

# The Three Hundred project: a large catalogue of theoretically modelled galaxy clusters for cosmological and astrophysical applications

Weiguang Cui,<sup>1★</sup> Alexander Knebe,<sup>1,2,3★</sup> Gustavo Yepes,<sup>1,2★</sup> Frazer Pearce,<sup>4</sup> Chris Power,<sup>3</sup> Romeel Dave,<sup>5</sup> Alexander Arth,<sup>6,7</sup> Stefano Borgani,<sup>8,9,10</sup> Klaus Dolag,<sup>6,11</sup> Pascal Elahi,<sup>3</sup> Robert Mostoghiu,<sup>1</sup> Giuseppe Murante,<sup>9</sup> Elena Rasia,<sup>9</sup> Doris Stoppacher,<sup>1,12</sup> Jesus Vega-Ferrero,<sup>13</sup> Yang Wang,<sup>14</sup> Xiaohu Yang,<sup>15,16</sup> Andrew Benson,<sup>17</sup> Sofía A. Cora,<sup>18,19</sup> Darren J. Croton,<sup>20</sup> Manodeep Sinha,<sup>20</sup> Adam R. H. Stevens,<sup>3</sup> Cristian A. Vega-Martínez,<sup>18</sup> Jake Arthur,<sup>4</sup> Anna S. Baldi,<sup>21,22</sup> Rodrigo Cañas,<sup>3</sup> Giammarco Cialone,<sup>21</sup> Daniel Cunnama,<sup>23,24</sup> Marco De Petris,<sup>21,25</sup> Giacomo Durando,<sup>26</sup> Stefano Ettori,<sup>27,28</sup> Stefan Gottlöber,<sup>29</sup> Sebastián E. Nuza,<sup>30,31</sup> Lyndsay J. Old,<sup>32</sup> Sergey Pilipenko,<sup>33</sup> Jenny G. Sorce<sup>34,29</sup> and Charlotte Welker<sup>3,35</sup>

*Affiliations are listed at the end of the paper*

Accepted 2018 August 1. Received 2018 July 27; in original form 2018 May 22

## ABSTRACT

We introduce the THE THREE HUNDRED project, an endeavour to model 324 large galaxy clusters with full-physics hydrodynamical re-simulations. Here we present the data set and study the differences to observations for fundamental galaxy cluster properties and scaling relations. We find that the modelled galaxy clusters are generally in reasonable agreement with observations with respect to baryonic fractions and gas scaling relations at redshift  $z = 0$ . However, there are still some (model-dependent) differences, such as central galaxies being too massive, and galaxy colours ( $g - r$ ) being bluer (about 0.2 dex lower at the peak position) than in observations. The agreement in gas scaling relations down to  $10^{13} h^{-1} M_{\odot}$  between the simulations indicates that particulars of the sub-grid modelling of the baryonic physics only has a weak influence on these relations. We also include – where appropriate – a comparison to three semi-analytical galaxy formation models as applied to the same underlying dark-matter-only simulation. All simulations and derived data products are publicly available.

**Key words:** galaxies: clusters: general – galaxies: clusters: intracluster medium – galaxies: general – galaxies: haloes.

## 1 INTRODUCTION

Galaxy clusters are the largest gravitationally bound objects in the Universe and as such they provide a host environment for testing both cosmology models and theories of galaxy evolution. Their formation depends both on the underlying cosmological framework and the details of the baryonic physics that is responsible for powerful feedback processes. Amongst others, these mechanisms regulate the observed properties of the intracluster medium (ICM), the size of the central brightest cluster galaxy (BCG), and the number and properties of the satellite galaxies orbiting within a common

dark-matter (DM) envelope. Clusters of galaxies can therefore be considered to be large cosmological laboratories that are useful for pinning down both cosmological parameters and empirical models of astrophysical processes acting across a range of coupled scales.

