations with observations. We then present the galaxy stellar mass function (GSMF) at different redshifts (Section 4), the global properties of massive haloes (Section 5) and the hot gas profiles of groups and clusters (Section 6). We discuss our results in Section 7 and summarize our findings in Section 8.

## 2 SIMULATIONS

### 2.1 Basic simulation properties

We utilize the cosmological hydrodynamic moving-mesh code AREPO (Springel 2010), which solves the Euler equations of hydrodynamics on a quasi-Lagrangian moving Voronoi mesh. On top of gravity, hydrodynamics, and a spatially uniform ionizing background, the FABLE simulations employ a set of sub-grid models for processes important for galaxy formation. The majority of our sub-grid models are unchanged from Illustris and are described in full detail in Vogelsberger et al. (2013) and Torrey et al. (2014). These include models for star formation (Springel & Hernquist 2003; Springel, Di Matteo & Hernquist 2005), radiative cooling (Katz, Weinberg & Hernquist 1996; Wiersma, Schaye & Smith 2009a), and chemical enrichment (Wiersma et al. 2009b). The sub-grid models for feedback from stars (Vogelsberger et al. 2013) and AGN (Di Matteo, Springel & Hernquist 2005; Springel et al. 2005; Sijacki et al. 2007) have been modified from the Illustris models and are described in Sections 2.3 and 2.4. The parameters of the feedback models have been calibrated to reproduce observations of the local GSMF (Section 4) and the gas mass fractions of massive haloes (Section 5). This same calibration strategy was adopted for the BAHAMAS simulations (McCarthy et al. 2017), which successfully reproduce a broad range of observed hot gas and stellar properties of massive systems. The calibration process involved a series of periodic boxes of length $40\,h^{-1}$ Mpc simulated with different parametrizations of stellar and AGN feedback. We note that, as this volume is relatively small, these simulations did not contain massive clusters and therefore gas fractions were initially matched to observations on group-scales ($M_{500} \lesssim 10^{14}\,M_\odot$). Results from a sample of these calibration simulations are presented in Appendix A. The FABLE suite of simulations consists of the calibration volume and a series of zoom-in simulations of individual galaxy groups and clusters, which apply our calibrated model to higher mass objects.

The $40\,h^{-1}$ (comoving) Mpc periodic box was evolved to $z = 0$ from initial conditions based on a Planck cosmology (Planck Collaboration XIII 2016) with cosmological parameters $\Omega_\Lambda = 0.6935$, $\Omega_M = 0.3065$, $\Omega_b = 0.0483$, $\sigma_8 = 0.8154$, $n_s = 0.9681$, and $H_0 = 67.9\,\mathrm{km\,s^{-1}\,Mpc^{-1}} = h \times 100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. The simulation follows $512^3$ dark matter particles of mass $m_{DM} = 3.4 \times 10^7\,h^{-1}\,M_\odot$ and approximately $512^3$ baryonic resolution elements (gas cells and star/BH particles) of typical mass $\overline{m}_b = 6.4 \times 10^6\,h^{-1}\,M_\odot$. The gravitational softening length was fixed to $2.393\,h^{-1}$ kpc in physical coordinates below $z = 5$ and fixed in comoving coordinates at higher redshifts. This choice is consistent with the $z = 0$ opti-

mal softening length for low-mass cluster galaxies according to the empirical rule determined by Power et al. (2003).[2]

For our series of zoom-in simulations, individual groups and clusters were chosen from a collisionless (dark matter-only) parent simulation and re-simulated at high resolution. The zoom-in regions were drawn from Millennium XXL, a periodic box of side length $3\,h^{-1}$ Gpc (Angulo et al. 2012). Six systems were selected logarithmically spaced in mass spanning the mass range from groups ($\sim 10^{13}\,M_\odot$) to massive clusters ($\sim 10^{15}\,M_\odot$). These systems were selected only by their total mass within the parent simulation, with no prior knowledge regarding, for example, their dynamical state. The high-resolution regions were chosen such that they are free from lower resolution particles out to approximately $5\,r_{500}$ at $z = 0$. Dark matter particles in the high-resolution region have a mass of $m_{DM} = 5.5 \times 10^7\,h^{-1}\,M_\odot$. The gravitational softening length was fixed at $2.8125\,h^{-1}$ kpc in physical coordinates for $z \leq 5$ and fixed in comoving coordinates for $z > 5$. The softening lengths of boundary particles outside the high-resolution region were allowed to vary with their mass. Mode amplitudes in the initial conditions were scaled to a Planck cosmology (Planck Collaboration XIII 2016) with $\Omega_\Lambda = 0.6911$, $\Omega_m = 0.3089$, $\Omega_b = 0.0486$, $\sigma_8 = 0.8159$, $n_s = 0.9667$, and $H_0 = 67.74\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. In presenting our results, we rescale the appropriate quantities to the cosmology of the periodic box described above, although we note that this is a very small effect due to the similarity between the cosmological parameters.

The simulations were processed on-the-fly with the friends-of-friends (FoF) and SUBFIND algorithms (Davis et al. 1985; Springel et al. 2001; Dolag et al. 2009) to identify gravitationally bound groups of dark matter, stars, and gas ('haloes'), using a linking length of 0.2 times the mean dark matter interparticle separation, and decompose them into self-bound sub-structures ('sub-haloes').

### 2.2 Cluster visualization

In Fig. 1, we present a visualization of the gas properties of the most massive FABLE cluster at $z = 0$. The main panel shows the gas distribution in a $20 \times 20 \times 5$ Mpc slice centred on the gravitational potential minimum of the cluster. The gas surface density and temperature are encoded by the image brightness and hue/saturation, respectively, according to the colour map shown in the bottom left of the figure. The structure of the cluster gas is clearly laid out, revealing an array of interconnected filaments and infalling galaxies and groups, examples of which are shown in the panels on the right-hand side of the figure. During the formation of this cluster, a combination of adiabatic compression and shocks has heated the intracluster gas to high, X-ray emitting temperatures ($\gtrsim 10^7$ K). This is demonstrated in the middle right-hand panel, which shows the X-ray surface brightness within a region $2\,r_{500}$ on a side centred on the cluster.

### 2.3 Star formation, stellar feedback, and chemistry

Star formation and chemistry are treated with the same implementation as in Illustris, as described in detail in Vogelsberger et al. (2013) and Torrey et al. (2014). In brief, the dense star-forming interstellar medium (ISM) is treated in a sub-resolution fashion using a slightly modified version of the Springel & Hernquist (2003)

---

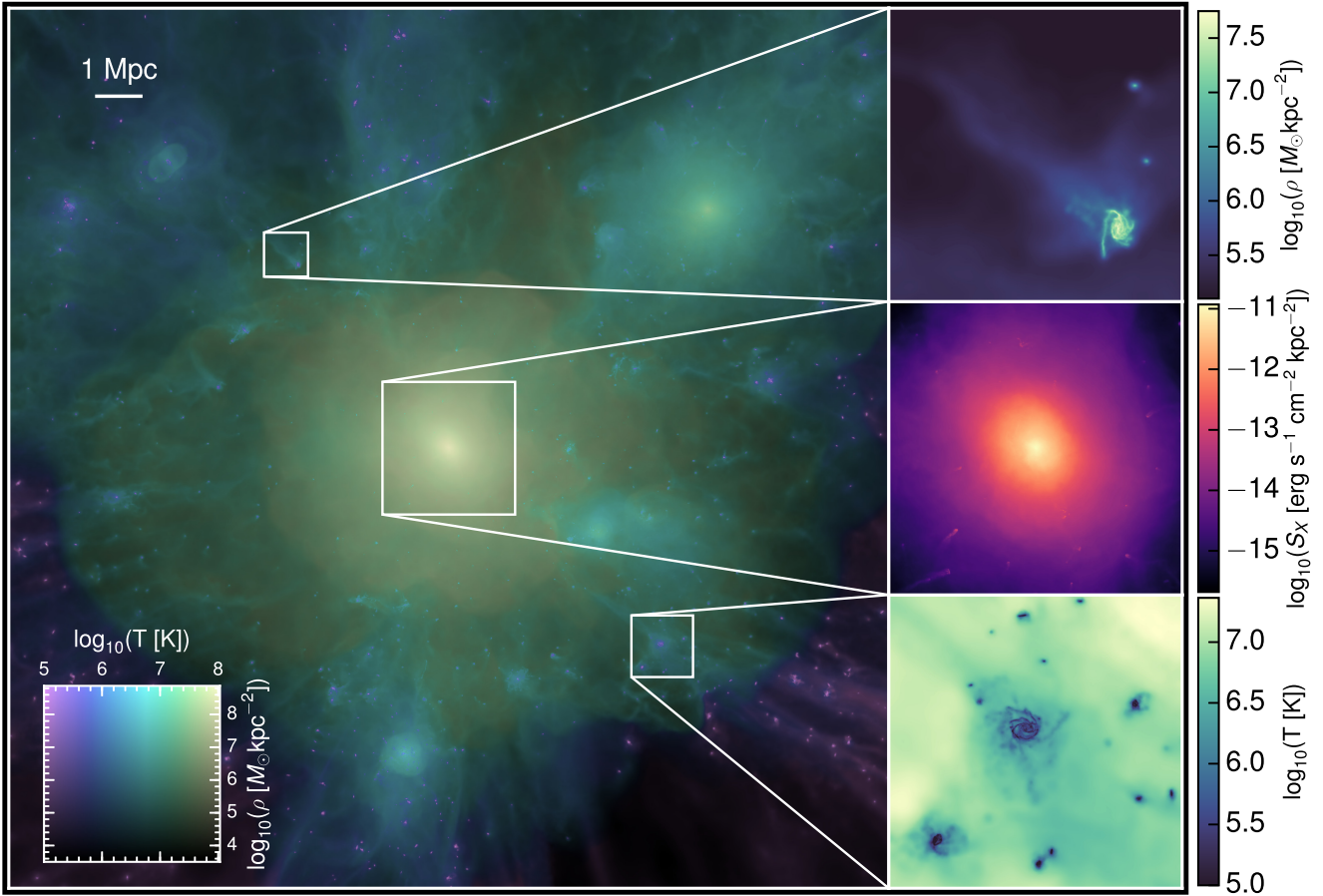[2]Note that this softening length is somewhat larger than used in Illustris.

**Figure 1.** A visualization of the gas properties of a massive FABLE cluster with $M_{500} = 9 \times 10^{14}\,\mathrm{M}_\odot$ at $z = 0$. The main panel shows a $20 \times 20 \times 5$ Mpc slice through the cluster with panels on the right-hand side zooming in on individual galaxies and the cluster core. In the main panel, the gas surface density is represented by the image brightness, whilst the gas temperature corresponds to the hue/saturation of the image according to the inset colour scale. The bottom right-hand panel shows the gas temperature distribution of a group of galaxies of various sizes and morphologies. The top right-hand panel shows the surface density of gas surrounding an individual spiral galaxy on the cluster outskirts, revealing the onset of ram pressure stripping of cold galactic gas as the galaxy approaches the dense cluster environment. Lastly, in the middle right-hand panel, we show the X-ray surface brightness in the 0.5–7 keV energy band within a region $2r_{500}$ on a side centred on the cluster.

sub-grid model. The ISM is modelled using an effective equation of state, where stars form stochastically from gas above a density threshold with a given star formation time-scale. Stellar mass-loss and metal enrichment are treated by calculating the mass and chemical composition of ejected material for each active star particle at each time-step and returning it to the nearby gas.

Stellar feedback in the form of galactic outflows is modelled via wind particles launched stochastically from star-forming gas with a velocity based on the local dark matter velocity dispersion (Vogelsberger et al. 2013). The wind particles are briefly decoupled from hydrodynamic interactions until they have left the local ISM and deposit their mass, metals, momentum, and thermal energy into the surrounding gas. In Illustris, the energy given to the wind particles is purely kinetic. These cold winds push gas out of the galactic disc, preventing overcooling within the dense ISM and lowering the galactic star formation rate. However, in this model, it is left fully up to hydrodynamical interactions to dissipate the kinetic energy to heat, which is often insufficient to prevent the ejected gas from quickly condensing back on to the galactic disc. We reduce this possibility by imparting one-third of the wind energy as thermal rather than kinetic. This slows the cooling of the ejected gas, causing it to remain outside of the dense ISM for a longer period of time,

increasing the overall effectiveness of stellar feedback. The same method has been used in Marinacci, Pakmor & Springel (2014) (50 per cent thermal energy), the Auriga galaxy simulations (50 per cent; Grand et al. 2017), and IllustrisTNG (10 per cent; Pillepich et al. 2018a).

### 2.4 Black hole accretion and feedback

#### 2.4.1 Seeding and accretion

Black hole (BH) formation is modelled assuming that 'seed' BHs of mass $10^5\,h^{-1}\,\mathrm{M}_\odot$ form regularly enough that every halo above a mass threshold of $5 \times 10^{10}\,h^{-1}\,\mathrm{M}_\odot$ contains a seed BH at its centre. Such haloes are identified by running a fast FoF algorithm on the fly. If a halo exceeds this mass threshold and does not already host a BH, the highest density gas cell in the halo is converted into a BH particle.

BHs are modelled as collisionless sink particles that are allowed to grow via mergers with other BHs and by accretion of ambient gas. The prescription for BH accretion is described in detail in Vogelsberger et al. (2013). Briefly, BHs accrete at an Eddington-limited Bondi–Hoyle–Lyttleton rate boosted by a constant factor $\alpha$

= 100 (Hoyle & Lyttleton 1939; Bondi & Hoyle 1944; Springer et al. 2005). A fraction $(1 - \epsilon_r)$ of the accreted mass is added to the mass of the BH, where $\epsilon_r = 0.1$ is the radiative efficiency. The remaining rest mass energy is made available as feedback energy.

### 2.4.2 AGN feedback

For feedback from BHs, we use an adapted form of the Illustris model. This model distinguishes between two states: a high accretion rate quasar mode (Di Matteo et al. 2005; Springer et al. 2005) and a low accretion rate radio mode (Sijacki et al. 2007). The quasar mode is dominant at high redshift, whilst the radio mode becomes increasingly important at lower redshifts as the average accretion rate diminishes.

In the quasar mode of feedback, a fraction $\epsilon_f = 0.1$ of the available feedback energy is coupled thermally and isotropically to the surrounding gas (see Springer et al. 2005 for details). In Illustris, BHs in the quasar-mode inject thermal energy into the surrounding gas *continuously*. However, this approach can result in artificial overcooling, whereby too little thermal energy is injected into too much mass and the energy is quickly radiated away. For this reason, Booth & Schaye (2009) developed a model for thermal AGN feedback in smoothed particle hydrodynamic (SPH) simulations in which BHs store feedback energy until it is sufficient to raise the temperature of a given number of SPH particles by a given amount. This model was subsequently employed in the cosmological hydrodynamical simulations EAGLE (Schaye et al. 2015), cosmo-OWLS (Le Brun et al. 2014), and BAHAMAS (McCarthy et al. 2017). We take a similar approach to reduce numerical overcooling in our simulations by introducing a duty cycle for the quasar-mode in which feedback energy is accumulated over a given time period, $\Delta t = 25$ Myr, before being released in a single event.

At low accretion rates below a fraction $\chi_{radio} = 0.01$ of the Eddington rate, we employ radio-mode feedback following Sijacki et al. (2007). In the radio-mode, hot buoyantly-rising bubbles are injected into the surrounding gas. The duty cycle of the radio-mode is controlled by the mass growth of the BH such that a bubble is injected once the BH has increased in mass by a fraction $\delta_{BH} \equiv \delta M_{BH}/M_{BH}$. The energy content of the bubble is then $E_{bub} = \epsilon_m \epsilon_r c^2 \delta M_{BH}$ for which we choose a radio-mode coupling efficiency $\epsilon_m = 0.8$. The total radio-mode feedback efficiency as given by the product $\epsilon_m \epsilon_r$ is 8 per cent, similar to Illustris (7 per cent). For the duty cycle, we set the threshold for bubble triggering to $\delta_{BH} = 0.01$, much smaller than $\delta_{BH} = 0.15$ as used in Illustris. Since the bubble energy is proportional to the increase in mass, this corresponds to more frequent but less energetic bubbles. For the bubble size and distance, we assume the scaling relations defined in Sijacki et al. (2007), with normalization constants $D_{bub, 0} = 30 \ h^{-1}$ kpc, $R_{bub, 0} = 50 \ h^{-1}$ kpc, $E_{bub, 0} = 10^{60}$ erg, and $\rho_{ICM,0} = 10^4 \ h^2 \ M_\odot$ kpc$^{-3}$ for the bubble distance, radius, energy content, and ambient density, respectively.

The remainder of the feedback energy that is not coupled to the surrounding gas by the quasar- or radio-mode goes into radiative electromagnetic feedback, which is approximated as an additional radiation field around the BH superposed with the redshift-dependent ultraviolet background. The full details of this model are given in Vogelsberger et al. (2013).

In Appendix A, we present several parametrizations of AGN feedback that were considered during the calibration process to demonstrate how changes to the duty cycle and energetics of the feedback modes impact the $z = 0$ GSMF and the stellar and gas mass fractions of massive haloes. In particular, we show that the introduction of a quasar-mode duty cycle can greatly suppress the growth of stellar mass in massive galaxies compared to continuous quasar-mode feedback.

## 3 COMPARING TO OBSERVATIONS

The physical properties of a simulation often do not correspond directly to those derived from real observations. This can result from a number of factors, including selection biases, projection effects, instrument systematics, and methodology. In this section, we will discuss how we can mitigate some of these effects in our comparisons with observations. We also discuss the derivation of quantities that are not intrinsic to the simulation. For example, X-ray luminosity is a complicated combination of continuum and line emission that is not followed in the simulations. We therefore follow a procedure described in Section 3.2 in which we create synthetic X-ray spectra in post-processing and analyse them similarly to observations.

### 3.1 Galaxy stellar mass functions

In the field of numerical galaxy formation, the GSMF is one of the most important observationally constrained properties that simulations should match. In constructing our GSMF, we define galaxies as the self-bound sub-haloes identified by SUBFIND and calculate their properties based on the corresponding sub-halo catalogue. With this definition, the total stellar mass of a galaxy is the total mass of star particles bound to the sub-halo. However, this does not necessarily correspond to the stellar mass of the galaxy as would be measured by an observer. This is because a significant fraction of the stellar mass in massive systems can exist in the form of diffuse intracluster light (ICL), which is difficult to quantify in stellar mass measurements due to its low surface brightness (e.g. Zibetti et al. 2005; Gonzalez, Zaritsky & Zabludoff 2007; Morishita et al. 2017). Some studies choose to integrate a galaxy's light within a 2D aperture of a given size (e.g. Li & White 2009), whilst others integrate Sersic or other fits to the light profiles (e.g. Baldry et al. 2012; Bernardi et al. 2013). This choice can have a significant impact on the derived stellar mass function (see e.g. Bernardi et al. 2013). We follow previous simulation studies in the literature and present our GSMF using multiple definitions of a galaxy's stellar mass: the total stellar mass bound to the corresponding sub-halo and two aperture masses. In one case, we follow Genel et al. (2014), who consider the stellar mass within a spherical aperture with radius equal to twice the stellar half-mass radius ($2r_{\star, 1/2}$). In the other, we follow Schaye et al. (2015) who define the galaxy stellar mass as the mass within a fixed spherical aperture of radius 30 physical kiloparsec (pkpc). In each case, only stellar mass bound to the sub-halo is considered.

### 3.2 X-ray properties

Galaxy clusters are permeated with diffuse gas heated to temperatures of the order of $10^{7-8}$ K. This intracluster medium (ICM) emits strongly in X-rays, providing an invaluable means of detecting and studying the properties of clusters. Previous studies have shown that the properties of hot gas derived from X-ray observations can be biased, for example due to complex thermal structure or clumping of the gas (e.g. Mazzotta et al. 2004; Rasia et al. 2005; Nagai, Vikhlinin & Kravtsov 2007; Khedekar et al. 2013). A reliable comparison between simulations and observations therefore requires actual simulation of the spectral properties of the X-ray emission.

We derive X-ray luminosities and spectroscopic temperatures in the following manner.

For a given halo, we consider gas within a projected aperture of radius $r_{500}$ centred on the minimum of the gravitational potential. This includes all gas along the line of sight (although for the zoom-in simulations we exclude gas outside the high-resolution region). This is intended to mimic X-ray observations, which measure projected luminosities and temperatures. We exclude cold gas with a temperature less than $3 \times 10^4$ K as the lack of molecular cooling in our simulations implies that the temperature of such gas can be significantly overestimated and it should contribute negligibly to the X-ray emission. We also exclude gas above the density threshold required for star formation, as the sub-grid multiphase model for star-forming gas does not reliably predict the thermal properties of the gas. In Appendix B1, we demonstrate that derived X-ray luminosities and spectroscopic temperatures are not particularly sensitive to the choice of temperature–density cut.

We project the emission measure of the gas on to temperature bins of width 0.02 keV between 0.01 and 24.0 keV, summing up the emission measure in each bin. Using the XSPEC package (Arnaud 1996, version 12.8.0), we then produce a mock X-ray spectrum by summing APEC emission models (Smith et al. 2001) generated for each temperature bin. For simplicity, we assume a constant metallicity of 0.3 times the solar value. This ensures that our X-ray analysis remains independent of the metal enrichment in the simulations, which is highly sensitive to the details of the feedback model. Even so, we find no significant difference in the derived X-ray luminosities and temperatures when using the metallicity of the gas in the simulations. The mock spectrum is convolved with the response function of *Chandra* and we adopt a large exposure time of $10^6$ s so as not to be limited by photon noise. We then fit the spectrum with a single-temperature APEC model in the 0.5–10 keV energy range with the temperature, metallicity, and normalization of the spectrum left as free parameters. From the best-fitting spectrum, we obtain the X-ray luminosity and spectroscopic temperature.