Concerted effort, from both observational and theoretical perspectives, has been devoted to improve our understanding of the formation and evolution of galaxy clusters. On the observational side, multiwavelength telescopes are designed to observe different properties of galaxy clusters: radio and far infrared data provide information on the cold gas; optical data focusses attention on the stellar properties and provides input to gravitational lensing analyses which target the DM component; millimetre and X-ray observations target the ICM. In parallel with these observational programmes, hydrodynamical simulations of the formation and evolution of galaxy clusters have been a very powerful tool to interpret

\* E-mail: cuiweiguang@gmail.com (WC); alexander.knebe@uam.es (AK); gustavo.yepes@uam.es (GY)

and guide observations for more than 20 yr (Evrard, Metzler & Navarro 1996; Bryan & Norman 1998). However, these extremely large objects with masses  $M \geq 10^{15} h^{-1} M_{\odot}$  are very rare and can only be found in large volumes  $V \gg (100 h^{-1} \text{Mpc})^3$ . But modelling such volumes with all the relevant DM and baryonic physics, while obtaining sufficient mass and spatial resolution at the same time, is challenging. Therefore, the most commonly used approach is to perform so-called ‘zoom’ simulations, i.e. selecting an object of interest from a parent DM simulation and only adding baryonic physics (at a much higher resolution) in a region about that object. This strategy has led to valuable results, but in order to be of statistical significance one would need to run hundreds – if not thousands – of such zoom simulations, which is what workers in the field are striving for at the moment.

Recent years have seen great advances in the direction of generating substantial samples of highly resolved galaxy cluster simulations that include all the relevant baryonic processes, e.g. the 500 ‘MUSIC’ clusters (Sembolini et al. 2013), the sample of 29 clusters of Planelles et al. (2013), the 10 ‘Rhapsody-G’ clusters (Wu et al. 2015), the 390 ‘MACSIS’ clusters (Barnes et al. 2017a), the 30 ‘Cluster-EAGLE’ (Barnes et al. 2017b), and 24 related ‘Hydrangea’ clusters (Bahé et al. 2017). The mass resolution of these zoom simulations varies from sample to sample covering the range of DM particle masses  $m_{\text{DM}} = 9.7 \times 10^6 h^{-1} M_{\odot}$  for ‘Hydrangea’ and ‘Cluster-EAGLE’ up to  $4.4 \times 10^9 h^{-1} M_{\odot}$  for the large ‘MACSIS’ sample. There are additionally cluster samples extracted from full box simulations, e.g. ‘cosmo-OWLS’ (Le Brun et al. 2014) and its follow-up ‘BAHAMAS’ (McCarthy et al. 2017) featuring hundreds of galaxy clusters, but the majority with masses lower than  $10^{14.5} h^{-1} M_{\odot}$  and at a mass resolution of  $m_{\text{DM}} \sim 4 \times 10^9 h^{-1} M_{\odot}$ .

In a series of precursor papers (i.e. the ‘nIFTy cluster comparison project’ introduced in Sembolini et al. 2016a,b), we investigated the differences in cluster properties arising from simulating *one* individual galaxy cluster with a variety of different numerical techniques including standard Smooth-Particle-Hydrodynamics (SPH), modern<sup>1</sup> SPH, and (moving) mesh codes. The results obtained there led us to the choice of using the modern SPH code GADGET-X which includes an improved SPH scheme and the implementation of black hole (BH) and active galactic nuclei (AGN) feedback compared to our fiducial GADGET-MUSIC code.

The primary goal of this paper is to introduce THE THREE HUNDRED project and its associated data set<sup>2</sup> that maximizes the ratio between number of objects and mass resolution: 324 regions of radius  $15 h^{-1} \text{Mpc}$  – having a cluster with mass  $M_{200} > 6.42 \times 10^{14} h^{-1} M_{\odot}$  at its centre – have been modelled with a combined mass resolution of  $m_{\text{DM}} + m_{\text{gas}} = 1.5 \times 10^9 h^{-1} M_{\odot}$ . This is, in fact, the same resolution as used for our previous ‘MUSIC’ clusters, but the difference here lies in an improved modelling of sub-grid physics and an application of a modern numerical SPH scheme. We detail the hydrosimulations, and the procedures for producing the cluster catalogue. We also present generic results, such as the dynamical state, baryon fraction, and optical/gas scaling relations. In addition, we add to the plots – where possible – the results from three semi-analytical galaxy formation models (SAMs) GALACTICUS, SAG, and SAGE, noting that they have been applied to the

**Table 1.** Parameters of THE THREE HUNDRED simulations.

	Value	Description
$\Omega_{\text{M}}$	0.307	Total matter density parameter
$\Omega_{\text{B}}$	0.048	Baryon density parameter
$\Omega_{\Lambda}$	0.693	Cosmological constant density parameter
$h$	0.678	Hubble constant in units of $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$
$\sigma_8$	0.823	Power spectrum normalization
$n_s$	0.96	Power index
$z_{\text{init}}$	120	Initial redshift
$\epsilon_{\text{phys}}$	6.5	Plummer equivalent softening in $h^{-1} \text{kpc}$
$L$	1	Size of the MDPL2 simulation box in $h^{-1} \text{Gpc}$
$R_{\text{resim}}$	15	Radius for each re-simulation region in $h^{-1} \text{Mpc}$
$M_{\text{DM}}$	12.7	DM particle mass in $10^8 h^{-1} M_{\odot}$
$M_{\text{gas}}$	2.36	Gas particle mass in $10^8 h^{-1} M_{\odot}$

same DM-only simulation that formed the basis for the selection of the clusters presented here (see Knebe et al. 2018, for the public release of the corresponding catalogues). Although this is not the first time that a joint analysis of hydrodynamical simulations with SAMs has been performed (e.g. Saro et al. 2010; Cui et al. 2011; Monaco et al. 2014; Guo et al. 2016), it is, to our knowledge, the first time such an approach has been applied to a large number of galaxy clusters. Detailed comparisons between the models and further investigation into different aspects of the cluster properties will be addressed in following companion papers.

The paper is structured as follows: we begin by describing the properties of the cluster sample in Section 2, which also includes a description of the hydrodynamical methods and of the SAMs. We briefly present our results for cluster bulk properties in Section 3, and for the relevant relations in different wavebands in Section 4. We conclude our results in Section 5.

## 2 THE GALAXY CLUSTER SAMPLE

The basis of our data set has been formed by extracting 324 spherical regions centred on each of the most massive clusters identified at  $z = 0$  by the ROCKSTAR<sup>3</sup> halo finder (Behroozi, Wechsler & Wu 2013) within the DM-only MDPL2, MultiDark simulation (Klypin et al. 2016).<sup>4</sup> The MDPL2 simulation utilizes the cosmological parameters shown in Table 1 which are those of the *Planck* mission (Planck Collaboration XIII 2016). The MDPL2 is a periodic cube of comoving length  $1 h^{-1} \text{Gpc}$  containing  $3840^3$  DM particles, each of mass  $1.5 \times 10^9 h^{-1} M_{\odot}$ .

### 2.1 The full-physics hydrodynamical simulations

The 324 clusters at the centre of each re-simulation region were selected initially as those with the largest halo virial mass<sup>5</sup> at  $z = 0$  with  $M_{\text{vir}} \gtrsim 8 \times 10^{14} h^{-1} M_{\odot}$ . The centres of their DM haloes serve as the centre of a spherical region with radius  $15 h^{-1} \text{Mpc}$ , for which initial conditions with multiple levels of mass refinement have been generated using the fully parallel GINNINGAGAP<sup>6</sup> code. DM particles within the highest resolution Lagrangian regions are split into DM and gas particles, according to the assumed cosmological baryon

<sup>1</sup>We define ‘modern’ as those SPH implementations that adopt an improved treatment of discontinuities.

<sup>2</sup>The data (ca. 50 TB of simulation data and 4TB of halo catalogues) are stored on a server to which access will be granted upon request to either AK or GY.