### 3.3 Halo masses

As previously discussed in Section 1, the most common method to obtain halo masses for calibrating cluster scaling relations is the X-ray hydrostatic method, which some studies have shown can underestimate the true halo mass. As there is no consensus as to the magnitude of the bias, any comparison between X-ray hydrostatic masses and true halo masses from the simulations must be carefully considered. We make use of the rising number of weak gravitational lensing studies to compare true masses from the simulations with both X-ray hydrostatic masses and halo masses measured with weak lensing. The latter are generally considered to be less biased due to the insensitivity of gravitational lensing to the equilibrium state of the gas or dark matter (for a review, see e.g. Hoekstra et al. 2013; Mandelbaum 2017). On the other hand, lensing mass measurements can possess significantly more scatter due to the effects of cluster sub-structure, cluster triaxiality, and projection effects (e.g. Corless & King 2007; Marian, Smith & Bernstein 2010; Meneghetti et al. 2010; Becker & Kravtsov 2011). Since there are also many more observed systems with available X-ray hydrostatic masses than weak-lensing masses, especially towards lower halo masses, we choose to compare spherical overdensity masses measured from the simulations to both measures where possible.

### 3.4 SZ properties

The thermal Sunyaev–Zel'dovich (tSZ) effect has long been recognized as a powerful tool for studying the physics of the ICM and the formation of large-scale structure (Birkinshaw 1999; Carlstrom, Holder & Reese 2002). The tSZ effect appears as a distortion in the cosmic microwave background (CMB) spectrum, arising from the inverse Compton scattering of CMB photons on energetic electrons in the ICM. When integrated over the volume of a system, the tSZ flux provides a relatively clean measure of its total thermal energy.

The tSZ signal is characterized by $Y_{500}$, the Comptonization parameter integrated over a sphere of radius $r_{500}$. Specifically,

$$D_A^2(z) Y_{500} \equiv \frac{\sigma_T}{m_e c^2} \int_0^{r_{500}} P \, \mathrm{d}V, \qquad (1)$$

where $D_A(z)$ is the angular diameter distance, $\sigma_T$ is the Thomson cross-section, $m_e$ is the electron rest mass, $c$ is the speed of light, and $P = n_e k T_e$ is the electron pressure, equal to the product of the electron number density and electron temperature.

We compare the tSZ signal of our simulated systems with Planck Collaboration XI (2013), who perform a stacking analysis on *Planck* multifrequency observations of a large sample of locally brightest galaxies (LBGs) selected from the Sloan Digital Sky Survey (SDSS). The selection criteria were designed to maximize the fraction of objects that are the central galaxies of their dark matter haloes. Correspondingly, we measure the tSZ signal only for central galaxies and use $r_{500}$ calculated in spheres centred on each galaxy. Planck Collaboration XI (2013) estimate the mean $Y_{500}$ in a series of stellar mass bins and convert this to a tSZ signal–halo mass relation by estimating an 'effective' halo mass for each stellar mass bin using a mock sample of LBGs from the semianalytic galaxy formation simulation of Guo et al. (2011). We also compare to the weak-lensing recalibration of the tSZ signal–halo mass relation given in Wang et al. (2016), which reduces the dependence of the halo mass estimates on the galaxy formation model and explicitly accounts for uncertainties in both the modelling and the lensing measurements.

Due to the limited angular resolution of *Planck*, the tSZ flux is actually measured within a (projected) aperture of radius $5\,r_{500}$. Planck Collaboration XI (2013) convert this measured flux, $Y_{5r_{500}}$, into the flux within a spherical aperture of radius $r_{500}$ by a conversion factor $Y_{500} = Y_{5r_{500}}/1.796$. This factor assumes the spatial template used in their matched filter, the universal pressure profile (Arnaud et al. 2010), and assumes no tSZ flux originates from beyond $5\,r_{500}$. In our comparisons, we remove the dependence on the assumed pressure distribution by converting the observationally inferred flux reported in Planck Collaboration XI (2013), $Y_{500}$, back into the actual measured flux, $Y_{5r_{500}}$. This is motivated by the work of Le Brun, McCarthy & Melin (2015) who show that $Y_{500}$ is highly sensitive to the assumed spatial template. Le Brun et al. (2015) generate synthetic tSZ maps from a cosmological hydrodynamical simulation from the cosmo-OWLS project that reproduces a range of global SZ, X-ray, and optical properties of local groups and clusters (Le Brun et al. 2014). Applying the same tools and assumptions as *Planck*, Le Brun et al. (2015) show that the inferred flux, $Y_{500}$, is biased high by a factor of ∼2 at $M_{500} = 2.6 \times 10^{13}\, \mathrm{M_\odot}$. This bias increases with decreasing halo mass, reaching nearly an order of magnitude overestimate below ∼$10^{13}\, \mathrm{M_\odot}$. The vast majority of the bias is due to the assumption of a fixed spatial template, which becomes an increasingly worse description of the hot gas in low-mass haloes.

The tSZ flux as measured by *Planck* for individual, low-mass systems may be biased due to source confusion (i.e. hot gas along the line-of-sight). Indeed, our tests have shown that the flux of individual low-mass systems can be significantly boosted by hot gas that overlaps with them in projection. However, Planck Collaboration XI (2013) estimate the mean tSZ flux from stacking a large number of systems in mass bins, which can remain unbiased by source confusion. Indeed, Le Brun et al. (2015) demonstrate that the mean tSZ flux is not significantly biased by uncorrelated confusion across the whole mass range of the Planck Collaboration XI (2013) stacking analysis. For this reason, we integrate the tSZ flux of a galaxy within a spherical aperture of radius $5\,r_{500}$ and note that this may underestimate the observed flux if correlated or uncorrelated source confusion leads to a significant bias in the observations.

### 3.5 ICM profiles

Radial profiles of the ICM characterize its distribution and thermodynamic history. The effects of non-gravitational processes such as AGN feedback cause ICM profiles to deviate from the self-similar relations predicted in the absence of such processes (e.g. Voit, Kay & Bryan 2005). A comparison of simulated profiles to observed ones is therefore a useful test of non-gravitational physics.

In X-ray observations, gas density profiles are derived from background-subtracted surface brightness profiles and gas temperature profiles are obtained from extracted spectra. Since these observations are sensitive only to hot X-ray emitting gas, in calculating the radial profiles of our simulated ICM we apply the same temperature–density cuts as used in our X-ray analysis described in Section 3.2. The gas is then divided into concentric spherical shells with logarithmically spaced radii centred on the minimum of the gravitational potential of the halo. For each radial bin, we calculate the volume-weighted electron number density, $n_e$, and the mass-weighted temperature, $T$. The former is defined as the total electron number divided by the volume of the bin. Since we may exclude some gas cells from the bin, we correct the bin volume by the ratio of the total volume of cells that are not excluded to the total volume of all cells in the bin.

## 4 THE GALAXY POPULATION

### 4.1 GSMF at $z = 0$

Fig. 2 shows the $z = 0$ GSMF for all galaxies in the $(40\,h^{-1}$ Mpc)$^3$ simulation volume. We plot the GSMF for three different definitions of a galaxy's stellar mass, as discussed in Section 3.1: these are the total bound stellar mass, the mass within twice the stellar half-mass radius $(2\,r_{\star,1/2})$ and the mass within a radius of 30 pkpc. We compare with the Illustris and EAGLE simulations (grey lines) and observational estimates of the GSMF from Li & White (2009), Baldry et al. (2012), Bernardi et al. (2013), and D'Souza, Vegetti & Kauffmann (2015) (symbols with error bars). We plot the $z = 0$ GSMF from Illustris for both the total bound mass and the mass within $2\,r_{\star,1/2}$ (Genel et al. 2014) and the $z = 0.1$ GSMF of EAGLE, which uses the mass within a 30 pkpc aperture (Schaye et al. 2015).

At the low-mass end of the GSMF, FABLE is in excellent agreement with the observational data. This contrasts with Illustris, which significantly overestimates the abundance of galaxies below the knee ($M_\star \lesssim 10^{10}\,\mathrm{M_\odot}$). Whilst some of this improvement is owed to changes to the stellar feedback model, we have determined that the dominant cause is a difference in the gravitational softening
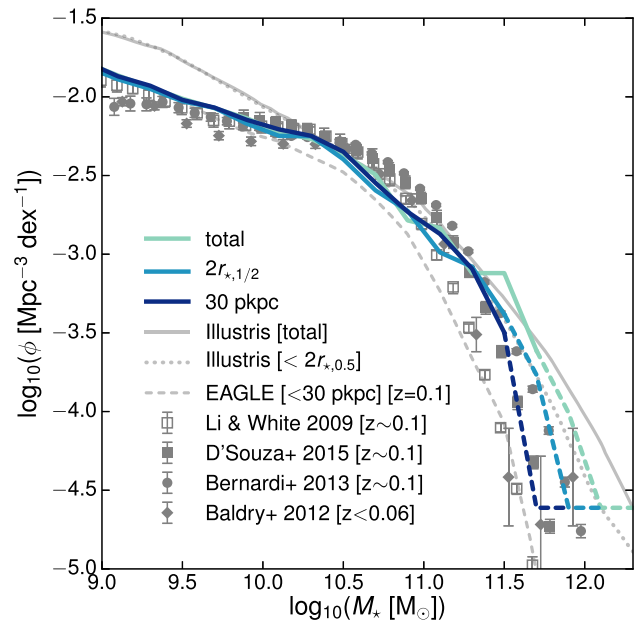


**Figure 2.** The FABLE GSMF at $z = 0$ for different definitions of galaxy stellar mass (lines) in comparison to observations (symbols with error bars). We consider three definitions of a galaxy's stellar mass: the total mass of all star particles bound to the sub-halo, those within twice the stellar half-mass radius and those within 30 pkpc. Lines become dashed at the high-mass end when there are fewer than 10 objects per 0.2 dex stellar mass bin. Data points show measurements with 1-sigma error bars from Li & White (2009), D'Souza et al. (2015), Bernardi et al. (2013) and Baldry et al. (2012). The grey solid and dotted lines are the $z = 0$ GSMFs of Illustris (Genel et al. 2014) using the total bound stellar mass of each galaxy and the mass within twice the stellar half-mass radius, respectively. The grey dashed line is the $z = 0.1$ GSMF of EAGLE (Schaye et al. 2015) using the stellar mass within a spherical 30 pkpc aperture. All mass functions assume a Chabrier (2003) initial mass function (IMF).

compared to Illustris. We have chosen somewhat larger softening lengths compared to Illustris and find that this results in fewer low-mass galaxies. Somewhat surprisingly, the difference is not only present near the resolution limit but extends almost two decades in stellar mass. We have tested this explicitly by performing a simulation identical to the FABLE periodic box described in Section 2.1 but with a gravitational softening length that is approximately 2.5 times smaller. Quantitatively, the effect is a systematic decrease of ∼0.2 dex in the mass function for stellar masses ranging from near the resolution limit ($\sim 10^8\,\mathrm{M_\odot}$) to just below the knee ($\sim 10^{10}\,\mathrm{M_\odot}$). This offset is already present at $z = 8$ and persists until $z = 0$. As there is no general consensus regarding the choice of gravitational softening lengths, our improved agreement at the low-mass end of the GSMF is largely serendipitous. We also note that, like Illustris, we are unlikely to be fully converged with respect to resolution.

The knee of the GSMF is in agreement with the data, although it is slightly lower than in Illustris. The difference is largely the result of changes to the stellar feedback model. As discussed in Section 2.3, we have modified the stellar feedback model of Illustris by assigning one-third of the galactic wind energy as thermal. We find that this reduces the abundance of galaxies mostly around the knee of the $z = 0$ GSMF, leaving lower masses ($M_\star \lesssim 10^{10}\,\mathrm{M_\odot}$) relatively unaffected. Physically, star formation is suppressed because warmer winds increase the buoyancy of the gas outflow, thereby reducing the rate at which gas is recycled back on to the galaxy.

The high-mass end of the total and $2\,r_{\star,1/2}$ GSMFs are very similar to those of Illustris, with a slight reduction in the abundance of the most massive galaxies. This is encouraging given that we have greatly reduced the burstiness of the radio mode of AGN feedback compared to Illustris and therefore reduced its ability to suppress star formation in massive haloes. Instead, we have managed to efficiently suppress star formation in massive galaxies through the use of a duty cycle for the quasar mode of feedback. In Appendix A, we show how this affects the high-mass end of the $z = 0$ GSMF and the gas fractions of massive haloes compared with continuous quasar-mode feedback.

The GSMF of EAGLE was calibrated to reproduce the mass function derived by Li & White (2009) for a complete spectroscopic sample from the SDSS (open symbols in Fig. 2). Li & White (2009) measure galaxy flux within a projected Petrosian aperture, which Schaye et al. (2015) show yields a GSMF similar to a spherical aperture of radius 30 pkpc for galaxies in EAGLE. At low masses, $M_\star \lesssim 10^{10}\,M_\odot$, our 30 pkpc aperture mass function is almost identical to that of EAGLE. Around the knee of the GSMF ($M_\star \lesssim 10^{11}\,M_\odot$), we are actually in better agreement with Li & White (2009) compared with EAGLE, which somewhat underestimates the GSMF there. At larger masses, we overestimate Li & White (2009) by ∼0.3 dex, although the difference may be exaggerated as Bernardi et al. (2017) argue that Li & White (2009) use mass-to-light ratios that are biased low for massive galaxies.

At the high-mass end of the GSMF, there is significant variation between different observational studies. Some of the dominant causes for this variation are discussed in Bernardi et al. (2017). In particular, the observed GSMF depends on how the total light associated with a galaxy is determined. The potential impact of this effect is apparent from the simulated total and aperture mass functions shown in Fig. 2, which differ by ∼0.3 dex on the vertical axis at the high-mass end. In addition, there is some freedom in the stellar population modelling used to estimate the mass associated with the total stellar light (i.e. the mass-to-light ratio).

The GSMF derived from the total bound stellar mass of galaxies slightly exceeds observational constraints at the high-mass end, suggesting that the FABLE simulations produce slightly too many massive galaxies. The degree to which this is true depends on what fraction of the total mass in massive galaxies is accounted for in observations. In particular, a significant fraction (∼30 per cent) of the total stellar mass in our massive galaxies is contained in the ICL, the diffuse nature of which makes it difficult to quantify from observations. Studies that measure the galaxy flux within a particular aperture will typically exclude the majority of the ICL associated with massive galaxies. For example, the Petrosian aperture used by Li & White (2009) is known to significantly underestimate the flux of galaxies with extended surface brightness profiles (Blanton et al. 2001; Graham et al. 2005; Bernardi et al. 2010, 2013). More recent studies such as Baldry et al. (2012), Bernardi et al. (2013), and D'Souza et al. (2015) attempt to measure a better estimate of the total flux of galaxies by integrating models fit to their surface brightness distributions. Baldry et al. (2012) fit Sersic profiles to $z < 0.06$ galaxies from the Galaxy And Mass Assembly (GAMA) survey, while Bernardi et al. (2013) fit Sersic-exponential models to a magnitude-limited sample of SDSS galaxies. D'Souza et al. (2015) use the same sample as Li & White (2009) but integrate the galaxy flux from exponential or de Vaucouleurs profile fits and derive flux corrections from stacked SDSS images, which provide a more accurate measurement of the total amount of light owing to the increased signal-to-noise ratio. The extra light returned by this method compared to Li & White (2009) results in a larger abun-

dance of massive galaxies, as evident in Fig. 2. These model profiles can only be fit to the central, high signal-to-noise regions of galaxies and must make assumptions about the outer regions of the galaxy profile. Which profile is the most appropriate at the high-mass end is still debated and can lead to a significant bias in the total estimated flux and resultant stellar mass estimate (see e.g. discussion in Bernardi et al. 2013).

Using an aperture of radius $2\,r_{\star,1/2}$, the FABLE GSMF is in very good agreement with Bernardi et al. (2013) at the high-mass end but is overestimated compared to D'Souza et al. (2015) and Baldry et al. (2012). Bernardi et al. (2017) show that the Sersic-exponential fits used by Bernardi et al. (2013) return galaxy fluxes similar to the corrected fluxes of D'Souza et al. (2015). Similarly, Bernardi et al. (2013) show that their luminosity function is in good agreement with that used in Baldry et al. (2012). This implies that the difference between these three studies at the massive end of the GSMF is due to differences in their assumed mass-to-light ratios rather than their methods for estimating galaxy flux. Bernardi et al. (2017) state that the Bernardi et al. (2013) model is oversimplified and may overestimate the mass-to-light ratio. On the other hand, the stellar population modelling used by D'Souza et al. (2015) results in a mass function that is ∼0.3 dex lower on the vertical axes above $10^{11.5}\,M_\odot$ compared to more recent estimates of the mass-to-light ratio based on the same IMF (Bernardi et al. 2017). Given that there is no consensus as to the best approach to stellar population modelling, in addition to the uncertainty in how to fit the light profiles of massive galaxies, there is arguably little point in tuning the simulated GSMF to a specific data set. Overall, we are satisfied that the difference between the simulated and observed GSMFs is similar to the variation between different observational studies.

### 4.2 GSMF at $z \leq 3$

In Fig. 3, we plot the GSMF at $z \leq 3$ in comparison to observational data from Muzzin et al. (2013) and Ilbert et al. (2013), two independent estimates both based on UltraVISTA DR1, and Tomczak et al. (2014), for the FourStar Galaxy Evolution Survey (ZFOURGE).

We continue to have good agreement with the data beyond $z = 0$. Although the FABLE model for AGN feedback has been calibrated to match the $z = 0$ GSMF, the agreement is not guaranteed at higher redshifts. The high-mass end of the GSMF is in good agreement with the data at each of the redshifts shown, except for a slight underestimate at $z = 2$. This may be due to small number statistics imposed by our finite box size, as the two highest occupied mass bins at $z = 2$ contain only one or two galaxies.

At $z \geq 1$, the low-mass end of the GSMF is somewhat overestimated, although this is not entirely unexpected given that this was also the case in Illustris. We have slightly altered the stellar feedback model of Illustris by implementing thermal winds; however, this largely affects galaxies around the knee of the GSMF rather than low-mass galaxies and any significant changes are limited to $z < 2$.

## 5 GLOBAL GROUP AND CLUSTER PROPERTIES

### 5.1 Stellar mass fractions

In Fig. 4, we plot the stellar mass fraction within $r_{500}$ as a function of halo mass at $z = 0$. We consider all FoF haloes in the periodic volume (open diamonds), plus the main FoF halo in each of the zoom-in simulations (filled diamonds). We do not discriminate between stars
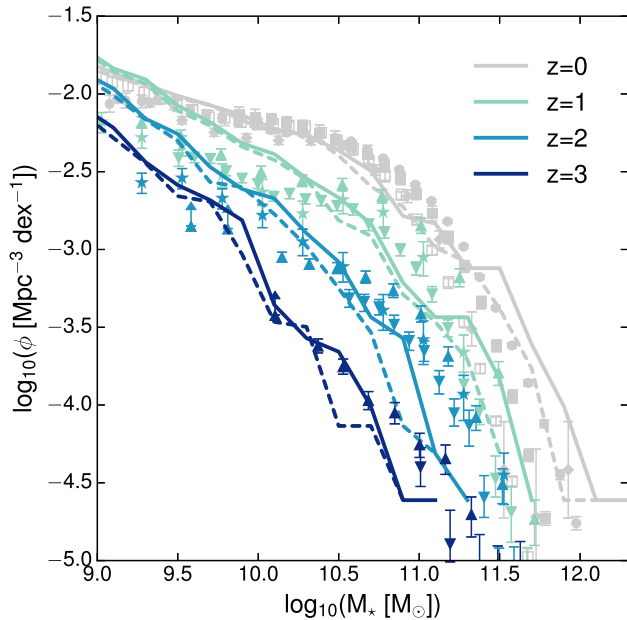
**Figure 3.** The GSMF at redshifts $0 \leq z \leq 3$ (lines) compared to observations (symbols with error bars). Two definitions are used for a simulated galaxy's stellar mass: all stellar mass bound to the sub-halo (solid lines) and bound stellar mass within twice the stellar half-mass radius (dashed lines). The $z = 0$ data are the same as in Fig. 2. At $z = 1$, we compare to observed GSMFs in the redshift ranges $1.0 \leq z < 1.5$, $0.8 < z < 1.1$, and $1.0 < z < 1.25$ for Muzzin et al. (2013) (downward triangles), Ilbert et al. (2013) (upward triangles), and Tomczak et al. (2014) (stars), respectively. At $z = 2$ and $z = 3$, we compare with the GSMFs for redshift ranges $2.0 < z < 2.5$ and $3.0 < z < 3.5$, respectively. Only stellar mass bins above the mass completeness limit are plotted in each case. Stellar masses have been converted to a Chabrier (2003) IMF where necessary by subtracting 0.25 or 0.05 dex for a Salpeter (1955) or Kroupa (2001) IMF, respectively.



**Figure 4.** Stellar mass fraction within $r_{500}$ as a function of halo mass at $z = 0$. Open diamonds show haloes from the $40\,h^{-1}$ Mpc periodic box, while filled diamonds show the main halo of each zoom-in simulation. The solid line shows the mean relation in bins of halo mass for comparison with the mean relation from Illustris at $z = 0$ (grey dashed line). Symbols with error bars show $z \simeq 0$ observations from Gonzalez et al. (2013), Sanderson et al. (2013), and Kravtsov et al. (2018). Following Chiu et al. (2015), the stellar mass measurements of Gonzalez et al. (2013) and Sanderson et al. (2013) have been reduced by 24 per cent to ensure all stellar masses are appropriate for a Chabrier IMF.

in satellite galaxies, the brightest central galaxy (BCG) or the ICL because the distinction between BCG and ICL is not well defined. We therefore compare to studies which take into account contributions from all three components (Gonzalez et al. 2013; Sanderson et al. 2013; Kravtsov, Vikhlinin & Meshcheryakov 2018). The grey dashed line is the mean relation from the Illustris simulation (Genel et al. 2014).

The FABLE simulations are a very good match to observed stellar mass fractions across a wide range of halo masses from galaxy groups (a few times $10^{13}\,\mathrm{M_\odot}$) to high-mass clusters ($\sim 10^{15}\,\mathrm{M_\odot}$). Although our AGN feedback model was tuned to reproduce the $z = 0$ GSMF, the cosmological volume used in the tuning process contains only one halo with $M_{500} \sim 10^{14}\,\mathrm{M_\odot}$. It is therefore reassuring that our model yields a realistic build-up of stellar mass even in dense cluster environments. We reproduce the observed trend with halo mass, with an increase in the stellar fraction towards lower mass systems. The relationship between stellar fraction and halo mass is similar between FABLE and Illustris, with an offset in the normalization. The difference in normalization at $M_{500} \gtrsim 10^{13}\,\mathrm{M_\odot}$ is due to more efficient suppression of star formation by AGN feedback in FABLE, consistent with the offset in the GSMFs. The offset at $M_{500} \lesssim 10^{13}\,\mathrm{M_\odot}$ is partly due to our choice of larger gravitational softening lengths compared with Illustris and partly our change to the stellar feedback model.

We caution that the observations plotted in Fig. 4 derive halo masses from X-ray data, which could potentially underestimate the
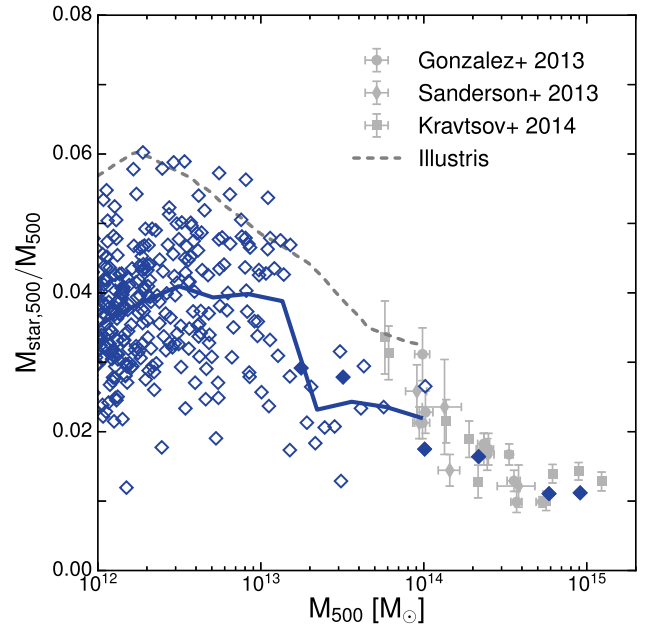
true mass (see Section 3.3). Correcting for such a bias (if present) would shift the observational data to lower stellar fractions, away from our relation. On the other hand, observations may underestimate the stellar mass of the BCG and associated ICL, which can contribute $\gtrsim 50$ per cent of the total stellar mass both observationally (e.g. Gonzalez et al. 2013) and in our simulations. The characteristically diffuse emission of the ICL makes it particularly difficult to quantify, and thus, a non-negligible fraction of a cluster's total stellar content may be missed.

## 5.2 Gas mass fractions

In Fig. 5, we plot the total gas mass fraction within $r_{500}$ as a function of halo mass at $z = 0$. The solid line shows the mean relation in bins of halo mass. We compare to a range of observational data (symbols with error bars) as well as the mean relation in Illustris (grey dashed line). Since observed gas fractions are obtained from X-ray observations, we exclude cold gas and star-forming gas that is followed only with a simplified multiphase model, in accordance with our simulated X-ray spectra (see Section 3.2). For completeness, we also show the mean gas fraction–halo mass relation when all gas is included (dashed line). The difference becomes significant below $\sim 10^{13}\,\mathrm{M_\odot}$; however, there is very little available X-ray data at these masses.

We have calibrated the FABLE AGN feedback model to observed gas fractions only for haloes in the $40\,h^{-1}$ Mpc periodic box, which are represented by the open diamonds in Fig. 5. Although the model was not calibrated to cluster scales ($\gtrsim 10^{14}\,\mathrm{M_\odot}$), the zoom-in simulations (filled diamonds) are in good agreement with the observa-
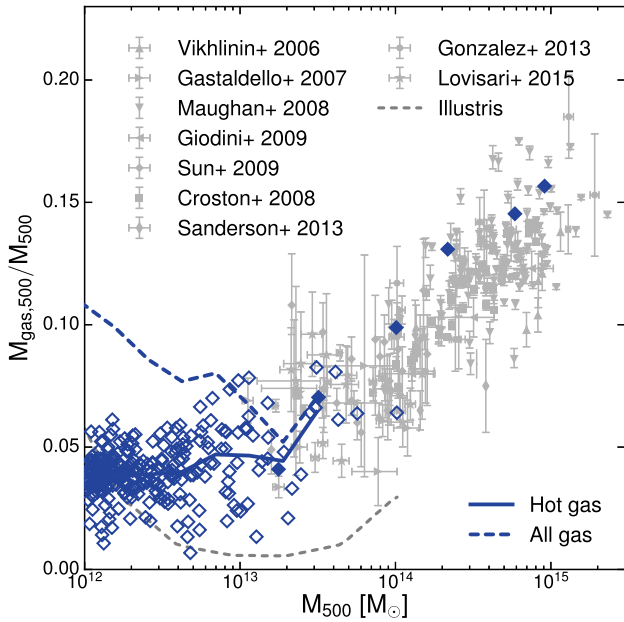
**Figure 5.** Gas mass fraction within $r_{500}$ as a function of halo mass at $z = 0$. Marker styles are the same as in Fig. 4. The solid line shows the mean relation of these points in halo mass bins. Cold and multiphase gas has been excluded as described in Section 3.2. The dashed line shows the mean relation when all gas is included. The grey dashed line is the mean relation from Illustris at $z = 0$ (Genel et al. 2014). Symbols with error bars show $z \simeq 0$ observations from Vikhlinin et al. (2006), Gastaldello et al. (2007), Maughan et al. (2008), Giodini et al. (2009), Sun et al. (2009), Croston et al. (2008), Sanderson et al. (2013), Gonzalez et al. (2013), and Lovisari, Reiprich & Schellenberger (2015).

tions even in massive clusters ($\approx 10^{15}\,\mathrm{M_\odot}$). This was certainly not guaranteed, as the much deeper potentials of clusters make it more difficult for AGN feedback to eject gas beyond $r_{500}$. Indeed, the mean gas fractions of our simulated clusters do lie slightly above the mean observed relation, which suggests that the AGN feedback may need to be slightly more efficient at these high masses. We also note that this offset would be exacerbated in the case of a significant X-ray mass bias, which would shift the observed data away from our relation.

At $M_{500} \approx 10^{13}$–$10^{14}\,\mathrm{M_\odot}$, the mean gas mass fraction in FABLE is significantly higher than in Illustris. In the Illustris model, radio-mode AGN feedback ejected large gas masses out of massive haloes, resulting in significantly underestimated gas fractions. In our updated model, radio-mode feedback events are more frequent but less energetic than in Illustris, and are therefore less able to eject gas from massive haloes. In Appendix A, we show that radio-mode feedback in FABLE is able to lower gas fractions in massive haloes but that it is also assisted by the quasar mode, whose periodic rather than continuous feedback reduces gas fractions by suppressing the accumulation of gas at early times.

### 5.3 X-ray luminosity–mass relation

Here, we present X-ray luminosity as a function of halo mass and gas mass in the FABLE simulations in comparison with observations. X-ray luminosities are calculated as described in Section 3.2 in one of two bands: a soft X-ray band in the range of 0.1–2.4 keV and a bolometric band in the range of 0.01–100 keV. We scale luminosity by $E(z)^{-1}$ and mass by $E(z)$, where $E(z) = \sqrt{\Omega_M(1+z)^3 + (1 - \Omega_M - \Omega_\Lambda)(1+z)^2 + \Omega_\Lambda}$ describes

the redshift evolution of the Hubble parameter. Halo masses are those measured directly from the simulation. The gas mass is the total mass of gas included in the creation of the synthetic X-ray spectrum, i.e. after excluding cold and multiphase gas (see Section 3.2).

In Fig. 6, we plot the $L_{500}$–$M_{500}$ and $L_{500}$–$M_{gas}$ relations in the soft X-ray band. At group scales, we compare to Lovisari et al. (2015), a complete X-ray-selected sample of 20 galaxy groups observed with *XMM–Newton*, and Eckmiller, Hudson & Reiprich (2011), a sample of 26 groups observed with *Chandra*. At the high-mass end, we compare to Pratt et al. (2009) who investigate the luminosity scaling relations of 31 local ($z < 0.2$) clusters from the Representative XMM–Newton Cluster Structure Survey (REXCESS; Böhringer et al. 2007). REXCESS halo masses were estimated iteratively from the $M_{500}$–$Y_X$ relation of Arnaud, Pointecouteau & Pratt (2007) and gas masses are taken from Croston et al. (2008). For the $L_{500}$–$M_{500}$ relation, we supplement the high-mass end with data from Giles et al. (2017), a complete sample of 34 galaxy clusters at $0.15 \leq z \leq 0.3$ observed with *Chandra*. Each of these studies uses X-ray hydrostatic mass estimates.

We also compare to the weak-lensing-calibrated $L_{500}$–$M_{500}$ relations of Leauthaud et al. (2010) and Mantz et al. (2016). Leauthaud et al. (2010) measure mean halo masses via stacked weak gravitational lensing for a sample of 206 X-ray selected groups in the COSMOS field. We have rescaled their mean halo mass, $\langle M_{200} \rangle$, to the mass within $r_{500}$ using the best-fitting Navarro-Frenk-White (NFW; Navarro, Frenk & White 1997) profile of each mass bin. Mantz et al. (2016) measure X-ray luminosities from *Chandra* and *ROSAT* data for 27 clusters with weak-lensing mass estimates as part of the *Weighing The Giants* project (Applegate et al. 2014; Kelly et al. 2014; von der Linden et al. 2014a). We also plot the $L_{500}$–$M_{gas}$ measurements of Mantz et al. (2016) for their full sample of 139 clusters.

Compared with observed $L_{500}$–$M_{500}$ relations based on X-ray hydrostatic masses, the FABLE simulations are in excellent agreement over the full-mass range from groups to massive clusters. However, there is tentative evidence that the predicted relation is overestimated compared with observations based on weak-lensing masses. This may point to the existence of an X-ray mass bias, which would manifest itself as a systematic difference between the $L_{500}$ and $M_{500}$ relations derived from weak-lensing masses and those derived from X-ray hydrostatic masses. If a significant X-ray mass bias is indeed present, the observed halo gas mass fractions shown in Fig. 5 would be shifted to lower values, away from our relation. In this case, the FABLE clusters would be too gas rich and as a result their X-ray luminosities would be too high relative to the weak-lensing-calibrated $L$–$M_{500}$ (assuming weak-lensing masses are less biased). Given that X-ray luminosity is very sensitive to the density distribution of the X-ray emitting gas, the luminosities of the simulated systems could be biased by high-density clumps with high-X-ray emission. However, the excellent agreement with the observed $L_{500}$–$M_{gas}$ relation in the right-hand panel of Fig. 6 suggests that this is not the case, i.e. that the gas content of our simulated haloes has a realistic clumping factor.

In Fig. 7, we show the $L_{500}$–$M_{500}$ and $L_{500}$–$M_{gas}$ relations using the bolometric X-ray luminosity. For the $L_{500}$–$M_{500}$ relation, we compare to REXCESS (Pratt et al. 2009) and Giles et al. (2017), both of whom use X-ray hydrostatic masses. We also compare with weak-lensing-calibrated data from Giles et al. (2016) for the XXL-100-GC sample, which consists of the 100 brightest clusters in the XXL survey (Pacaud et al. 2016). Halo masses for the XXL-100-GC are estimated from the weak-lensing mass–temperature relation
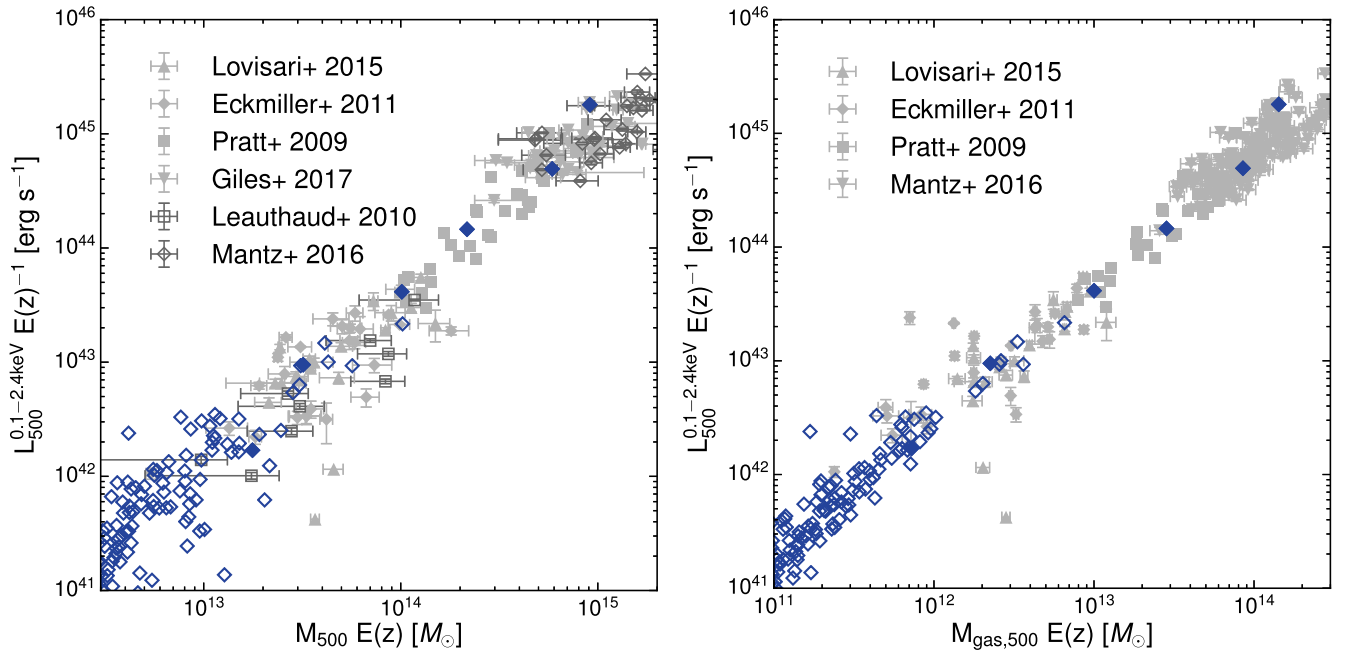
**Figure 6.** X-ray Luminosity in the 0.1–2.4 keV band as a function of total mass (left) and gas mass (right) at $z = 0$. Marker styles are the same as in Fig. 4. In the left-hand panel, light grey symbols with error bars represent observational data for which total masses were estimated from X-ray observations assuming hydrostatic equilibrium (Lovisari et al. 2015; Pratt et al. 2009; Eckmiller et al. 2011; Giles et al. 2017). Dark grey symbols with error bars show observational data for which total masses were estimated from weak gravitational lensing (Leauthaud et al. 2010; Mantz et al. 2016). In the right-hand panel, we compare to X-ray luminosities and gas masses from Lovisari et al. (2015), Eckmiller et al. (2011), Pratt et al. (2009), and Mantz et al. (2016).

presented in Lieu et al. (2016). For $L_{500}$–$M_{gas}$, we complement the low-mass end with data from Zhang et al. (2011), who analyse 62 galaxy clusters in the HIFLUGCS sample with *XMM–Newton* and *ROSAT* data. At the high-mass end of $L_{500}$–$M_{gas}$, we compare with Mahdavi et al. (2013), a sample of 50 clusters with weak-lensing mass estimates and X-ray data from *Chandra* or *XMM–Newton*. Note that the luminosity and gas mass for this sample are measured within $r_{500}$ as derived from weak lensing.

Again we achieve excellent agreement with the relation based on X-ray hydrostatic masses. However, as was hinted at in the soft-band $L_{500}$–$M_{500}$ relation, there appears to be a systematic offset between observed bolometric $L_{500}$–$M_{500}$ relations based on weak-lensing masses and those based on X-ray hydrostatic masses. Under the assumption that weak-lensing masses are less biased, this suggests that the X-ray luminosities of the FABLE groups and clusters may be slightly too high for a given halo mass. This conclusion is, however, complicated by the relatively large measurement uncertainties and overall scatter in weak-lensing mass estimates compared to X-ray masses (see e.g. Meneghetti et al. 2010) and the relatively poor sampling of the $L_{500}$–$M_{500}$ plane by weak-lensing data.

### 5.4 X-ray luminosity–temperature relation

In Fig. 8, we plot bolometric X-ray luminosity as a function of X-ray spectroscopic temperature. In the left-hand and right-hand panels, we consider luminosities and temperatures measured with and without the cluster core, respectively. We define a core radius of 0.15 $r_{500}$ as used in the comparison studies. We use the spectroscopic temperature estimates of FABLE groups and clusters for consistency with the X-ray luminosity determinations, although we show in Appendix B2 that the difference between spectroscopic and mass-

weighted temperature is small ($\lesssim 0.1$ dex) for temperatures above ≈0.5 keV.

We compare against data from Osmond & Ponman (2004), Maughan et al. (2012), Zou et al. (2016), REXCESS (Pratt et al. 2009), and Sun (2012). The Osmond & Ponman (2004) sample contains 35 systems with group-scale X-ray emission observed with *ROSAT*. Although not statistically representative, the sample corresponds to some of the lowest temperature systems that have been observed. The data from Maughan et al. (2012), Zou et al. (2016), and Sun (2012) each consist of groups and clusters observed with *Chandra*. The Maughan et al. (2012) sample consists of 114 clusters originally described in Maughan et al. (2008), the Zou et al. (2016) sample is statistically complete and contains 23 groups and low-mass clusters from the 400 d survey (Burenin et al. 2007), and Sun (2012) presents the X-ray luminosity–temperature relation of the Sun et al. (2009) sample of 43 galaxy groups with which we compare radial profiles of the ICM in Section 6.

We find that the FABLE groups and clusters lie on the upper end of the scatter in the observations. This is true whether or not the cluster core is excised, which implies that our X-ray luminosities or temperatures are not biased by, for example, an overabundance of dense cool cores. At $T_{500} \sim 1$ keV, we are actually in good agreement with the Sun (2012) data; however, we expect that their sample is biased towards high-X-ray luminosities due to their selection criteria, which require that the group emission can be traced to at least $r_{2500}$ (≈0.47 $r_{500}$; Sun et al. 2009). Indeed, their average luminosity at $T_{500} \sim 1$ keV is notably higher than the Osmond & Ponman (2004) and Zou et al. (2016) samples. Similarly, Le Brun et al. (2014) find that the Sun et al. (2009) sample has a significantly higher mean X-ray luminosity for groups with masses $M_{500} \sim 10^{13-13.5}$ M$_\odot$ compared with the galaxy group studies of
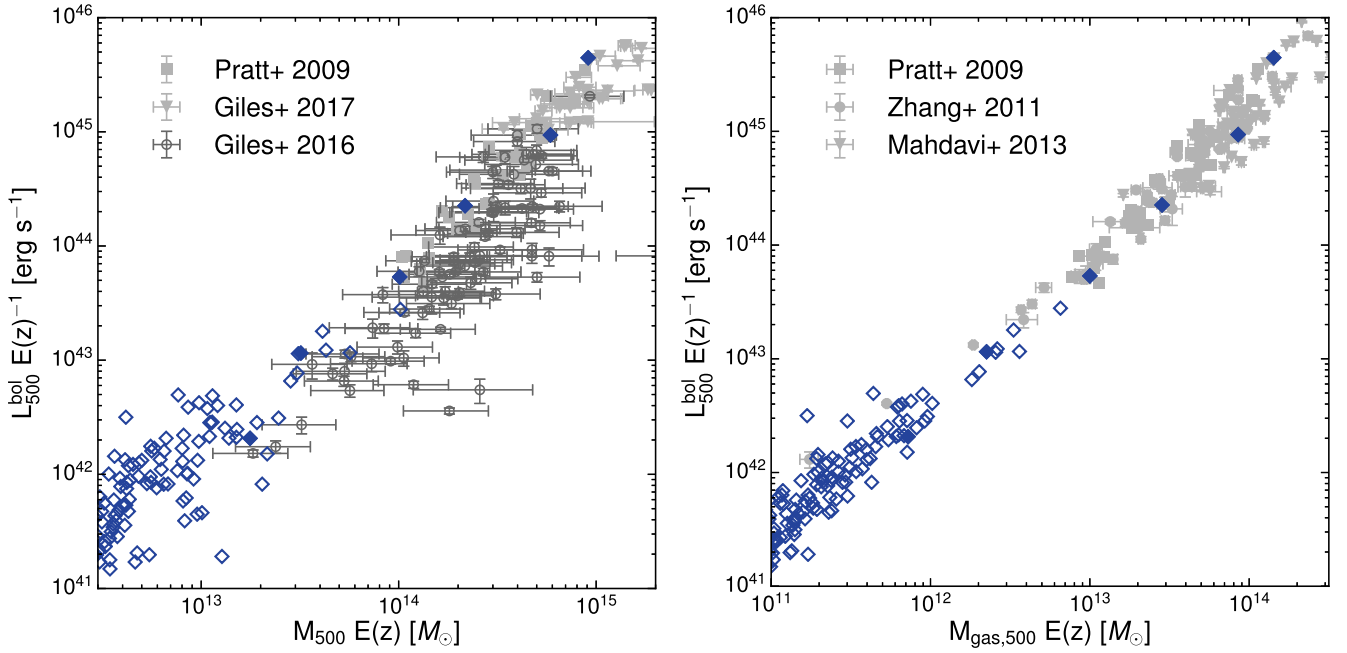
**Figure 7.** Bolometric (0.01–100 keV) X-ray luminosity as a function of total mass (left) and gas mass (right) at $z = 0$. In the left-hand panel, light grey symbols with error bars represent observational data based on X-ray hydrostatic masses (Pratt et al. 2009; Giles et al. 2017). Dark grey symbols with error bars are the XXL-100-GC clusters (Giles et al. 2016) for which total masses were estimated from the internally calibrated weak-lensing mass–temperature relation presented in Lieu et al. (2016). In the right-hand panel, we compare with data from Pratt et al. (2009), Zhang et al. (2011), and Mahdavi et al. (2013).
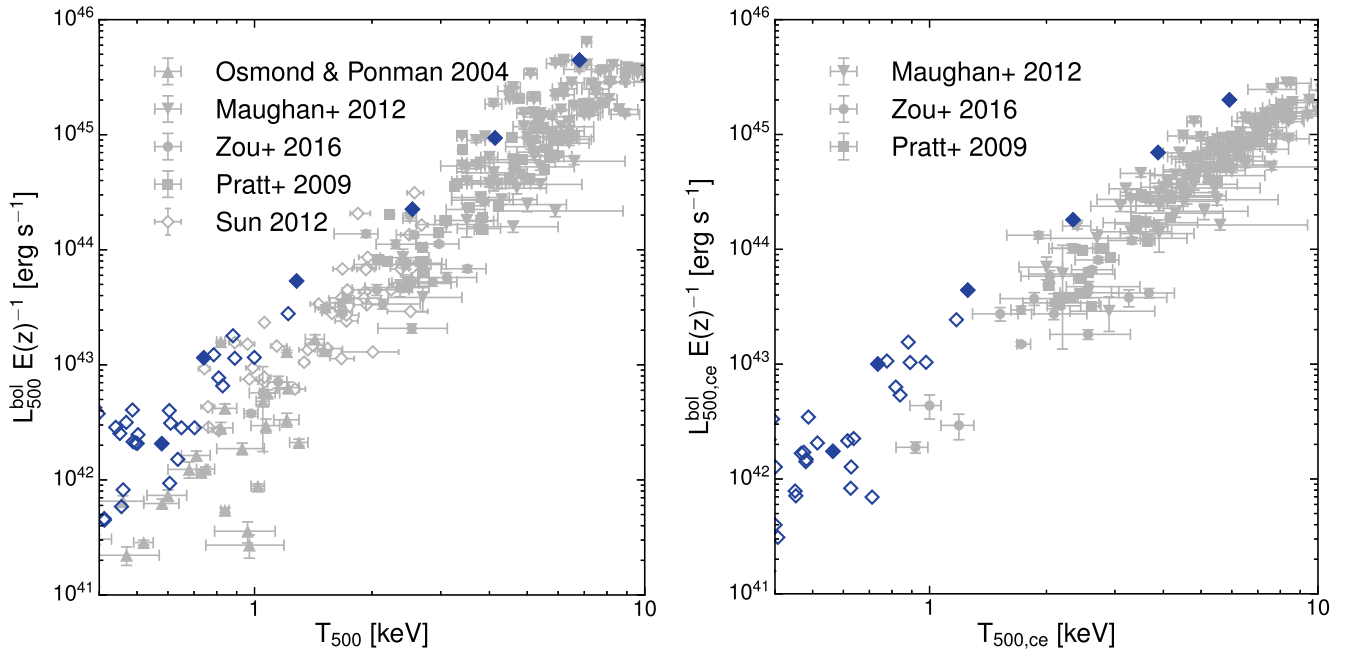


**Figure 8.** Bolometric (0.01–100 keV) X-ray luminosity as a function of spectroscopic temperature measured in the $(0 - 1)$ $r_{500}$ aperture (left) and the $(0.15 - 1)$ $r_{500}$ aperture (right) at $z = 0$. Data from several X-ray studies are shown for comparison (Osmond & Ponman 2004; Pratt et al. 2009; Maughan et al. 2012; Sun 2012; Zou et al. 2016).