<sup>3</sup><https://bitbucket.org/gcfstanford/rockstar>

<sup>4</sup>The MultiDark simulations are publicly available at the <https://www.cosmosim.org> data base.

<sup>5</sup>The halo virial mass is defined as the mass enclosed inside an overdensity of  $\sim 98$  times the critical density of the universe (Bryan & Norman 1998).

<sup>6</sup><https://github.com/ginnungagapgroup/ginnungagap>

**Table 2.** Baryonic models for the two simulation codes.

Baryon physics	GADGET-MUSIC	GADGET-X
Gas treatment		
Homogeneous UV background	Haardt & Madau (2001)	Haardt & Madau (1996)
Cooling	Metal independent	Metal dependent (Wiersma, Schaye & Smith 2009)
Star formation and stellar feedback		
Stellar model	Springel & Hernquist (2003)	Tornatore et al. (2007)
Threshold for star forming	$0.1 \text{ cm}^{-3}$	$0.1 \text{ cm}^{-3}$
IMF	Salpeter (1955)	Chabrier (2003)
Kinetic feedback	Springel & Hernquist (2003)	Springel & Hernquist (2003)
Wind velocity	$400 \text{ km s}^{-1}$	$350 \text{ km s}^{-1}$
Thermal feedback	2-phase model (Yepes et al. 1997)	Only set the hot phase temperature
Gas mass-loss	via galactic winds	No
BH and AGN feedback		
BH seeding	No	$M_{\text{bh}} = 5 \times 10^6 h^{-1} M_{\odot}$ for $M_{\text{FoF}} \geq 2.5 \times 10^{11} h^{-1} M_{\odot}$
BH growth	No	Individual accretion of hot and cold gas
AGN feedback	No	Steinborn et al. (2015)

fraction listed in Table 1. Our mass resolution is a factor of three better than that used for the 390 ‘MACSIS’ clusters (Barnes et al. 2017a). We further highlight that our re-simulation regions have the same mass resolution as the original DM-only simulation upon which the SAMs are based. The DM particles outside this region are successively degraded in multiple layers (with a shell thickness of  $\sim 4 h^{-1}$  Mpc) with lower mass resolution particles (increased by eight times for each layer) that eventually provide the same tidal fields yet at a much lower computational costs than in the original simulation.<sup>7</sup> The size of the re-simulated region is much larger than the virial radius of the cluster it surrounds. As such, each region also contains many additional groups and filamentary structure that may or may not be physically associated with the cluster they surround.

The initial conditions – also publicly available – were run with the ‘modern’ SPH code GADGET-X and snapshots of the simulations stored for a set of pre-selected redshifts. A total of 128 different snapshots have been stored for each simulation from redshift  $z = 17$  to 0. We also ran the same simulations with our fiducial GADGET-MUSIC code (Sembolini et al. 2013). Both codes are based on the gravity solver of the GADGET3 Tree-PM code (an updated version of the GADGET2 code; Springel 2005). While both use smooth-particle hydrodynamics (SPH) to follow the evolution of the gas component, they apply different SPH techniques as well as rather distinct models for the sub-resolution physics. GADGET-X includes an improved SPH scheme (Beck et al. 2016) with artificial thermal diffusion, time-dependent artificial viscosity, high-order Wendland C4 interpolating kernel, and wake-up scheme. These improvements advance the SPH capability of following gas-dynamical instabilities and mixing processes by better describing the discontinuities and reducing the clumpiness instability of gas. They also minimize the viscosity away from shock regions and especially in rotating shears. GADGET-MUSIC uses the classic entropy-conserving SPH formulation with a 40 neighbour M3 interpolation kernel. The differences in baryon treatment have been summarized in Table 2. For more details and the implications of the code differences we refer the reader to our comparison papers (Sembolini et al. 2016a,b).