Osmond & Ponman (2004), Rozo et al. (2009), and Leauthaud et al. (2010).

It is not clear whether the small discrepancy between the simulation prediction and the observed relations is a result of overestimated total luminosities, underestimated global temperatures, or a combination of the two. The former explanation may be consistent with

our comparison to weak-lensing studies of the $L_{500}$–$M_{500}$ relation (Figs 6 and 7), for which there is some evidence that our luminosities are too high for a given halo mass. However, the poor sampling and relatively large scatter of the weak-lensing measurements means that we cannot rule out the possibility that the gas in our groups and clusters possess too low average temperatures. In the following

section, we therefore investigate the role of temperature using the total mass–temperature relation.

## 5.5 Mass–temperature relation

In Fig. 9, we show the total mass as a function of spectroscopic temperature in FABLE compared to observational samples that use either X-ray hydrostatic masses (left-hand panel) or weak-lensing masses (right-hand panel).

For the X-ray mass comparison, we show data from Lovisari et al. (2015), Eckmiller et al. (2011), REXCESS (Pratt et al. 2009), and Mahdavi et al. (2013). We calculate the spectroscopic temperature within $r_{500}$, appropriate for REXCESS and Mahdavi et al. (2013), but caution that the data of Lovisari et al. (2015) and Eckmiller et al. (2011), which populate group-scale masses, cover only a fraction of $r_{500}$ and this fraction is not consistent between different systems.

For the weak-lensing mass comparison, we show the mass–temperature relation derived by Lieu et al. (2016) for 38 clusters from the XXL-100-GC sample that lie within the footprint of the Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS), using the CFHTLenS shear catalogue for the mass measurements (Heymans et al. 2012; Erben et al. 2013). As in Lieu et al. (2016), we measure spectroscopic temperatures within a projected aperture of radius 300 pkpc. We also show the best-fitting $M_{500}$–$T_{300\mathrm{kpc}}$ relation from Lieu et al. (2016) for an extended sample including 10 galaxy groups from COSMOS (Kettula et al. 2013) and 48 massive clusters from the Canadian Cluster Comparison Project (CCCP; Hoekstra et al. 2015).

We find that the FABLE groups and clusters lie systematically above the $M_{500}$–$T_{500}$ relation derived from X-ray masses (left-hand panel of Fig. 9). As the size of the offset is similar to the mean offset between the simulated and observed $L_{500}$–$T_{500}$ relations, both discrepancies could be the result of systematically low average temperatures. On the other hand, the FABLE systems are a good match to the weak-lensing-calibrated $M_{500}$–$T_{300\mathrm{kpc}}$ relation from Lieu et al. (2016) (the right-hand panel of Fig. 9), which would otherwise suggest that their average temperatures are not significantly biased. Although the statistical uncertainties are large, the Lieu et al. (2016) $M_{500}$–$T_{300\mathrm{kpc}}$ relation has a significantly higher normalization than the $L_{500}$–$T_{500}$ relations based on X-ray hydrostatic masses. The difference in aperture may have a small effect at high temperatures, since a 300 pkpc aperture is smaller than $r_{500}$ for the majority of $T > 1$ keV systems and could yield a slightly higher temperature measurement due to the declining temperature profiles of such systems. However, this would lower rather than boost the normalization of the $M_{500}$–$T$ relation. Furthermore, the difference between these measures is small in our simulations ($\lesssim 0.1$ dex) and Giles et al. (2016) find no systematic differences between the two temperatures for the XXL-100-GC sample. This implies that the offset between the X-ray and weak-lensing calibrated $M_{500}$–$T$ relations is due to an offset in halo mass. Under the assumption that weak-lensing masses are unbiased, the Lieu et al. (2016) results suggest that X-ray hydrostatic masses are biased low and our agreement with their results suggests that FABLE systems possess realistic global temperatures.

We find that the simulation predictions of the $M_{500}$–$T$ relation are rather robust with respect to variations of the feedback modelling. The better agreement of the simulations with the weak-lensing-calibrated relation may hence provide further circumstantial evidence that weak-lensing-mass measurements are less biased. We caution, however, that changes in the feedback modelling beyond those considered here may results in larger variations in the predicted normalization.

## 5.6 SZ–mass relation

In Fig. 10, we plot the mean tSZ flux–halo mass relation calculated as described in Section 3.4. As in Planck Collaboration XI (2013), we self-similarly scale the tSZ flux to redshift $z = 0$ and to a fixed angular diameter distance of 500 Mpc, yielding the tSZ signal, $\tilde{Y}_{5r_{500}}$, in units of square arcminutes.

The FABLE simulations produce a power-law relation extending from massive galaxies to clusters in good agreement with both the original Planck Collaboration XI (2013) relation, which is based on halo masses derived from a semi-analytic galaxy formation model, and the weak-lensing-calibrated relation from Wang et al. (2016). At $\sim 5 \times 10^{12}\,\mathrm{M}_\odot$, there is a slight upturn in the observed relation not seen in the simulations; however, these two mass bins correspond to detections of less than $2\sigma$. Furthermore, Planck Collaboration XI (2013) state that the three lowest mass bins are noticeably affected by dust contamination and may be more uncertain than the statistical measures indicate. Indeed, Greco et al. (2015) explicitly model the dust emission from each LBG in the sample and find that, for the low-mass LBGs with $M_{500} \lesssim 10^{13.3}\,\mathrm{M}_\odot$, the stacked signal from dust emission is comparable to or larger than the stacked tSZ signal.

# 6 ICM PROFILES

## 6.1 Density profiles

In Fig. 11, we show spherically averaged radial density profiles of the ICM in FABLE groups and clusters at $z = 0$. In the left-hand panel, we compare group-scale systems to the density profiles of the Sun et al. (2009) sample of groups observed with *Chandra*. The Sun et al. (2009) sample consists of 43 groups with X-ray hydrostatic mass estimates in the range $1.6 \times 10^{13}\,\mathrm{M}_\odot \leq M_{500} \leq 1.8 \times 10^{14}\,\mathrm{M}_\odot$ with a median mass of $7.3 \times 10^{13}\,\mathrm{M}_\odot$. The sample was drawn from *Chandra* archival data with the requirement that gas properties could be derived out to at least $r_{2500}$ ($\approx 0.47\ r_{500}$). We compare to a halo mass-selected sample of simulated groups with $3.2 \times 10^{13}\,\mathrm{M}_\odot \leq M_{500} \leq 2.2 \times 10^{14}\,\mathrm{M}_\odot$ and a median mass of $5.7 \times 10^{13}\,\mathrm{M}_\odot$. The simulated density profiles are in good agreement with the Sun et al. (2009) profiles beyond $\sim 0.3\ r_{500}$ but slightly underestimate the median density at smaller radii. At $\gtrsim 0.3\ r_{500}$, the agreement is consistent with our match to observed gas mass fractions (see Fig. 5), as this region contains $\sim 90$ per cent of the total gas mass.

In Section 5, we argued that if there is a significant X-ray mass bias then the FABLE systems may be overluminous in X-rays and that this may be the result of an overabundance of gas (the observed gas fractions in Fig. 5 being biased high). In this case, we would expect to overestimate the gas density in the outer regions compared to the data. The fact that this is not seen does not, however, rule out the possibility of a significant X-ray mass bias. One explanation is the difference between the halo mass distributions of the observed and simulated samples. Although we have tried to match the median halo masses of the observed and simulated samples as closely as possible given the small sample size, the median mass of the simulated sample is 22 per cent lower than that of the Sun et al. (2009) sample and would decrease further in the case of an X-ray mass bias (a 45 per cent difference if X-ray masses are biased low by 30 per cent). As the gas content of massive haloes is a relatively strong function of their halo mass (see e.g. Fig. 5), this could result in a mismatch between the two samples. Another explanation is that the average X-ray luminosity of the Sun et al. (2009) sample is biased high due to their requirement that group emission be
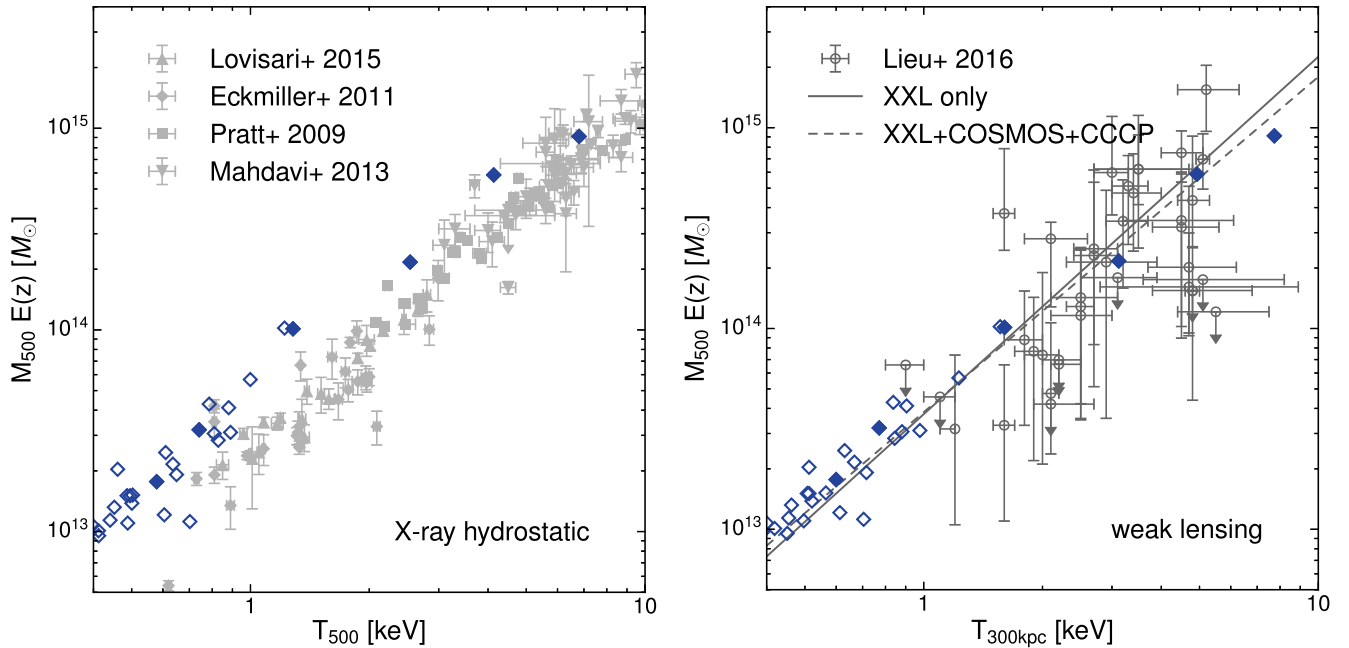
**Figure 9.** Total mass as a function of spectroscopic temperature compared to observations based on X-ray hydrostatic masses (left) and weak-lensing masses (right). In the left-hand panel, temperatures are measured within a projected aperture of radius $r_{500}$. Symbols with error bars show data from Lovisari et al. (2015), Eckmiller et al. (2011), Pratt et al. (2009), and Mahdavi et al. (2013). In the right-hand panel, we compare to Lieu et al. (2016) whom measure weak-lensing masses for a sample of the brightest clusters in the XXL survey (symbols with error bars). We mimic Lieu et al. (2016) by measuring temperatures within a projected aperture of radius 300 pkpc. The solid line is the best fit to the XXL data and the dashed line is the best fit to an extended sample including additional groups and clusters from COSMOS (Kettula et al. 2013) and CCCP (Hoekstra et al. 2015).
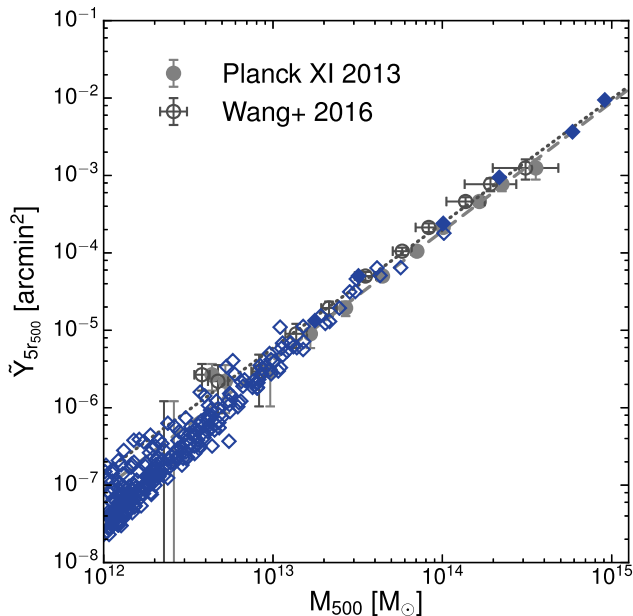


**Figure 10.** tSZ flux as a function of halo mass at $z = 0$. The tSZ flux is calculated as described in Section 3.4 and has been scaled to $z = 0$ and a fixed angular diameter distance of 500 Mpc. Symbols with error bars show the mean tSZ signal in bins of halo mass for a sample of SDSS LBGs presented in Planck Collaboration XI (2013). Filled circles correspond to halo masses derived by Planck Collaboration XI (2013) from a semi-analytic galaxy formation model and open circles show the recalibrated halo masses and associated uncertainties determined by Wang et al. (2016) using stacked weak-lensing analyses. The dashed line shows the best-fitting relation from Planck Collaboration XI (2013), and the dotted line shows the best-fitting relation for the Wang et al. (2016) recalibration.

traced out to a significant fraction of $r_{500}$. Indeed, as we discussed in Section 5.4, the Sun et al. (2009) groups with $T_{500} \sim 1$ keV possess slightly higher X-ray luminosities compared to the Osmond & Ponman (2004) and Zou et al. (2016) samples, such that the $L_{500}^{bol}$–$T_{500}$ relation of our simulated groups is actually in good agreement with the Sun et al. (2009) data. This could explain why our simulated groups match the density profile of the Sun et al. (2009) groups yet seem to overestimate the X-ray luminosity relative to other studies.

At $r \lesssim 0.3 \, r_{500}$ the simulated density profiles lie within the observed scatter but largely fall below the median observed profile. This suggests that our AGN feedback model may displace slightly too much gas from the central regions. On the other hand, the observed groups are detected via their X-ray flux, which may preferentially select systems with high central densities compared to our halo mass-selected sample, particularly if the Sun et al. (2009) sample is biased towards high luminosities relative to other X-ray selected samples.

In the right-hand panel of Fig. 11, we compare our simulated clusters to the density profiles of the REXCESS clusters, a representative sample of 31 clusters observed with *XMM–Newton* (Böhringer et al. 2007; Croston et al. 2008). The REXCESS clusters were chosen such that $r_{500}$ lies well within the field of view of *XMM–Newton*, allowing detailed local background modelling and increased measurement precision at large radii (Croston et al. 2008). The sample is unbiased with respect to cluster morphology or dynamical state, containing a representative distribution of relaxed, cool-core, and morphologically disturbed clusters (as defined in Pratt et al. 2009), which correspond to the solid, dotted, and dashed lines in Fig. 11, respectively. The REXCESS sample covers the mass range of $1.0 \times 10^{14} \, M_\odot \leq M_{500} \leq 7.8 \times 10^{14} \, M_\odot$ with a median mass of $2.6 \times 10^{14} \, M_\odot$. Our comparison sample consists of all five FABLE
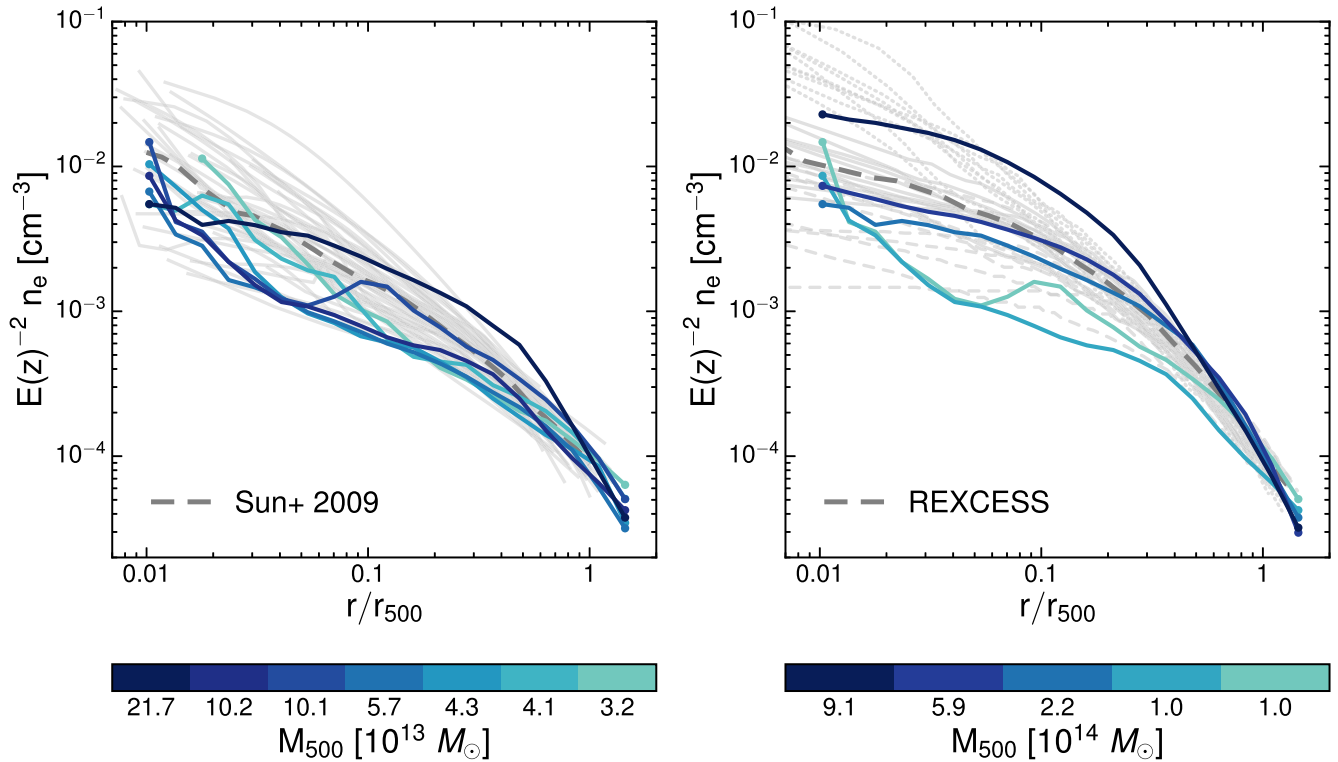
**Figure 11.** Density profiles of the ICM in FABLE groups and clusters at $z = 0$ in comparison to observed profiles. Lines show individual profiles of simulated systems colour coded by halo mass. All profiles have been self-similarly scaled in redshift. In the left-hand panel, we plot the density profiles of the Sun et al. (2009) galaxy groups (grey lines) and the profiles of a mass-selected sample of simulated group-scale systems with a similar median halo mass (see the main text). The thick dashed line shows the median of the Sun et al. (2009) sample. In the right-hand panel, we show density profiles for all cluster-scale systems with $M_{500} \geq 10^{14}$ M$_{\odot}$ in comparison to those of the REXCESS cluster sample (grey lines; Croston et al. 2008). For the observed sample, solid lines correspond to relaxed clusters, dashed lines to disturbed clusters, and dotted lines to cool-core clusters according to the definitions of Pratt et al. (2009). The thick dashed line shows the median REXCESS profile.

clusters with $M_{500} \geq 1.0 \times 10^{14}$ M$_{\odot}$ and has a comparable median mass of $2.1 \times 10^{14}$ M$_{\odot}$.