<sup>7</sup>The initial conditions for these clusters are publicly available in GADGET format and can be downloaded from <http://music.ft.uam.es> upon request. We have also produced higher resolution initial conditions corresponding to an equivalent resolution of  $7680^3$  particles, for a subsample of the cluster catalogue.

All data were then analysed with a standardized pipeline that includes the AHF<sup>8</sup> (Knollmann & Knebe 2009) halo finder which self-consistently includes both gas and stars in the halo finding process. For each halo, we compute the radius  $R_{200}$ , that is the radius  $r$  at which the density  $M(< r)/(4\pi r^3/3)$  drops below  $200\rho_{\text{crit}}$ .<sup>9</sup> Here  $\rho_{\text{crit}}$  is the critical density of the Universe at the respective redshift. Subhaloes are defined as haloes which lie within the  $R_{200}$  region of a more massive halo, the so-called host halo. As subhaloes are embedded within the density of their respective host halo, their own density profile usually shows a characteristic upturn at a radius  $R_t \lesssim R_{200}$ , where  $R_{200}$  would be their actual radius if they were found in isolation. We use this ‘truncation radius’  $R_t$  as the outer edge of the subhalo and hence subhalo properties (i.e. mass, density profile, velocity dispersion, rotation curve) are calculated using the gravitationally bound particles inside the truncation radius  $R_t$ . For a host halo that contains the mass of their subhaloes, we calculate properties using the radius  $R_{200}$ . Halo merger trees, that link objects between different redshifts, were constructed using MERGERTREE that forms part of the AHF package. We have calculated luminosities in different spectral bands from the stars within the haloes by applying the stellar population synthesis code STARDUST (see Devriendt, Guiderdoni & Sadat 1999, and references therein for more details). This code computes the spectral energy distribution from far-UV to radio, for an instantaneous starburst of a given mass, age, and metallicity. The stellar contribution to the total flux is calculated assuming a Kennicutt initial mass function (Kennicutt 1998).

The full data set consists of 324 re-simulated regions, which cover a much larger volume (out to  $15 h^{-1}$  Mpc in radius) than the central halo’s virial radius and hence our sample includes many other objects outside that sphere. These objects are composed of haloes, groups, and filaments, which allow us to investigate the preprocessing of the galaxy cluster as well as its large-scale environment. As some of the objects close to the boundary could be contaminated by low-resolution particles in the hydrodynamic simulations, we explicitly checked that all the objects included in the comprehensive catalogue do not contain any low-resolution particles. In what follows we refer to this data set, which consists of all the uncon-

<sup>8</sup><http://popia.ft.uam.es/AHF>

<sup>9</sup>Similarly, the subscript 500 used in this paper later are for haloes defined with enclosed overdensities of 500 times the critical density of the universe.

**Table 3.** Salient differences between the three SAMs. We only list here whether or not the model has been re-calibrated to the MDPL2 simulation, how it treats orphan galaxies (i.e. galaxies devoid of a DM halo), whether it features intracluster stars, and how luminosities are available. There are certainly many more differences in the exact implementation of the baryonic physics, but we refer the reader to the model presentation for those details.

SAM	Re-calibration	Orphan galaxies	Intracluster stars	Luminosities
GALACTICUS	No	Yes, but without positions/velocities	No	Yes
SAG	Yes	Yes, with full orbit integration	Yes	Yes
SAGE	Yes	No	Yes	Only for a sub-volume via TAO

taminated haloes from all the simulations as the ‘comprehensive’ sample (see the appendix for details).