Overall the density profiles of our simulated clusters are a good match to the REXCESS clusters. At $r \gtrsim 0.3\,r_{500}$, the densities of the three most massive FABLE clusters are somewhat high compared to the median observed profile, which is consistent with the excess gas we might expect in our simulated clusters in the case of an X-ray mass bias. Indeed, the cumulative gas fraction (not shown) rises more steeply between $\sim 0.1$ and $0.6\,r_{500}$ than the REXCESS profiles (Pratt et al. 2010). This suggests that AGN feedback may act too violently, pushing gas from the cluster centre and causing it to pile up at larger radii. A similar trend was found for clusters in IllustrisTNG (Barnes et al. 2017a).

The simulated profiles are similar in shape to the relaxed or disturbed REXCESS clusters. Of the five simulated clusters only one has a central density comparable to observed (weak) cool-core clusters. Potentially, heating of the cluster centre by AGN feedback might be preventing cool-cores from forming in the same proportion as observed at $z \sim 0$ (10 of 31 REXCESS clusters and $\sim 30$–40 per cent in SZ surveys, e.g. Planck Collaboration XI 2011; Andrade-Santos et al. 2017; Rossetti et al. 2017). A larger sample will be needed to assess this in detail. Reproducing the observed fraction of cool-core galaxy clusters in cosmological simulations with feedback is a notoriously difficult problem (e.g. Borgani & Kravtsov 2011; Kravtsov & Borgani 2012). There has been some recent success in this area (e.g. Rasia et al. 2015; Barnes et al. 2017a; Hahn et al. 2017); however, these simulations tend to underestimate the

observed cool-core fraction when compared to low-redshift SZ-selected samples. It is clear that an AGN feedback model that is capable of reproducing the global properties of clusters does not necessarily provide an effective description of the physical processes responsible for the creation of cool cores: additional processes such as AGN-driven turbulence, cosmic rays, stabilization from magnetic fields or anisotropic thermal conduction may be required.

## 6.2 Temperature profiles

In Fig. 12, we plot the dimensionless temperature profiles at $z = 0$ for the same group- and cluster-scale samples described in the previous section. We facilitate a comparison between different halo masses by normalizing the profiles by the characteristic temperature,

$$kT_{500} = \mu m_p G M_{500} / 2 r_{500}, \qquad (2)$$

the temperature of an isothermal sphere of mass $M_{500}$ and radius $r_{500}$. Here, $\mu$ is the mean molecular weight, which we take as $\mu = 0.59$, and $m_p$ is the proton mass.

In the left-hand panel of Fig. 12, we compare to the deprojected temperature profiles of the Sun et al. (2009) groups. We find that beyond the core ($\gtrsim 0.2\,r_{500}$) the temperature profiles of the FABLE groups have a similar slope to the median observed profile. The normalization is somewhat lower than observed; however, this may be the result of an X-ray hydrostatic mass bias in the Sun et al. (2009) mass estimates: since the halo mass, $M_{500}$, is related to the
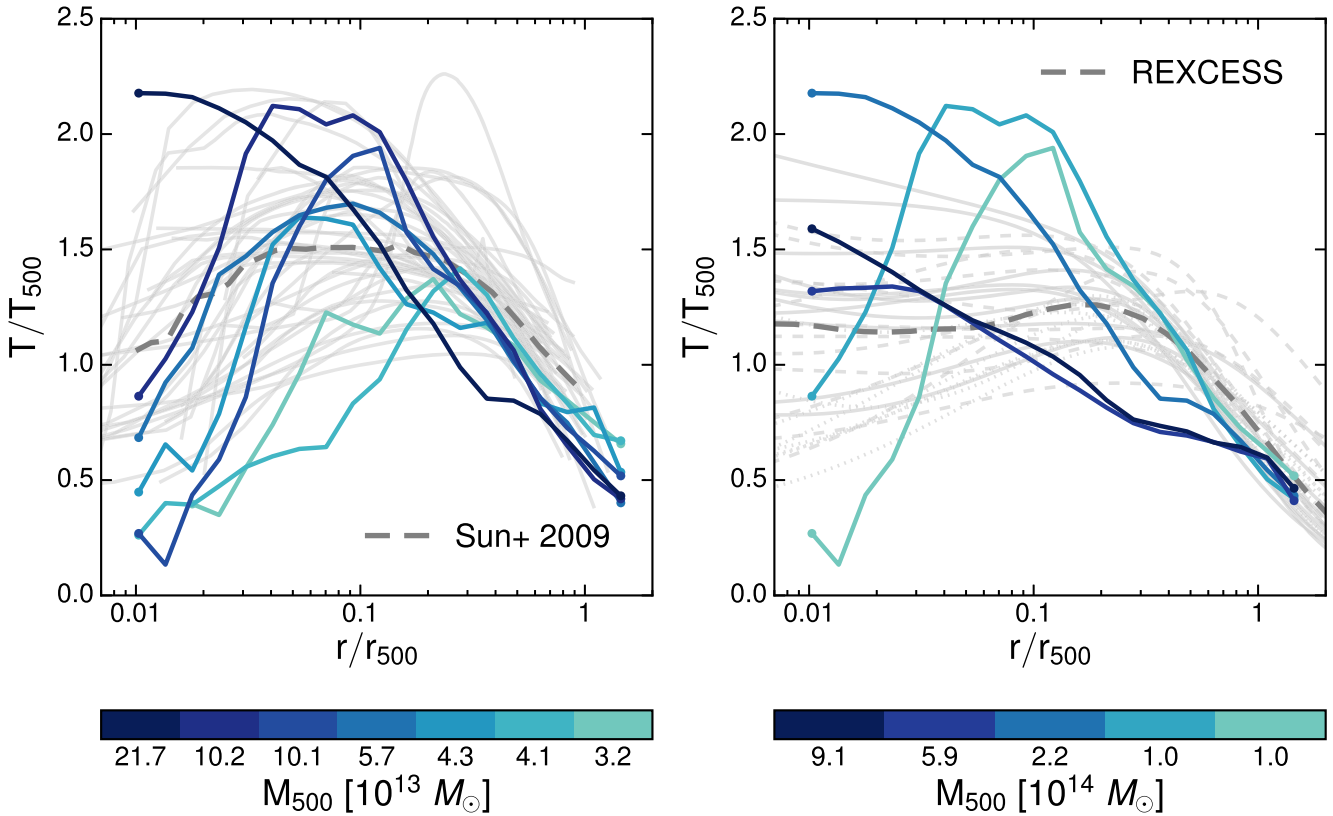
**Figure 12.** Dimensionless temperature profiles at $z = 0$ for the same simulated and observed samples as shown in Fig. 11. Profiles are normalized by the characteristic temperature defined in equation (2). In the left-hand panel, solid grey lines show the individual deprojected temperature profiles of the Sun et al. (2009) sample. The thick dashed line shows the median profile. In the right-hand panel, grey lines show individual temperature profiles for REXCESS clusters derived from the best-fitting pressure (Arnaud et al. 2010) and entropy (Pratt et al. 2010) profiles. Line styles are the same as shown in Fig. 11.

characteristic temperature scaling as $T_{500} \propto M_{500}^{2/3}$, a bias towards lower masses would shift the observed dimensionless temperature profiles to higher values. Within the core, the simulated temperature profiles show similar scatter to observed and in general follow the same shape as the observed profiles.

For the most massive simulated system in the group sample, the temperature rises steadily towards the centre rather than dropping within the core. This may be a side effect of our relatively simple model for radio-mode AGN feedback, which injects bubbles of thermal energy at irregular intervals. In reality, such bubbles are expected to be supported by non-thermal pressure from, e.g. cosmic rays and should only contribute to the observed temperature profile once the injected energy has thermalized.

In the right-hand panel of Fig. 12, we compare to the dimensionless temperature profiles of the REXCESS clusters. The REXCESS temperature profiles rise with roughly constant slope from $r_{500}$ down to $\sim 0.3\,r_{500}$ before dropping slowly or levelling out within the cluster core. Close to $r_{500}$, there is a slight offset between the predicted and observed temperatures. As for the Sun et al. (2009) group comparison, this could be explained by a bias in the X-ray hydrostatic mass estimates. Between $\sim 0.3$ and $1\,r_{500}$, three of the five simulated temperature profiles show a similar slope to observed but for the two most massive clusters in the sample the slope is somewhat shallower.

The REXCESS temperature profiles are not as strongly peaked as the Sun et al. (2009) profiles. Hence, the two FABLE systems with $M_{500} \sim 1 \times 10^{14}\,M_\odot$, which are included in both the group- and cluster-scale samples, are a reasonable match to one or more of

the Sun et al. (2009) systems but reach higher peak temperatures compared to the REXCESS clusters. This is unsurprising given that the halo masses of these systems are on the boundary between groups and clusters and correspond only to the very lowest masses in the REXCESS sample. The three most massive FABLE clusters possess temperature profiles that continue to rise within the cluster core ($\lesssim 0.1\,r_{500}$), unlike the observed temperature profiles, which tend to level out. As noted above, this could be because AGN bubble feedback is modelled via injection of thermal energy rather than of non-thermal components. This may result in an overheating of the ICM in the central regions.

### 6.3 Entropy profiles

In Fig. 13, we plot the dimensionless entropy profiles at $z = 0$ for the same group- and cluster-scale samples described in Section 6.1. We use the widely adopted definition of ICM 'entropy', $K = kT/n_e^{2/3}$, and normalize by the characteristic entropy scale, $K_{500} = kT_{500}/n_{e,500}^{2/3}$, which reflects the mass variation expected in a self-similar model. The characteristic temperature $kT_{500}$ is defined in equation (2). $n_{e,500}$ is the mean electron density within $r_{500}$ and is defined as

$$n_{e,500} = 500 f_b \rho_c(z)/(\mu_e m_p),\qquad(3)$$

assuming the global baryon fraction $f_b = \Omega_b/\Omega_M$ in a universe with critical density $\rho_c(z)$ at redshift $z$ and a mean molecular weight per free electron $\mu_e$. We adopt a value of $\mu_e = 1.14$ and use $f_b$ and $\rho_c(z)$ corresponding to our assumed cosmology.
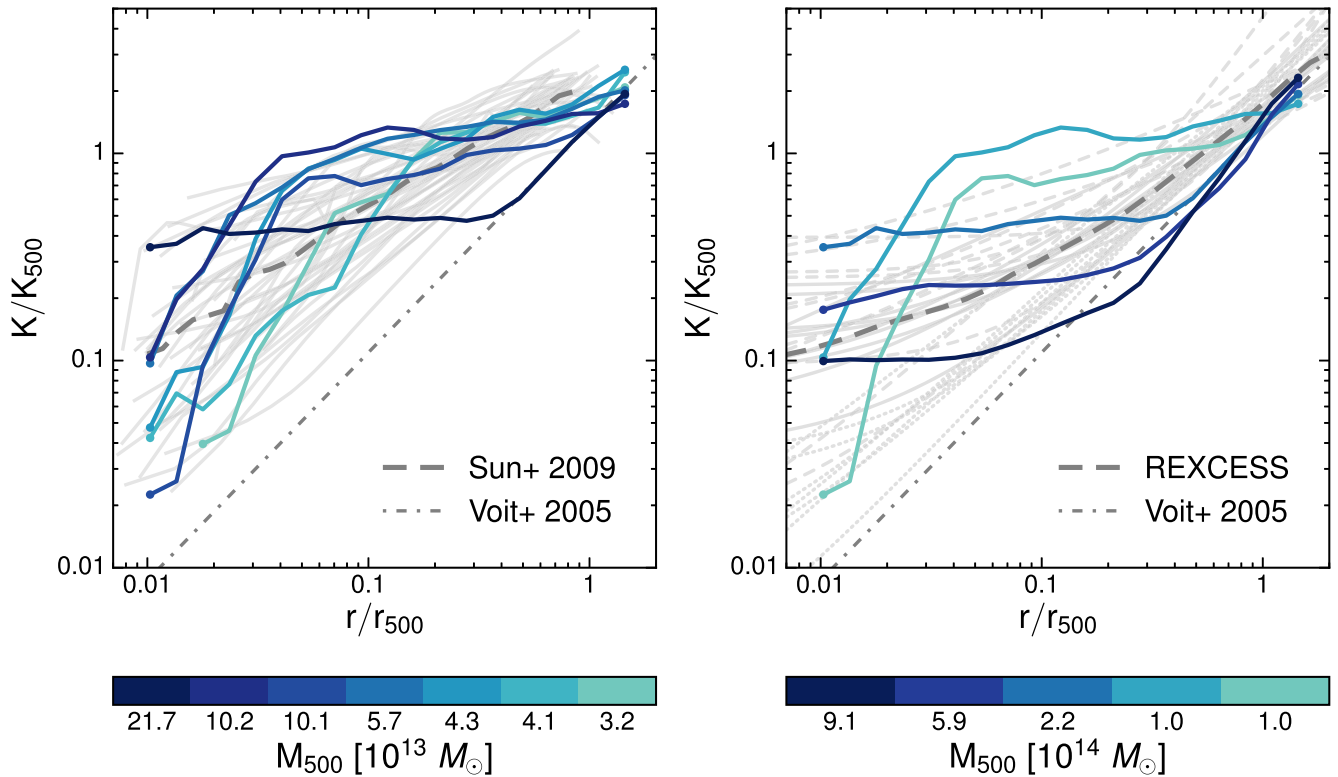
**Figure 13.** Dimensionless entropy profiles at $z = 0$ for the same simulated and observed samples as shown in Fig. 11. Profiles are scaled by the characteristic entropy scale $K_{500}$ as defined in the text. The dash–dotted line shows the baseline ICM entropy profile derived by Voit (2005) from non-radiative simulations. In the left-hand panel, grey lines show the entropy profiles of the Sun et al. (2009) groups derived from the density and temperature profiles shown in Figs 11 and 12. In the right-hand panel, grey lines show the best-fitting entropy profiles of the REXCESS clusters (Pratt et al. 2010). Line styles are the same as shown in Fig. 11.

In the left-hand panel of Fig. 13, we plot the entropy profiles of the group-scale systems. Close to $r_{500}$ the entropy profiles tend to the power-law prediction from Voit (2005), which was derived from hydrodynamic simulations in the absence of non-gravitational processes. At smaller radii, the entropy profiles deviate from this relation due to non-gravitational processes such as AGN feedback, which ejects and heats gas in the group centre, and star formation, which removes low entropy gas. The simulated entropy profiles are in good agreement with the Sun et al. (2009) profiles over the full range of radii. The exception is the most massive system, which follows the non-radiative relation before quickly flattening into an isentropic core at $\sim 0.4\, r_{500}$. This is consistent with the temperature profile of this system, which rises steadily towards the cluster centre instead of dropping slowly within the core.

In the right-hand panel of Fig. 13, we see that the observed entropy profiles of the REXCESS clusters run approximately parallel to the Voit (2005) relation at large radii and slowly flatten towards the cluster centre. This change in slope occurs less rapidly than in the less massive Sun et al. (2009) groups, since the deeper potentials of more massive haloes make non-gravitational processes such as AGN feedback less effective. The cool-core REXCESS clusters have mostly power-law-like entropy profiles while the disturbed clusters tend to have cored entropy profiles.

The two FABLE clusters with $M_{500} \sim 1 \times 10^{14}\, \mathrm{M_\odot}$ have fairly flat entropy profiles down to $\sim 0.05\, r_{500}$ before dropping rapidly within the core. These profiles lie within the intrinsic scatter of the Sun

et al. (2009) sample but are unlike the entropy profiles of the REXCESS clusters. This suggests that these systems are more similar to high-mass groups than low-mass clusters. For the three most massive FABLE clusters, the entropy at $\gtrsim 0.3\, r_{500}$ closely follows the baseline relation of Voit (2005). This suggests that AGN feedback has little effect on the thermodynamics of the gas at large radii. The observed profiles lie somewhat higher than the baseline and the simulations; however, the observed profiles may be slightly overestimated in the case of an X-ray mass bias. At smaller radii, the three most massive clusters show flat central entropy distributions. These cored entropy profiles are characteristic of many of the relaxed and disturbed REXCESS clusters, although the cores extend to somewhat larger radii than observed, which is likely related to the slight overprediction of the gas density in these systems at $\sim 0.3\, r_{500}$ (see Section 6.1). This is consistent with a picture in which AGN feedback in FABLE clusters is overly effective at heating and expelling gas in the central regions but is relatively ineffective at large radii.

### 6.4 Pressure profiles

In Fig. 14, we plot the pressure profiles of the ICM at $z = 0$ for the group- and cluster-scale samples described in Section 6.1. We normalize the pressure profiles, $P(r) = kT(r)n_e(r)$, by the characteristic pressure, $P_{500} = kT_{500}n_{e,500}$, where $kT_{500}$ and $n_{e,500}$ are defined in equations (2) and (3).

In the left-hand panel of Fig. 14, we compare to the Sun et al. (2009) pressure profiles. At $r \gtrsim 0.3\, r_{500}$, the simulated profiles are a good match to the data. Similar to the dimensionless temperature
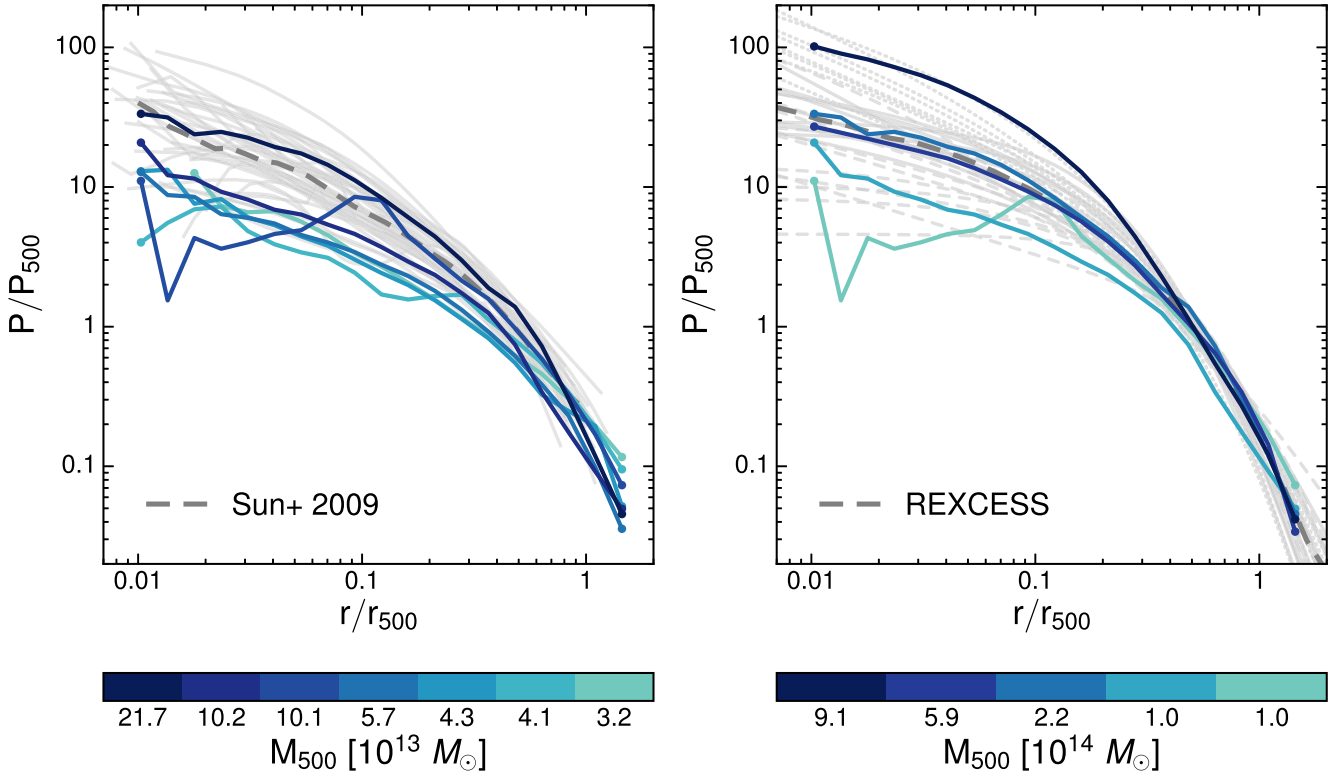
**Figure 14.** Dimensionless pressure profiles at $z = 0$ for the same simulated and observed samples as shown in Fig. 11. Profiles are scaled by the characteristic pressure $P_{500}$ as defined in the text. In the left-hand panel, grey lines show the pressure profiles of the Sun et al. (2009) groups derived from the density and temperature profiles shown in Figs 11 and 12. In the right-hand panel, grey lines show the best-fitting pressure profiles of the REXCESS clusters (Pratt et al. 2010). Line styles are the same as shown in Fig. 11.

and entropy profiles, there is a slight offset in normalization with respect to the median observed profile, possibly due to an X-ray mass bias. Inside $\sim 0.3\, r_{500}$, the most massive system remains in good agreement with the data; however, the central pressure in the less massive simulated groups is slightly underestimated. This is consistent with the density profiles shown in Fig. 11, which are slightly underestimated at $r \lesssim 0.3\, r_{500}$ compared to the median observed profile. As we discussed in Section 6.1, this suggests that AGN feedback may be ejecting too much gas from the central regions of galaxy groups, although selection effects in the Sun et al. (2009) sample may also play a role.

In the right-hand panel of Fig. 14, we find excellent agreement with the REXCESS pressure profiles across the full range of radii. In the outskirts of the FABLE clusters ($r \geq 0.5\, r_{500}$), the dimensionless pressure profiles coincide over a wide range of halo masses. The same is true for the observed clusters, which suggests that both the simulated and observational samples represent fairly self-similar populations. At small radii, there is a departure from the self-similar scaling due to the effects of non-gravitational processes. The dispersion increases towards the cluster centre in both the observed and simulated samples to a similar degree.

## 7 DISCUSSION

The FABLE simulations employ an updated version of the Illustris galaxy formation model. Specifically, we have updated the sub-grid models for feedback from stars and AGN in order to reproduce the $z = 0$ GSMF and the present-day gas mass fractions of massive haloes. The latter were not considered during the Illustris calibration

and were severely underestimated with respect to observations. By adopting a model that reproduces observed gas fractions, we have obtained significantly more realistic galaxy groups and clusters, while, at the same time, maintaining a good match to the observed GSMF in the field.

The FABLE model produces a very similar GSMF to Illustris (see Fig. 2), despite significantly changing the way in which AGN feedback regulates star formation in massive galaxies. In Illustris, quasar-mode feedback was continuous such that a relatively small amount of accreted feedback energy was injected into the surrounding gas at every time-step. This often resulted in too little thermal energy being input into too much gas, allowing the energy to be efficiently radiated away. The quasar-mode was therefore inefficient at suppressing star formation and strong radio-mode feedback was needed to suppress stellar mass build-up in massive haloes. This was achieved with a long duty cycle for the radio-mode such that feedback events were infrequent but highly energetic. The major side-effect of this model was that the radio-mode acted too violently on the gas, resulting in severely underestimated gas fractions in Illustris at $z = 0$ (see Fig. 5). In FABLE, we have introduced a modification to the quasar mode in the form of a duty cycle. Rather than a continuous injection of thermal energy as in Illustris, the available feedback energy is stored over a 25 Myr time period before being injected into the surrounding gas in a single event. This heats the gas to much higher temperatures, resulting in a longer cooling time and overall more efficient feedback. The updated quasar-mode feedback is more effective at suppressing the stellar mass build-up of massive galaxies, thereby allowing the radio mode to operate on a much shorter duty cycle. This gentler form of radio-mode feedback has a

smaller impact on the gas content of massive haloes and has enabled us to produce groups and clusters with realistic gas fractions (see Fig. 5). We point out that the agreement with observed gas fractions at cluster scales ($M_{500} \gtrsim 10^{14}\,\mathrm{M_\odot}$) was not guaranteed, since the feedback model was not calibrated at such scales. By applying our calibrated model to cluster-scale objects simulated using the zoom-in technique, we are able to compare the predictions of the FABLE model to a variety of observational constraints across a wide range of halo masses. We demonstrate very good agreement with observations for a number of group and cluster properties, including stellar mass fractions, X-ray luminosity–mass relations, integrated SZ flux, and radial profiles of the ICM.

Yet there remain some discrepancies with our model compared to observations. In particular, we find that the X-ray luminosity–temperature ($L_{500}$–$T$) relation lies on the upper end of the observed scatter (see Fig. 8). From the $L_{500}$–$T$ relation alone it is unclear whether the cause of this offset is dominated by overestimated X-ray luminosities or underestimated spectroscopic temperatures. We aim to gain some insight on the discrepancy by comparing to other observed scaling relations such as the halo mass–temperature ($M_{500}$–$T$) relation and the X-ray luminosity–halo mass ($L_{500}$–$M_{500}$) relation. However, the difference between relations based on X-ray hydrostatic mass estimates versus weak-lensing mass estimates means that the conclusion is dependent upon which is used for the comparison.

If we compare the $M_{500}$–$T_{500}$ relation of the FABLE simulations to observational data based on X-ray derived halo masses (the left-hand panel of Fig. 9), we would conclude that the average X-ray temperatures of our systems are systematically underestimated by ∼0.2 dex. This is large enough to explain the discrepancy in $L_{500}$–$T_{500}$ without affecting the $L_{500}$–$M_{500}$ relations, which are in excellent agreement with observed $L_{500}$–$M_{500}$ relations based on X-ray-derived halo masses (Figs 6 and 7).

On the other hand, we have very good agreement with the $M_{500}$–$T_{300\mathrm{kpc}}$ relation of Lieu et al. (2016), which uses halo masses measured via weak lensing (the right-hand panel of Fig. 9). This would suggest that our simulated systems possess realistic global temperatures. Similarly, our agreement with the weak-lensing-calibrated $Y$–$M_{500}$ relation (Fig. 10) implies that the mass-weighted temperatures of our simulated systems are realistic (the integrated tSZ flux being proportional to the total thermal energy content of the gas). Given that our spectroscopic temperature estimates are not systematically lower than the mass-weighted temperature (see Appendix B2), this provides further evidence that the discrepancy in $L_{500}$–$T_{500}$ relation is unlikely to be due to underestimated temperatures.

There is a significant offset in normalization between the weak-lensing $M_{500}$–$T$ relation of Lieu et al. (2016) and the $M_{500}$–$T$ relations based on X-ray masses, albeit with large scatter in the weak-lensing data. In fact, Lieu et al. (2016) perform a comparison of the normalization of different $M_{500}$–$T$ relations from the literature and find that relations based on weak-lensing masses favour ∼40 per cent higher normalizations than those based on X-ray hydrostatic masses. This implies a systematic difference between halo masses measured from weak lensing and masses measured from X-rays. Indeed, a number of observational studies have found that, within $r_{500}$, X-ray hydrostatic masses are biased low compared to weak-lensing masses by ∼25–30 per cent (e.g. Donahue et al. 2014; von der Linden et al. 2014b; Hoekstra et al. 2015; Simet et al. 2017). A slightly larger X-ray mass bias of ∼40 per cent, as seemingly preferred by the Lieu et al. (2016) sample, is large enough to reconcile the results of cluster abundance studies with cosmological

constraints from *Planck* measurements of the primary CMB (Planck Collaboration XXIV 2016).

Under the assumption that weak-lensing masses are less biased than X-ray hydrostatic masses, we would deduce that the spectroscopic temperatures of FABLE groups and clusters are realistic and that the discrepancy in $L_{500}$–$T_{500}$ is largely the result of overestimated X-ray luminosities. This is consistent with the $L_{500}$–$M_{500}$ relations shown in Figs 6 and 7, which suggest that the predicted X-ray luminosities may be overestimated as a function of halo mass compared to observations based on weak-lensing (rather than X-ray) mass measurements. Furthermore, if X-ray hydrostatic masses are biased low, then the observational constraints on stellar and gas mass fractions plotted in Figs 4 and 5 would be biased high. For example, Eckert et al. (2016) find that the weak-lensing-calibrated gas fraction of XXL-100-GC clusters is significantly lower than independent results based on X-ray hydrostatic masses. Since we have calibrated our feedback model to reproduce observed gas mass fractions assuming a negligible X-ray mass bias, this would imply that FABLE groups and clusters are too gas rich. The excess gas could then explain our overpredicted X-ray luminosities at fixed temperature.

On the other hand, a large X-ray hydrostatic mass bias ($\gtrsim 30$ per cent) implies a baryon depletion factor significantly exceeding that predicted by numerical simulations (e.g. Eckert et al. 2016). Moreover, a number of studies find results consistent with little to no bias (e.g. Gruen et al. 2014; Israel et al. 2014; Smith et al. 2015; Applegate et al. 2016; Maughan et al. 2016; Andreon et al. 2017). In a future paper, we will investigate the level of X-ray and hydrostatic mass bias in the FABLE simulations as a function of cosmic time (Henden et al.in preparation). With this knowledge in hand we hope to explain the origin of the discrepancies between the predicted and observed scaling relations.

Few cosmological hydrodynamical simulations manage to convincingly reproduce the observed X-ray scaling relations. The cosmo-OWLS and BAHAMAS projects obtain a good match to the observed X-ray luminosity–halo mass relation with relatively low-resolution simulations and when modelling a significant X-ray mass bias. At much higher resolution, the C-EAGLE clusters, which employ the EAGLE galaxy formation model, are slightly overluminous for a given halo mass due to the clusters being too gas rich. Similarly, the hydrodynamical cluster simulations presented in Truong et al. (2018) possess lower than observed temperatures and approximately 30 per cent higher X-ray luminosities than observed, although the latter discrepancy they suggest is at least partly due to sample selection. Mesh-based cosmological hydrodynamical simulations encounter similar issues. For example, the Rhapsody-G (Hahn et al. 2017) suite of cluster zoom-in simulations, which use an Eulerian adaptive mesh refinement (AMR) method, show X-ray luminosities as a function of halo mass consistently higher than observed (by ∼20 per cent at $M_{500} \approx 10^{15}\,\mathrm{M_\odot}$ and about a factor of 2 at $M_{500} \approx 10^{14}\,\mathrm{M_\odot}$). Interestingly, Hahn et al. (2017) find that the normalization of the X-ray luminosity–halo mass relation is insensitive to the AGN feedback parameters, including drastic changes to the length of the duty cycle, contrary to the results of SPH simulations with a similar AGN feedback model (e.g. Le Brun et al. 2014).

In our simulations, which are run with the AREPO moving-mesh code, we find that changes to the duty cycle and energetics of AGN feedback can have a large impact on the gas mass fractions of massive haloes (see Appendix A). We expect that the FABLE model could likely be adjusted to produce somewhat lower gas mass fractions and thus lower X-ray luminosities in massive haloes by lengthening the duty cycle of radio-mode feedback. This would

make individual events more energetic and therefore more effective at ejecting gas beyond the virial radius. In Appendix A, we show that increasing the burstiness of the radio mode in this way can significantly lower halo gas fractions without drastically altering the $z = 0$ GSMF, which is already in good agreement with observations in our fiducial model.

On the other hand, further tuning of our relatively simplistic model for thermal bubble feedback is unlikely to significantly improve the ICM profiles of our simulated clusters, which show indications of overheating and removal of too much gas in the central regions, while gas at $\gtrsim 0.5\,r_{500}$ is relatively unaffected (see e.g. entropy profiles in Section 6.3). A similar though more extreme predicament applies to the C-EAGLE clusters (Barnes et al. 2017c), which show lower than observed gas density in the core and entropy profiles with significantly larger cores and higher central entropies than observed. Barnes et al. (2017c) suggest that AGN feedback in C-EAGLE is too active at late times, increasing the central entropy of clusters and preventing the formation of cool-core systems with steep central density and entropy profiles. The FABLE clusters also do not contain an obvious strong cool-core system, although a larger sample will be needed to assess this issue in detail. Other numerical works such as Rasia et al. (2015) and Hahn et al. (2017) have reproduced the observed dichotomy between cool-core and non-cool-core clusters; however, this does not necessarily imply good agreement with observational constraints on the global properties of clusters, since both models tend to produce clusters with somewhat higher than observed X-ray luminosities (Hahn et al. 2017; Truong et al. 2018). IllustrisTNG also produce a fraction of cool-core clusters that is in agreement with observations between $0.25 < z < 1.0$; however, the cool-core fraction is underpredicted at $z < 0.25$, with more clusters showing close to isentropic cores at $z = 0$ than observed (Barnes et al. 2017a).

Given that essentially all numerical simulations are unable to convincingly reproduce the observed thermodynamic profiles of cluster core regions, it seems vital that simulation models for AGN feedback should continue to be improved, for example, by inflating bubbles self-consistently with AGN jet feedback (Bourne & Sijacki 2017; Weinberger et al. 2017b), while also incorporating additional physical processes that have previously been neglected, such as cosmic rays (Jacob & Pfrommer 2017; Pfrommer et al. 2017), outflows driven by radiation pressure from AGN (Costa et al. 2017, 2018; Ishibashi, Fabian & Maiolino 2018) and anisotropic thermal conduction (Kannan et al. 2016a,b). In addition, further improvement to the realism of simulated groups and clusters will rely on a better understanding of the issues of mass bias and selection effects so that reliable comparisons to unbiased observational data sets can be performed.

# 8 CONCLUSIONS

We have presented the new FABLE suite of cosmological hydrodynamical simulations, which is based on the framework of the Illustris project. The simulations consist of a $(40\,h^{-1}\,\mathrm{Mpc})^3$ cosmological volume and a number of zoom-in simulations of individual galaxy groups and clusters. We have employed the moving mesh code AREPO and an updated version of the Illustris galaxy formation model. We have adapted the sub-grid models for stellar and AGN feedback in order to reproduce galaxy groups and clusters with more realistic gas fractions compared to Illustris, whilst maintaining a similarly high level of agreement with the observed present-day GSMF. In this paper, we have presented various other comparisons with observations, including X-ray and SZ scaling relations and radial profiles

of the ICM over a wide range of halo masses. Our main conclusions are as follows:

(i) We obtain very good agreement with observed GSMFs. The high-mass end of the $z = 0$ mass function is similar to that of Illustris, despite significant changes to the way in which AGN feedback suppresses the build-up of stellar mass. While the FABLE model was calibrated to reproduce the $z \approx 0$ mass function, the fact that the agreement with observations continues to higher redshift is a success of the model.

(ii) The stellar mass fractions of FABLE galaxy groups and clusters are also an excellent match to low-redshift observations, including in massive clusters that were not present in the calibration volume.

(iii) The $z = 0$ halo gas mass fractions represent a major improvement over Illustris and are now in good agreement with observations. This can be attributed to much less energetic but more frequent thermal energy injections in the radio-mode of AGN feedback, which remove less gas from haloes compared to the Illustris model.

(iv) The predicted X-ray luminosity–total mass ($L_{500}$–$M_{500}$) relations are in excellent agreement with observed relations based on X-ray hydrostatic mass estimates but seem to overestimate the X-ray luminosity for a given halo mass when compared with weak-lensing mass estimates. The difference between observed relations is consistent with a significant X-ray mass bias. Similarly, a comparison of observed total mass–spectroscopic temperature ($M_{500}$–$T$) relations reveals a systematic difference between those based on X-ray hydrostatic masses and weak-lensing masses. The FABLE simulations are in good agreement with $M_{500}$–$T$ data based on weak-lensing masses but have significantly lower global temperatures/higher masses compared to relations using only X-ray data.

(v) The slope of the predicted X-ray luminosity–spectroscopic temperature ($L_{500}$–$T_{500}$) relation is in excellent agreement with observations. The normalization of the relation lies, however, on the upper end of the scatter in the data. The size of this offset is similar to the offset with weak-lensing-based $L_{500}$–$M_{500}$ relations and X-ray-only $M_{500}$–$T_{500}$ relations. This implies that the discrepancy in $L_{500}$–$T_{500}$ could be due to either overestimated X-ray luminosities or underestimated global temperatures. We lean towards the former explanation, as this is consistent with the general expectation that weak-lensing masses are less biased than X-ray hydrostatic masses. An improved understanding of mass bias will be important for making further progress here.

(vi) The simulations are also in excellent agreement with the mean SZ flux–total mass ($Y_{5r_{500}}$–$M_{500}$) relation derived from *Planck* observations of locally bright galaxies. This implies that the global temperatures of our simulated systems are not significantly underestimated, consistent with our match to the weak-lensing $M_{500}$–$T$ relation.

(vii) In general, the radial profiles of the ICM are a good match to observations outside $\sim 0.3\,r_{500}$, where the majority of the ICM is located. Density and pressure profiles of the ICM are in good agreement with observations of both group- and cluster-scale systems. The group-scale profiles have slightly lower-than-observed density/pressure within $\sim 0.3\,r_{500}$; however, this may be (partly) due to selection effects in the observed sample. The temperature and entropy profiles of $\lesssim 10^{14}\,\mathrm{M_\odot}$ haloes are also in good agreement with observations, while for more massive systems, the scatter in the simulated profiles somewhat exceeds that of the observed samples.

The FABLE simulations represent a major improvement over Illustris in the galaxy group and cluster regime. The baryonic content and global X-ray and tSZ properties of the ICM are a good match to

observations across a wide range of scales. The ICM is realistically distributed with residual deviations arising in the thermodynamic properties only towards the cluster centre. Our results are consistent with numerous other simulation studies and suggest that a subtle interplay between AGN feedback and a number of supplementary physical phenomena may be needed to explain the observational properties of galaxy clusters and groups in the core and to the outskirts. In future papers we aim to study the formation and evolution of cluster galaxies and make predictions for various quantities relevant to cluster cosmology, such as the scatter and redshift evolution of cluster scaling relations.

## REFERENCES

Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
Andrade-Santos F. et al., 2017, ApJ, 843, 76
Andreon S., Trinchieri G., Moretti A., Wang J., 2017, A&A, 606, A25
Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046
Applegate D. E. et al., 2014, MNRAS, 439, 48
Applegate D. E. et al., 2016, MNRAS, 457, 1522
Arnaud K., 1996, in Jacoby G., Barnes J., eds, ASP Conf. Ser. Vol. 101, Astronomical Data Analysis Software and Systems V. Astron. Soc. Pac., San Francisco, p. 17
Arnaud M., Pointecouteau E., Pratt G. W., 2007, A&A, 474, L37
Arnaud M., Pratt G. W., Piffaretti R., Böhringer H., Croston J. H., Pointecouteau E., 2010, A&A, 517, A92
Bahé Y. M. et al., 2017, MNRAS, 470, 4186
Baldry I. K. et al., 2012, MNRAS, 421, 621
Barnes D. J., Kay S. T., Henson M. A., McCarthy I. G., Schaye J., Jenkins A., 2017b, MNRAS, 465, 213
Barnes D. J. et al., 2017a, preprint (arXiv:1710.08420)
Barnes D. J. et al., 2017c, MNRAS, 471, 1088
Becker M. R., Kravtsov A. V., 2011, ApJ, 740, 25
Benson B. A. et al., 2014, Millimeter, Submillimeter, Far-Infrared Detect. Instrum. Astron. VII, 9153, 91531P
Bernardi M., Meert A., Sheth R. K., Fischer J. L., Huertas-Company M., Maraston C., Shankar F., Vikram V., 2017, MNRAS, 467, 2217
Bernardi M., Meert A., Sheth R. K., Vikram V., Huertas-Company M., Mei S., Shankar F., 2013, MNRAS, 436, 697

Bernardi M., Shankar F., Hyde J. B., Mei S., Marulli F., Sheth R. K., 2010, MNRAS, 404, 2087
Bhattacharya S., Heitmann K., White M., Lukić Z., Wagner C., Habib S., 2011, ApJ, 732, 122
Biffi V., Sembolini F., De Petris M., Valdarnini R., Yepes G., Gottlöber S., 2014, MNRAS, 439, 588
Birkinshaw M., 1999, Phys. Rep., 310, 97
Blanton M. R. et al., 2001, AJ, 121, 2358
Bondi H., Hoyle F., 1944, MNRAS, 104, 273
Booth C. M., Schaye J., 2009, MNRAS, 398, 53
Borgani S., Kravtsov A., 2011, Adv. Sci. Lett., 4, 204
Bourne M. A., Sijacki D., 2017, MNRAS, 472, 4707
Burenin R. A., Vikhlinin A., Hornstrup A., Ebeling H., Quintana H., Mescheryakov A., 2007, ApJS, 172, 561
Böhringer H. et al., 2007, A&A, 469, 363
Carlstrom J. E., Holder G. P., Reese E. D., 2002, ARA&A, 40, 643
Chabrier G., 2003, PASP, 115, 763
Chisari N. E. et al., 2018, preprint (arXiv:1801.08559)
Chiu I., Mohr J., Mcdonald M., Bocquet S., Ashby M. L., 2015, MNRAS, 455, 258
Cohn J. D., White M., 2008, MNRAS, 385, 2025
Corless V. L., King L. J., 2007, MNRAS, 380, 149
Costa T., Rosdahl J., Sijacki D., Haehnelt M. G., 2018, MNRAS, 479, 2079
Costa T., Rosdahl J., Sijacki D., Haehnelt M. G., 2018, MNRAS, 473, 4197
Courtin J., Rasera Y., Alimi J.-M., Corasaniti P.-S., Boucher V., Füzfa A., 2010, MNRAS, 410, no
Crocce M., Fosalba P., Castander F. J., Gaztañaga E., 2010, MNRAS, 403, 1353
Croston J. H. et al., 2008, A&A, 487, 431
Cui W., Borgani S., Murante G., 2014, MNRAS, 441, 1769
Cusworth S. J., Kay S. T., Battye R. A., Thomas P. A., 2014, MNRAS, 439, 2485
Dark Energy Survey Collaboration, 2005, preprint (arXiv:0510346)
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Dolag K., Borgani S., Murante G., Springel V., 2009, MNRAS, 399, 497
Dolag K., Komatsu E., Sunyaev R., 2016, MNRAS, 463, 1797
Donahue M. et al., 2014, ApJ, 794, 136
Dubois Y. et al., 2014, MNRAS, 444, 1453
D'Souza R., Vegetti S., Kauffmann G., 2015, MNRAS, 454, 4027
Eckert D. et al., 2016, A&A, 592, A12
Eckmiller H. J., Hudson D. S., Reiprich T. H., 2011, A&A, 535, A105
Erben T. et al., 2013, MNRAS, 433, 2545
Fabjan D., Borgani S., Rasia E., Bonafede A., Dolag K., Murante G., Tornatore L., 2011, MNRAS, 416, 801
Fabjan D., Borgani S., Tornatore L., Saro A., Murante G., Dolag K., 2010, MNRAS, 401, 1670
Foëx G., Böhringer H., Chon G., 2017, A&A, 606, A122
Gaspari M., Brighenti F., Temi P., Ettori S., 2014, ApJ, 783, L10
Gastaldello F., Buote D. A., Humphrey P. J., Zappacosta L., Bullock J. S., Brighenti F., Mathews W. G., 2007, ApJ, 669, 158
Genel S. et al., 2014, MNRAS, 445, 175
Giles P. A. et al., 2016, A&A, 592, A3
Giles P. A. et al., 2017, MNRAS, 465, 858
Giodini S. et al., 2009, ApJ, 703, 982
Gonzalez A. H., Sivanandam S., Zabludoff A. I., Zaritsky D., 2013, ApJ, 778, 14
Gonzalez A. H., Zaritsky D., Zabludoff A. I., 2007, ApJ, 666, 147
Graham A. W., Driver S. P., Petrosian V., Conselice C. J., Bershady M. A., Crawford S. M., Goto T., 2005, Astron. J., 130, 1535
Grand R. J. J. et al., 2017, MNRAS, 467, 179
Greco J. P., Colin Hill J., Spergel D. N., Battaglia N., 2015, ApJ, 808, 151
Gruen D. et al., 2014, MNRAS, 442, 1507
Guo Q. et al., 2011, MNRAS, 413, 101
Hahn O., Martizzi D., Wu H.-Y. Y., Evrard A. E., Teyssier R., Wechsler R. H., 2017, MNRAS, 470, 166
Henderson S. W. et al., 2016, J. Low Temp. Phys., 184, 772