## 2.2 The semi-analytical models

The aforementioned MDPL2 DM-only simulation has been populated with galaxies by three distinct SAMs, i.e. GALACTICUS (Benson 2012), SAG (Cora et al. 2018), and SAGE (Croton et al. 2016), and the public release of the resulting catalogues presented in Knebe et al. (2018). The same 324 regions (using the same radius cut) have also been extracted from the SAMs’ halo and galaxy catalogue that covers the entire  $1 h^{-1} \text{Gpc}^3$  volume of the parent MDPL2 simulation. This data set constitutes the counterpart sample of the hydrodynamical catalogue, which will be referred as the comprehensive sample as well. This allows for a direct comparison of the same galaxy clusters as modelled by our cosmological simulation codes detailed above. We briefly summarize the salient differences between these SAMs in Table 3, referring the reader to Knebe et al. (2018) for a more detailed presentation of the three models. Note that SAGE calculates luminosities in post-processing via the Theoretical Astrophysical Observatory (TAO,<sup>10</sup> Bernyk et al. 2016), which is currently only possible for a sub-volume of the full  $1 h^{-1} \text{Gpc}^3$ . Therefore, SAGE will not enter any luminosity-related plots.

## 3 CLUSTER BULK PROPERTIES

Before quantifying the differences in various cluster properties, we first illustrate in Fig. 1 the distributions of simulated galaxies and DM within a cluster ( $r \leq R_{200}$ ) from one of our re-simulated regions, from both hydrodynamical simulations (upper row) and from SAMs (lower row). Each galaxy is represented by a sphere with size proportional to stellar mass that includes halo stars for the two hydrodynamical simulation, but only uses the stellar mass of the central galaxy for the SAMs. Their colours are based on their Sloan digital sky survey (SDSS)  $r$ -,  $g$ -, and  $u$ - band luminosities. The background colour map indicates the DM density field, which is produced by the PY-SPHVIEWER code (Benitez-Llambay 2015). The two circles mark the radii  $R_{200}$  (outer) and  $R_{500}$  (inner).

It is apparent that the galaxies marked in the different panels are neither exactly in the same position nor do they have the same size for the hydrodynamical simulations. This is not surprising given that the dynamics within the virialized region is non-linear and so small differences in orbit become rapidly amplified. That said, the underlying DM density field is visually similar with a large infalling group to the south-east. Both  $R_{200}$  and  $R_{500}$  are recovered well by the re-simulation. The galaxies also differ due to the varying treatment

of baryonic processes, as seen in e.g. Sembolini et al. (2016a,b), Elahi et al. (2016), Cui et al. (2016b), and Arthur et al. (2017). Note that the galaxy positions are identical for the two SAMs as they reflect the positions of the DM haloes in the underlying DM-only simulation which are the same. The apparent larger sizes for the hydrodynamical galaxies can be related back to the inclusion of halo stars. In agreement with previous studies (e.g. Ragone-Figueroa et al. 2013; Cui et al. 2014a, 2016b), the galaxy stellar masses are significantly larger for GADGET-MUSIC, which does not include a model for AGN feedback.