Henson M. A., Barnes D. J., Kay S. T., McCarthy I. G., Schaye J., 2017, MNRAS, 465, 3361

Heymans C. et al., 2012, MNRAS, 427, 146

Hoekstra H., Bartelmann M., Dahle H., Israel H., Limousin M., Meneghetti M., 2013, Space Sci. Rev., 177, 75

Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., 2015, MNRAS, 449, 685

Hoyle F., Lyttleton R., 1939, Proc. Camb. Philol. Soc., 35, 405

Ilbert O. et al., 2013, A&A, 556, A55

Ishibashi W., Fabian A. C., Maiolino R., 2018, MNRAS, 476, 512

Israel H., Reiprich T. H., Erben T., Massey R. J., Sarazin C. L., Schneider P., Vikhlinin A., 2014, A&A, 564, A129

Jacob S., Pfrommer C., 2017, MNRAS, 467, 1478

Kannan R., Springel V., Pakmor R., Marinacci F., Vogelsberger M., 2016a, MNRAS, 458, 410

Kannan R., Vogelsberger M., Pfrommer C., Weinberger R., Springel V., Hernquist L., Puchwein E., Pakmor R., 2016b, ApJ, 837, 1

Katz N., Weinberg D. H., Hernquist L., 1996, ApJS, 105, 19

Kelly P. L. et al., 2014, MNRAS, 439, 28

Kettula K. et al., 2013, ApJ, 778, 74

Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M. S., 2015, MNRAS, 450, 1349

Khedekar S., Churazov E., Kravtsov A., Zhuravleva I., Lau E. T., Nagai D., Sunyaev R., 2013, MNRAS, 431, 954

Kravtsov A. V., Borgani S., 2012, ARA&A, 50, 353

Kravtsov A. V., Vikhlinin A. A., Meshcheryakov A. V., 2018, Astron. Lett., 44, 8

Kroupa P., 2001, MNRAS, 322, 231

Laureijs R. et al., 2011, preprint (arXiv:1110.3193)

Leauthaud A. et al., 2010, ApJ, 709, 97

Le Brun A. M., McCarthy I. G., Melin J. B., 2015, MNRAS, 451, 3868

Le Brun A. M. C., McCarthy I. G., Schaye J., Ponman T. J., 2014, MNRAS, 441, 1270

Le Brun A. M. C., McCarthy I. G., Schaye J., Ponman T. J., 2017, MNRAS, 466, 4442

Li C., White S. D. M., 2009, MNRAS, 398, 2177

Lieu M. et al., 2016, A&A, 592, A4

Lovisari L., Reiprich T. H., Schellenberger G., 2015, A&A, 573, A118

LSST Dark Energy Science Collaboration, 2012, preprint (arXiv:1211.0310)

Mahdavi A., Hoekstra H., Babul A., Bildfell C., Jeltema T., Henry J. P., 2013, ApJ, 767, 116

Mahdavi A., Hoekstra H., Babul A., Henry J. P., 2008, MNRAS, 384, 1567

Mandelbaum R., 2017, preprint (arXiv:1710.03235)

Mantz A. B. et al., 2016, MNRAS, 463, 3582

Marian L., Smith R. E., Bernstein G. M., 2010, ApJ, 709, 286

Marinacci F., Pakmor R., Springel V., 2014, MNRAS, 437, 1750

Mathiesen B. F., Evrard A. E., 2001, ApJ, 546, 100

Maughan B. J., Giles P. A., Randall S. W., Jones C., Forman W. R., 2012, MNRAS, 421, 1583

Maughan B. J., Giles P. A., Rines K. J., Diaferio A., Geller M. J., Van Der Pyl N., Bonamente M., 2016, MNRAS, 461, 4182

Maughan B. J., Jones C., Forman W., Van Speybroeck L., 2008, ApJS, 174, 117

Mazzotta P., Rasia E., Moscardini L., Tormen G., 2004, MNRAS, 354, 10

McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, MNRAS, 465, 2936

McCarthy I. G., Schaye J., Bower R. G., Ponman T. J., Booth C. M., Vecchia C. D., Springel V., 2011, MNRAS, 412, 1965

McCarthy I. G. et al., 2010, MNRAS, 406, 822

Meneghetti M., Rasia E., Merten J., Bellagamba F., Ettori S., Mazzotta P., Dolag K., Marri S., 2010, A&A, 514, A93

Merloni A. et al., 2012, preprint (arXiv:1209.3114)

Miralda-Escude J., Babul A., 1995, ApJ, 449, 18

Morishita T., Abramson L. E., Treu T., Schmidt K. B., Vulcani B., Wang X., 2017, ApJ, 846, 139

Murray S. G., Power C., Robotham A. S. G., 2013, MNRAS, 434, L61

Muzzin A. et al., 2013, ApJ, 777, 18

Nagai D., Vikhlinin A., Kravtsov A. V., 2007, ApJ, 655, 98

Nandra K. et al., 2013, preprint (arXiv:1306.2307)

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Osmond J. P. F., Ponman T. J., 2004, MNRAS, 350, 1511

Pacaud F. et al., 2016, A&A, 592, A2

Pfrommer C., Pakmor R., Schaal K., Simpson C. M., Springel V., 2017, MNRAS, 465, 4500

Pike S. R., Kay S. T., Newton R. D. A., Thomas P. A., Jenkins A., 2014, MNRAS, 445, 1774

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

Planck Collaboration XI, 2011, A&A, 536, A11

Planck Collaboration XI, 2013, A&A, 557, A52

Planck Collaboration XIII, 2016, A&A, 594, A13

Planck Collaboration XXIV, 2016, A&A, 594, A24

Planelles S., Borgani S., Fabjan D., Killedar M., Murante G., Granato G. L., Ragone-Figueroa C., Dolag K., 2014, MNRAS, 438, 195

Planelles S., Schleicher D. R. G., Bykov A. M., 2015, Space Sci. Rev., 188, 93

Power C., Navarro J. F., Jenkins A., Frenk C. S., White S. D. M., Springel V., Stadel J., Quinn T., 2003, MNRAS, 338, 14

Pratt G. W., Croston J. H., Arnaud M., Böhringer H., 2009, A&A, 498, 361

Pratt G. W. et al., 2010, A&A, 511, 85

Puchwein E., Sijacki D., Springel V., 2008, ApJ, 687, L53

Puchwein E., Springel V., Sijacki D., Dolag K., 2010, MNRAS, 406, 936

Rasia E., Mazzotta P., Borgani S., Moscardini L., Dolag K., Tormen G., Diaferio A., Murante G., 2005, ApJ, 618, L1

Rasia E. et al., 2012, New J. Phys., 14, 055018

Rasia E. et al., 2015, ApJ, 813, L17

Richard J. et al., 2010, MNRAS, 404, 325

Rines K. J., Geller M. J., Diaferio A., Hwang H. S., 2016, ApJ, 819, 63

Rossetti M., Gastaldello F., Eckert D., Della Torre M., Pantiri G., Cazzoletti P., Molendi S., 2017, MNRAS, 468, 1917

Rozo E. et al., 2009, ApJ, 699, 768

Rudd D. H., Zentner A. R., Kravtsov A. V., 2008, ApJ, 672, 19

Salpeter E., 1955, ApJ, 121, 161

Sanderson A. J. R., O'Sullivan E., Ponman T. J., Gonzalez A. H., Sivanandam S., Zabludoff A. I., Zaritsky D., 2013, MNRAS, 429, 3288

Schaye J. et al., 2015, MNRAS, 446, 521

Semboloni E., Hoekstra H., Schaye J., Van Daalen M. P., Mccarthy I. G., 2011, MNRAS, 417, 2020

Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, MNRAS, 380, 877

Simet M., Battaglia N., Mandelbaum R., Seljak U., 2017, MNRAS, 466, 3663

Smith G. P. et al., 2015, MNRAS, 456, L74

Smith R. K., Brickhouse N. S., Liedahl D. A., Raymond J. C., 2001, ApJ, 556

Springel V., 2010, MNRAS, 401, 791

Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776

Springel V., Hernquist L., 2003, MNRAS, 339, 289

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726

Sun M., 2012, New J. Phys., 14, 045004

Sun M., Voit G. M., Donahue M., Jones C., Forman W., Vikhlinin A., 2009, ApJ, 693, 1142

Takahashi T. et al., 2010, in Arnaud M., Murray S. S., Takahashi T., eds, Proc. SPIE Conf. Ser. Vol. 7732, Space Telescopes and Instrumentation 2010: Ultraviolet to Gamma Ray. SPIE, Bellingham, p. 77320Z

Takahashi T. et al., 2016, in den Herder J. W., Takahashi T., Bautz M., eds, Proc. SPIE Conf. Ser. Vol. 9905, Space Telescopes and Instrumentation 2016: Ultraviolet to Gamma Ray. SPIE, Bellingham, p. 99050U

Tinker J. L. et al., 2008, ApJ, 688, 709

Tomczak A. R. et al., 2014, ApJ, 783, 85

Torrey P., Vogelsberger M., Genel S., Sijacki D., Springel V., Hernquist L., 2014, MNRAS, 438, 1985

Truong N. et al., 2018, MNRAS, 474, 4089

van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 3649

Velliscig M., van Daalen M. P., Schaye J., McCarthy I. G., Cacciato M., Le Brun A. M. C., Vecchia C. D., 2014, MNRAS, 442, 2641

Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., 2006, ApJ, 640, 691

Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031

Vogelsberger M. et al., 2014, MNRAS, 444, 1518

Voit G., 2005, Rev. Mod. Phys., 77, 207

Voit G. M., Kay S. T., Bryan G. L., 2005, MNRAS, 364, 909

von der Linden A. et al., 2014a, MNRAS, 439, 2

von der Linden A. et al., 2014b, MNRAS, 443, 1973

Wang W., White S. D. M., Mandelbaum R., Henriques B., Anderson M. E., Han J., 2016, MNRAS, 456, 2301

Watson W. A., Iliev I. T., D'Aloisio A., Knebe A., Shapiro P. R., Yepes G., 2013, MNRAS, 433, 1230

Weinberger R., Ehlert K., Pfrommer C., Pakmor R., Springel V., 2017b, MNRAS, 470, 4530

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

Wiersma R. P. C., Schaye J., Smith B. D., 2009a, MNRAS, 393, 99

Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009b, MNRAS, 399, 574

Wu X., Fang L., 1997, ApJ, 483, 62

Zhang Y.-Y. et al., 2010, ApJ, 711, 1033

Zhang Y. Y., Andernach H., Caretta C. A., Reiprich T. H., Boehringer H., Puchwein E., Sijacki D., Girardi M., 2011, A&A, 526, A105

Zibetti S., White S. D. M., Schneider D. P., Brinkmann J., 2005, MNRAS, 358, 949

Zou S., Maughan B. J., Giles P. A., Vikhlinin A., Pacaud F., Burenin R., Hornstrup A., 2016, MNRAS, 463, 820

# APPENDIX A: TESTING AGN FEEDBACK MODELS

The FABLE simulations implement a series of physical models based on those of the Illustris galaxy formation simulation. With the Illustris model as our starting point, we have updated the sub-grid model for AGN feedback to reproduce the massive end of the present-day GSMF and the hot gas content of massive haloes with $M_{500} \approx 10^{13}$–$10^{14}\, \mathrm{M}_\odot$. Multiple variations of AGN feedback were tested, with our preferred model corresponding to the fiducial model presented in this paper.

Here, we present three additional variations of our AGN feedback model and their corresponding impact on the $z = 0$ GSMF and the stellar and gas fractions of massive haloes. Each model has been implemented in a periodic box of length 40 $h^{-1}$ (comoving) Mpc on a side with initial conditions as described in Section 2.1. These simulations differ only by the AGN feedback parameters listed in Table A1. The meaning of each parameter is explained in Section 2.4.2 and qualitative descriptions of the different models are given below. For reference, we also list the parameters used in Illustris. We point out that our choice for the radiative efficiency, $\epsilon_r$, is half that of Illustris. However, the coupling efficiency of the quasar mode, $\epsilon_f$, has been doubled such that the fraction of accreted rest mass energy used for quasar-mode feedback (the product of $\epsilon_r$ and $\epsilon_f$) is kept the same.

The 'weak radio' AGN feedback model is similar to the Illustris model but has a significantly shorter radio-mode duty cycle. Specifically, we have greatly reduced the fractional increase in black hole mass required to trigger a radio-mode event, $\delta_{\mathrm{BH}}$, such that bubbles are created more frequently but with less energy. In addition, although the radio-mode coupling efficiency, $\epsilon_m$, is approximately the same as Illustris; our lower radiative efficiency means that the overall efficiency with which AGN converts their accreted rest mass energy into bubbles (the product of $\epsilon_r$ and $\epsilon_m$) is approximately half that of Illustris. In the model 'stronger radio', we increase the length of the radio-mode duty cycle such that bubbles are an order of magnitude more energetic compared to the 'weak radio' model, although correspondingly less frequent. We also increase the radio-mode coupling efficiency by a factor of 2 relative to the 'weak radio' model so that the overall efficiency is approximately the same as Illustris. The third model, 'quasar duty cycle', implements a duty cycle for the quasar-mode of feedback as described in Section 2.4.2. In this model, AGN accumulates feedback energy over a period $\Delta t = 25$ Myr before releasing the energy in a single event. The fourth model is our fiducial model and combines the stronger radio-mode feedback of the second model with the quasar-mode duty cycle of the third model. In addition, we have lowered the accretion rate threshold for switching between quasar- and radio-mode feedback, $\chi_{\mathrm{radio}}$, such that black holes spend overall more time in the quasar mode. We point out that our fiducial model and the Illustris model convert a very similar fraction of accreted mass into feedback energy ($\epsilon_r \epsilon_f$ and $\epsilon_r \epsilon_m$ for the quasar mode and radio mode, respectively). The major difference compared to the Illustris model is in the duty cycles of the two modes.

In Fig. A1, we compare the $z = 0$ GSMFs of the four models. The 'stronger radio' model produces a slightly lower abundance of galaxies at the high-mass end compared with the 'weak radio' model. This implies that less frequent but more energetic bubble injections are slightly more efficient at suppressing star formation in massive haloes, partly by displacing gas from the dense central regions and partly by heating the surrounding gas to higher temperatures. The change is remarkably small given that there is an order-of-magnitude difference between the duty cycle parameter, $\delta_{\mathrm{BH}}$, of the two models. This explains the need for strong radio-mode feedback in the Illustris model to reduce stellar mass build-up to the degree seen in observations of the high-mass end of the GSMF.

**Table A1.** Parameter values for the AGN feedback models studied in this paper in comparison to Illustris. The parameter $\chi_{\mathrm{radio}}$ is the fraction of the Eddington accretion rate below which BHs operate in the radio mode and above which they operate in the quasar mode; $\epsilon_r$ is the radiative efficiency of BH accretion, $\epsilon_f$ is the thermal coupling efficiency of the quasar mode, $\Delta t$ is the accumulation time period between quasar-mode feedback events in Myr, $\epsilon_m$ is the coupling efficiency of the radio mode, and $\delta_{\mathrm{BH}}$ is the fractional increase in BH mass required to trigger a radio-mode feedback event.

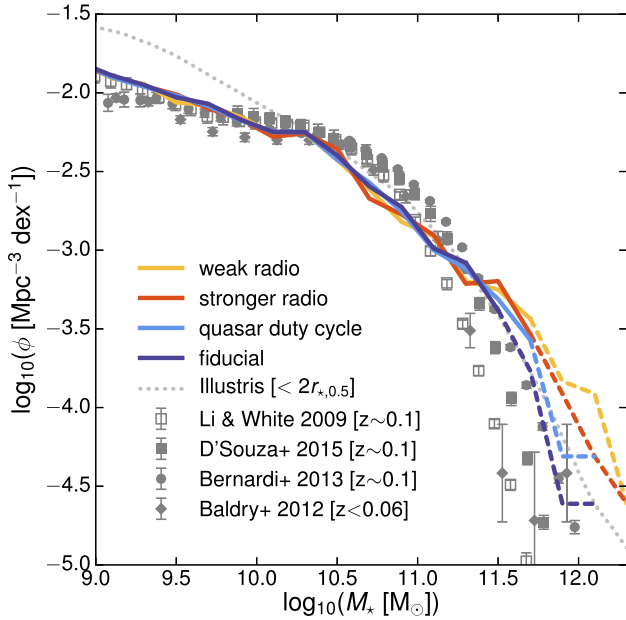| | $\chi_{\mathrm{radio}}$ | $\epsilon_r$ | $\epsilon_f$ | $\Delta t$ (Myr) | $\epsilon_m$ | $\delta_{\mathrm{BH}}$ |
|---|---|---|---|---|---|---|
| ILLUSTRIS | 0.05 | 0.2 | 0.05 | – | 0.35 | 0.15 |
| WEAK RADIO | 0.05 | 0.1 | 0.1 | – | 0.4 | 0.001 |
| STRONGER RADIO | 0.05 | 0.1 | 0.1 | – | 0.8 | 0.01 |
| QUASAR DUTY CYCLE | 0.05 | 0.1 | 0.1 | 25 | 0.4 | 0.001 |
| FIDUCIAL | 0.01 | 0.1 | 0.1 | 25 | 0.8 | 0.01 |

**Figure A1.** The GSMF at $z = 0$ for different AGN feedback models (lines) compared to observations (symbols with error bars). Lines become dashed at the high-mass end when there are fewer than 10 objects per 0.2 dex stellar mass bin. Here, the stellar mass of a galaxy is defined as the mass of stars bound to the sub-halo within twice the stellar half-mass radius. The grey dotted line shows the equivalent stellar mass function in Illustris. The observational data are as shown in Fig. 2.

The 'quasar duty cycle' model introduces a quasar-mode duty cycle to the 'weak radio' model, which leads to a significant reduction in the abundance of massive galaxies. This implies that periodic heating is much more effective at suppressing star formation than

continuous thermal feedback. Physically, this is because the gas is heated to higher temperatures, reducing cooling losses and slowing the rate at which the gas can condense to form stars. The GSMF of the 'quasar duty cycle' model is similar to Illustris at the high-mass end, despite using a far gentler form of radio-mode feedback. In Fig. A2, we will see that this change has significant consequences for the gas content of massive haloes. Our fiducial model combines a quasar-mode duty cycle with the 'stronger radio' model to reduce the abundance of massive galaxies even further, in good agreement with the data from Bernardi et al. (2013).

In the left-hand panel of Fig. A2, we plot the median stellar mass fraction as a function of halo mass for each of the models at $z = 0$. We note that the stellar fractions were not considered during the process of calibrating our fiducial AGN feedback model. The stellar fractions are largely consistent with the difference in the GSMFs between different models. That is, stronger radio-mode feedback reduces the total stellar mass of massive haloes but the introduction of a quasar-mode duty cycle has a significantly larger effect. The models converge at $\sim 10^{12}\,M_\odot$ since black holes in these haloes have generally not grown sufficiently in mass for AGN feedback to be effective. At $M_{500} \sim 7 \times 10^{13}\,M_\odot$, there is a slight overlap with the observational data and we find that the 'quasar duty cycle' and fiducial models, which implement a quasar-mode duty cycle, are in agreement with the observed stellar fractions, while the two models with continuous quasar-mode feedback slightly overpredict them. In Fig. 4, we demonstrate that the agreement with observations continues to much higher halo masses with our fiducial model, although we caution that this conclusion is dependent on the question of mass bias in X-ray hydrostatic mass estimates, as we discuss in Section 7.

In the right-hand panel of Fig. A2, we compare the median gas mass fraction of each model as a function of halo mass at $z = 0$. The 'stronger radio' model yields considerably lower gas fractions than the 'weak radio' model. This is owed to its more energetic bubble injections, which are able to eject gas from massive haloes more
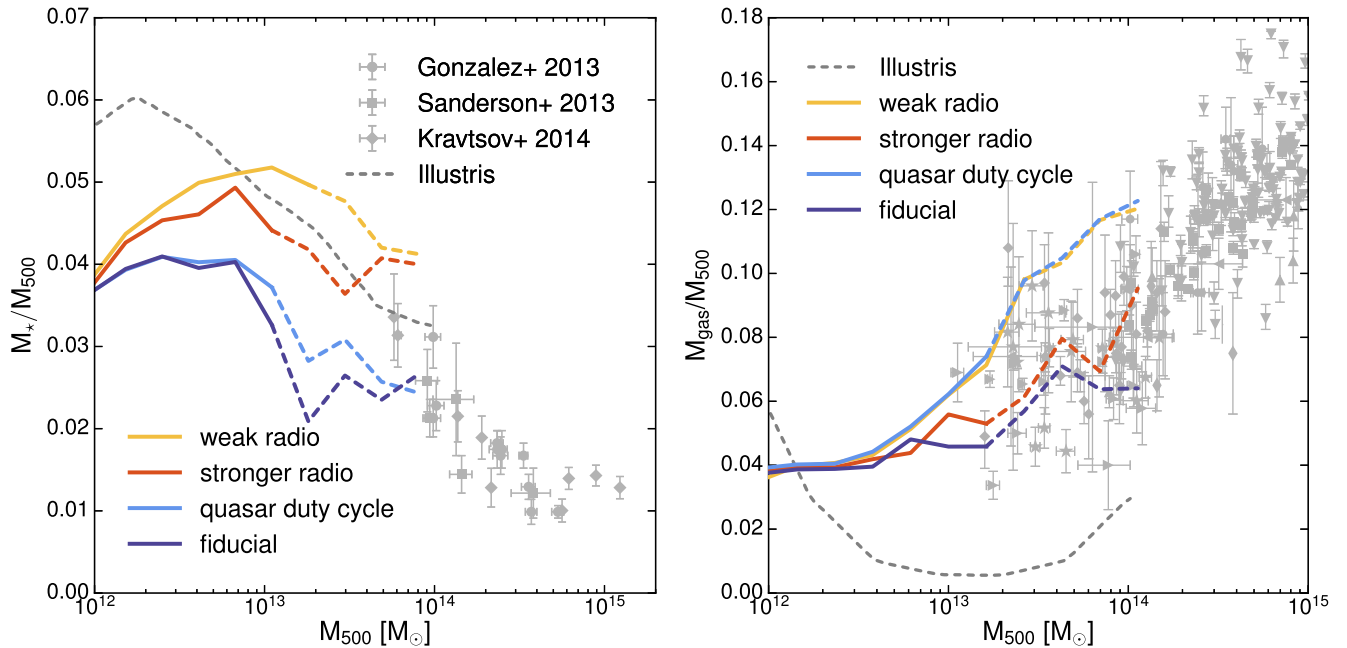


**Figure A2.** Stellar mass (left) and gas mass (right) fractions inside $r_{500}$ as a function of halo mass at $z = 0$ for different AGN feedback models. Lines show the mean relation in halo mass bins of width 0.2 dex. Lines become dashed when there are fewer than 10 haloes per bin. The grey dashed line shows the mean relation in Illustris. The total gas mass excludes cold and multiphase gas as described in Section 3.2. The observational data (symbols with error bars) are as shown in Figs 4 and 5.

efficiently. This also explains why all four models yield much larger gas fractions than the Illustris model, which injected approximately an order of magnitude more energy per bubble feedback event. The 'quasar duty cycle' model yields gas mass fractions in massive haloes almost identical to the 'weak radio' model, however, because the 'weak radio' run converts significantly more gas into stars, the total mass of baryons accumulated by $z = 0$ is lower in the model with a quasar-mode duty cycle. This is because a quasar-mode duty cycle slows the accumulation of gas on to the virialized region of the halo by heating the ICM to higher temperatures. This effect is consistent with SPH simulations such as Le Brun et al. (2014), which show that discontinuous thermal AGN feedback can have a strong impact on the gas content of massive haloes.

In summary, we find that periodic quasar-mode AGN feedback significantly reduces stellar mass build-up in massive haloes compared to the continuous case. This has allowed us to reproduce the observed abundance of galaxies at the high-mass end of the present-day GSMF without resorting to extremely strong radio-mode feedback, which was responsible for ejecting too much gas from massive haloes in the Illustris model. Using a far gentler form of radio-mode feedback in combination with a quasar-mode duty cycle allows us to reproduce the $z = 0$ GSMF and the local gas fractions of massive haloes simultaneously.

## APPENDIX B: MODELLING X-RAY LUMINOSITIES AND TEMPERATURES

### B1 Temperature–density cuts

In generating synthetic X-ray spectra for our simulated groups and clusters, we exclude gas with a temperature less than $3 \times 10^4$ K and gas above the density threshold required for star formation (see Section 3.2). These 'fiducial' temperature–density cuts ensure that

the derived X-ray luminosity and spectroscopic temperature are not biased by cold or star-forming gas, which should in reality produce negligible X-ray emission. In this section, we investigate whether the choice of cut can have a significant impact on the derived X-ray properties by comparing our fiducial cuts to a recalibrated version of the method used in Rasia et al. (2012).

Rasia et al. (2012, hereafter R12) identify a separated phase of cooling gas in their simulated clusters satisfying the condition

$$T < N \times \rho_{\mathrm{gas}}^{0.25}, \tag{B1}$$

where $T$ and $\rho_{\mathrm{gas}}$ are the temperature and density of the gas, respectively, and $N$ is a normalization factor. This relation follows from the polytropic equation for an ideal gas, $T \propto \rho_{\mathrm{gas}}^{\gamma-1}$, assuming a polytropic index of $\gamma = 1.25$. R12 consider a small range of cluster masses and assume a fixed normalization factor, $N$, equal to $3 \times 10^6$ with density in units of g cm$^{-3}$ and temperature in keV. Because our sample covers a much wider range of masses, we scale this normalization with the virial temperature of the halo, $T_{500}^{\mathrm{vir}} \propto M_{500}/r_{500}$, relative to the mean halo mass of the R12 sample, $\bar{M}_{500} = 5.61 \times 10^{14} \, h^{-1} \, \mathrm{M}_\odot$.

In Fig. B1, we show the temperature–density distribution for all gas within $r_{500}$ in the case of a high- and low-mass system: a galaxy cluster with $M_{500} = 5.9 \times 10^{14} \, \mathrm{M}_\odot$ and a galaxy group with $M_{500} = 4.6 \times 10^{13} \, \mathrm{M}_\odot$. Dashed lines indicate the fiducial cuts and solid lines the rescaled R12 cut. The fiducial and R12 cuts exclude 0.4 and 0.5 per cent of the total gas mass in the high-mass system, respectively, and 6.6 and 7.9 per cent in the lower mass system, respectively.

The R12 method was designed to excise compact sources with strong X-ray emission from synthetic X-ray images. In X-ray observations such structures would either be unresolved or would be excised during the analysis. The R12 procedure excludes the cold and dense gas clumps associated with these features without requir-
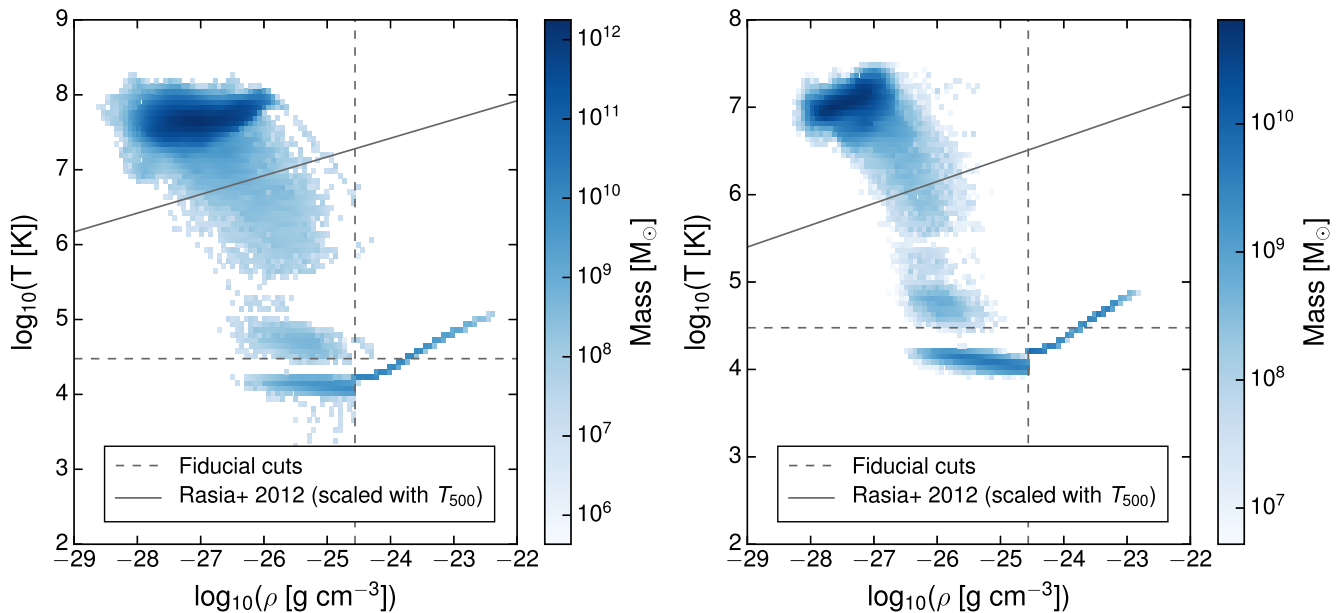


**Figure B1.** Temperature–density histogram for all gas within $r_{500}$ at $z = 0$ in a cluster with $M_{500} = 5.9 \times 10^{14} \, \mathrm{M}_\odot$ (left) and a group with $M_{500} = 4.1 \times 10^{13} \, \mathrm{M}_\odot$ (right). The colour scale quantifies the total mass of gas in each bin. The solid line shows the temperature–density cut corresponding to the Rasia et al. (2012) method described in the text. Gas below the solid line is excluded using this method and accounts for 0.5 and 7.9 per cent of the total gas mass within $r_{500}$ for the high- and low-mass system, respectively. The horizontal and vertical dashed lines show the fiducial temperature and density cuts used to exclude cold gas and gas followed only with the simple multiphase model for star formation, as discussed in Section 3.2. These cuts exclude 0.4 and 6.6 per cent of the total gas mass within $r_{500}$ for the high- and low-mass system, respectively.
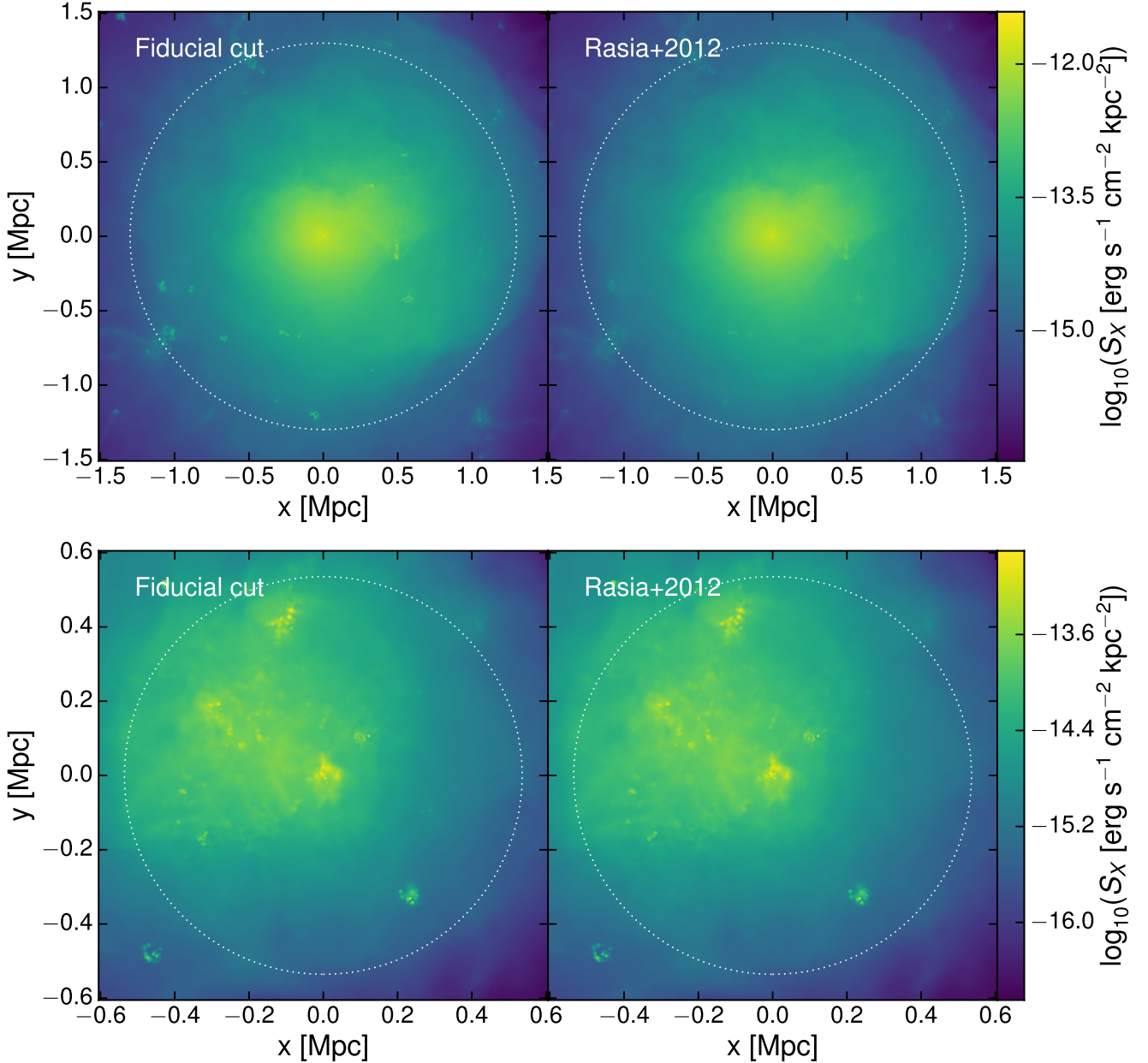
**Figure B2.** Projected X-ray surface brightness maps in the 0.5–10 keV energy band for the same systems shown in Fig. B1 for different temperature–density cuts. The upper panels show a cluster with $M_{500} = 5.9 \times 10^{14} \, M_\odot$ and the lower panels a group with $M_{500} = 4.1 \times 10^{13} \, M_\odot$. The white circle shows $r_{500}$ in each case. In the left-hand panels, the fiducial temperature and density cuts have been used to exclude cold and star-forming gas. In the right-hand panels, gas has been excluded according to the Rasia et al. (2012) method described in the text. Maps were derived from X-ray fluxes tabulated using APEC emission models generated for the same temperature bins and metallicity as used in the synthetic X-ray analysis described in Section 3.2. For a given gas cell, we interpolate an X-ray flux from the table and normalize it to the emission measure of the cell. We integrate the flux along the full length of the box, although for the more massive object, which was simulated using the zoom-in technique, gas outside the high-resolution region has been excluded from the integration.

ing time-consuming post-processing of the synthetic X-ray images. In Fig. B2, we assess the effect of the (rescaled) R12 cut on the X-ray surface brightness maps of the same cluster and group objects shown in Fig. B1. The left- and right-hand panels correspond to the fiducial and R12 temperature–density cuts, respectively. We calculate X-ray flux in the 0.5–10 keV band, which is a typical choice for *Chandra* data analyses and corresponds to the same energy range we use to fit our synthetic spectra (Section 3.2).

For the more massive object, the R12 cut removes most of the apparent sub-structure. These structures persist after excluding sub-

haloes from the analysis, which implies that they are not gravitationally bound to a dark matter sub-structure. Rather, they belong to a separated cooling phase that is at least partially excluded by the R12 method. For the lower mass system on the other hand, there is no discernible difference in the X-ray surface brightness distributions between the two cuts. This is in part because the 0.5–10 keV energy band is fairly insensitive to cool gas. From Fig. B1, we see that the gas excluded by either method has temperatures below 0.5 keV ($\approx 6 \times 10^6$ K). Such gas contributes relatively little to the X-ray surface brightness because much of the X-ray emission falls
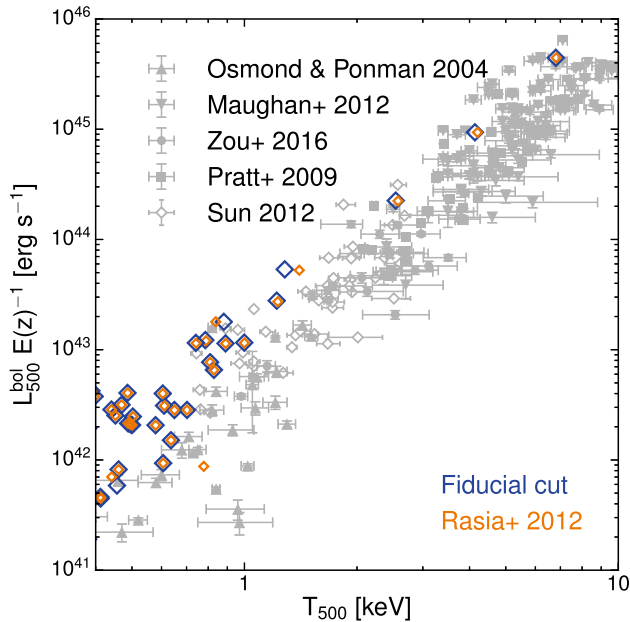
**Figure B3.** Bolometric X-ray luminosity as a function of spectroscopic temperature measured within $r_{500}$ at $z = 0$ for different density–temperature cuts. Blue diamonds show the luminosities and temperatures derived when the fiducial cuts are used to exclude cold and star-forming gas. Orange diamonds show the same systems when gas is excluded using the Rasia et al. (2012) method described in the text. Observational data from Osmond & Ponman (2004), Maughan et al. (2012), Zou et al. (2016), Pratt et al. (2009), and Sun (2012) are shown for comparison.



**Figure B4.** Mass-weighted versus X-ray spectroscopic temperature for FA-BLE haloes at $z = 0$. The solid line shows equality. The mass-weighted temperature is calculated within the spherical radius $r_{500}$ while the spectroscopic temperature is measured either within the same spherical aperture (3D; orange diamonds) or within a projected aperture of radius $r_{500}$ integrated through the length of the box (2D; blue diamonds). The 2D temperature was used for the X-ray scaling relations presented in this paper. The spectroscopic temperature is derived from a single-temperature fit to a synthetic X-ray spectrum convolved with the *Chandra* response function (see Section 3.2). Cold and multiphase gas has been excluded from both the mass-weighted and spectroscopic temperature measurements. The increase in spectroscopic temperature relative to mass-weighted at low temperatures is largely a result of *Chandra's* effective area, which drops rapidly below $\sim0.5$ keV.

outside the 0.5–10 keV energy band. We caution that this is somewhat dependent on our choice of scaling for the normalization, $N$, of the R12 cut shown in equation (B1). Using a higher normalization for the same object would exclude gas at higher temperatures, potentially removing some of the X-ray bright sub-structure.

Fig. B3 shows the bolometric X-ray luminosity as a function of X-ray spectroscopic temperature when using either the fiducial or R12 temperature–density cuts. We find that the derived X-ray luminosity and temperature is fairly insensitive to the choice of cut. At cluster scales ($T \gtrsim 3$ keV), we find little difference between the two methods. This implies that the X-ray bright structures, which are present in the fiducial case but are excluded with the R12 cut, have a minimal effect on the derived X-ray properties of clusters. In a handful of group-scale systems, the R12 cut produces a notable shift in the $L_{500}$–$T_{500}$ plane, typically towards higher temperatures. Unlike the group shown in the lower panel of Fig. B2, these few objects show relatively cold, dense clumps of gas with strong X-ray emission that are largely excluded by the R12 cut. The lower luminosities and temperatures of these groups compared to clusters means that such clumps can have a more significant impact on the derived X-ray properties. However, this is true of only a small fraction of objects, while for the majority of groups there is little difference between the fiducial and R12 cuts.

In summary, we find that our derived X-ray luminosities and temperatures are not particularly sensitive to the choice of temperature–density cut. Excluding gas using a rescaled R12 cut does not cause a significant systematic shift in the derived $L_{500}$–$T_{500}$ relation compared to our fiducial cuts. In addition, our results suggest that the X-ray properties of clusters are not significantly biased by the presence of X-ray bright sub-structure.
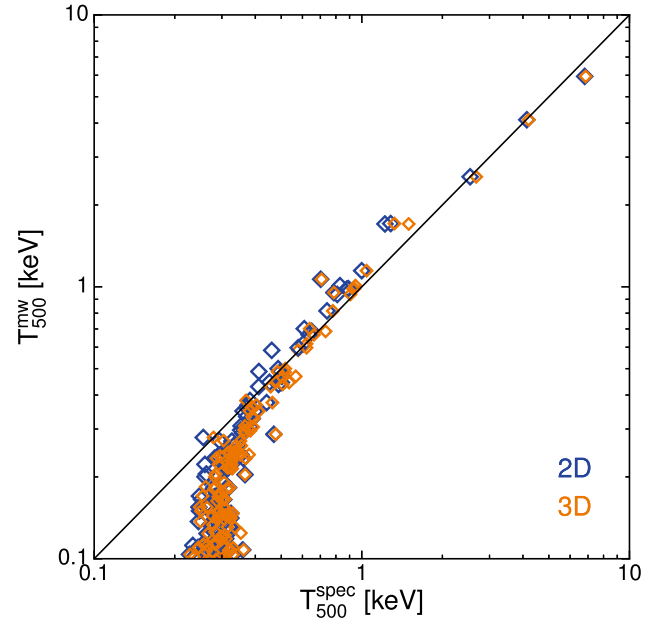
## B2 Spectroscopic and mass-weighted temperature

A number of studies have demonstrated that there can be significant biases in the temperature of hot gas derived from X-ray analyses compared to the mass-weighted temperature (e.g. Mazzotta et al. 2004; Rasia et al. 2005; Nagai et al. 2007). Given that the mass-weighted temperature is a direct measure of the thermal energy content of the gas, such a discrepancy could have important consequences for the scaling relations derived from X-ray observations.

In Fig. B4, we plot the mass-weighted temperature, $T_{500}^{\mathrm{mw}}$, against the X-ray spectroscopic temperature, $T_{500}^{\mathrm{spec}}$, as derived from our synthetic X-ray spectra (see Section 3.2) for all haloes with $T_{500}^{\mathrm{spec}} > 0.2$ keV. We calculate the mass-weighted temperature within the spherical radius $r_{500}$ and derive the spectroscopic temperature within either the same spherical aperture of radius $r_{500}$ (3D) or a projected aperture of radius $r_{500}$ integrated along the full length of the box (2D). The latter measure was used in the scaling relations presented in this paper as this is more akin to observations.

For systems with $T_{500}^{\mathrm{spec}} \gtrsim 0.5$ keV, the 2D spectroscopic temperature estimate is $\sim0.1$ dex lower than the mass-weighted temperature. This is in agreement with other studies using mock X-ray observations of simulated clusters (e.g. Mathiesen & Evrard 2001; Biffi et al. 2014). Unlike these studies, which find a systematic difference across the full $\sim1$–10 keV range, at $T_{500}^{\mathrm{spec}} > 2$ keV, we do not see a systematic offset between mass-weighted and spectroscopic temperatures. However, we only possess three systems at

these temperatures and the $T_{500}^{\mathrm{mw}}$–$T_{500}^{\mathrm{spec}}$ relation shows some object-to-object scatter. Part of the offset at $0.5\,\mathrm{keV} \lesssim T_{500}^{\mathrm{spec}} \lesssim 2\,\mathrm{keV}$ can be attributed to contamination of the cluster emission by cooler gas along the line of sight. This is demonstrated by the 3D spectroscopic temperature estimates, which are on average slightly higher than the 2D temperatures. In some cases, the cold gas clumps discussed in Appendix B1 bias the spectroscopic temperature slightly low compared to the mass-weighted temperature, although the difference is notable in only a handful of systems. The remaining difference is fairly small and may be explained by complex thermal structure in the ICM (e.g. multiple temperature components) that is inadequately modelled by a single-temperature fit to the integrated spectrum.

At temperatures below ~0.5 keV the spectroscopic temperature rises sharply compared to the mass-weighted temperature. This is largely due to the effective area of *Chandra*, which drops rapidly at energies below ~0.5 keV. It is thus difficult to reliably determine gas temperatures of systems in this regime from *Chandra* observations. This is not of particular concern for this study, as there are very few global temperature measurements at ~0.5 keV with which to compare.

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