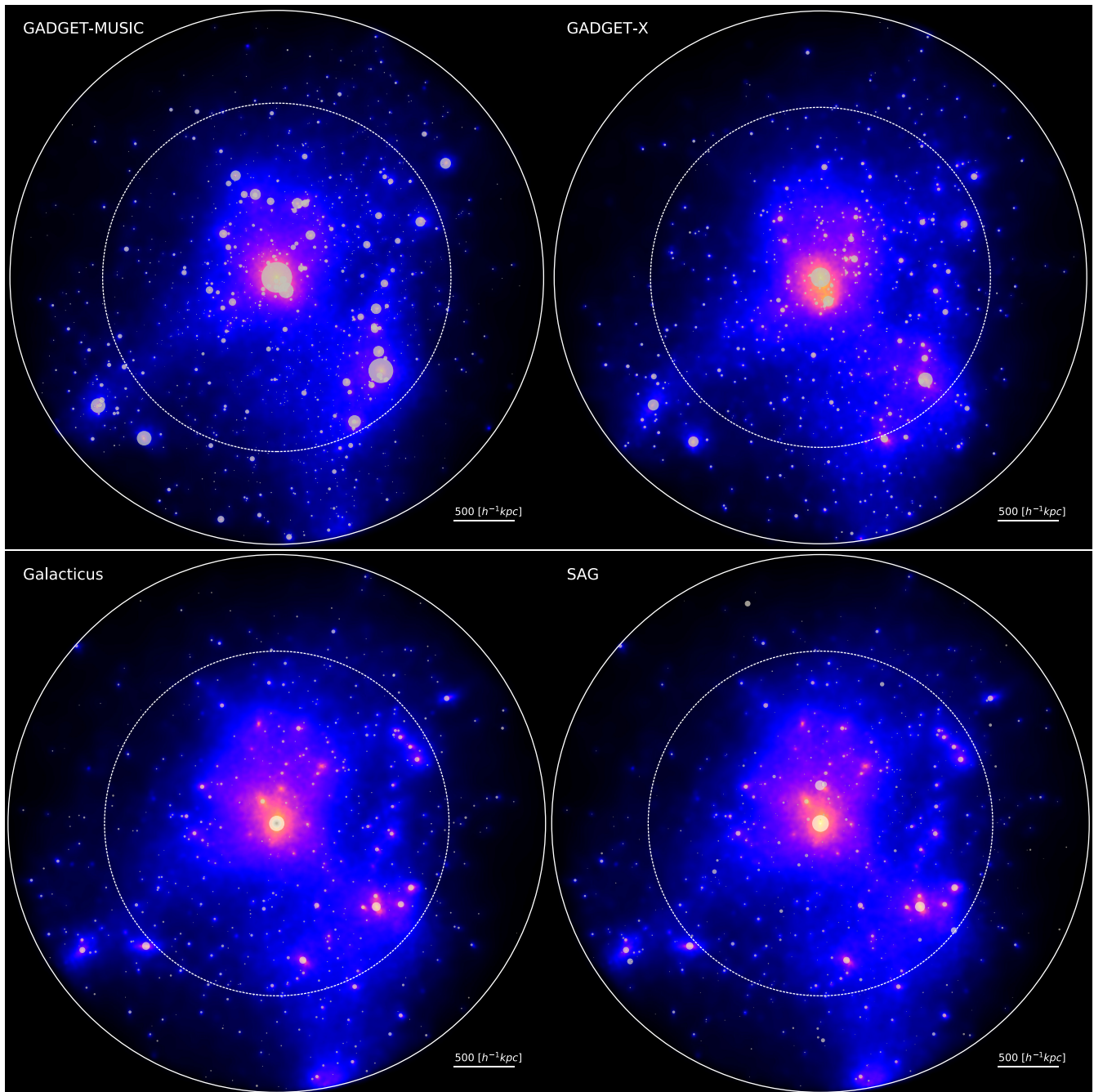
## 3.1 Halo properties

In this section, we focus on the results from the hydrosimulations, noting that the properties of the haloes of the SAM galaxies are identical to the MDPL2 halo properties presented elsewhere (Klypin et al. 2016; Knebe et al. 2018)

### 3.1.1 Baryon effects on halo mass

In order to compare individual clusters between the original MDPL2 simulation and the 324 re-simulated regions the haloes need to be matched. There is generally a direct 1-to-1 alignment between the largest object within the original simulation and the re-simulated region, as illustrated in Fig. 1. For the analysis presented here both the original MDPL2 region and the re-simulated region have been (re)processed using AHF. This ensures exact consistency between the halo finder definitions, i.e. it avoids effects introduced by using results from different halo finders (Knebe et al. 2011, 2013). Further, AHF can extract haloes self-consistently from simulations including gas and stars as well as DM. We use the halo centre position as the primary criteria for matching the clusters and select the one with the nearest mass when there are multiple matches. As previously mentioned the exact halo positions will have moved slightly from those in the original DM-only simulation but these changes are generally small (at the level of a few per cent of the virial radius in most cases, Cui et al. 2016b). Occasionally the differences are larger, typically due to the presence of an ongoing merger. It has been shown that halo finders struggle to uniquely track the main halo through a merger and rather treat the two participating objects as a host-subhalo system (Behroozi et al. 2015). Furthermore, the cluster centre can flip between different density peaks (subhaloes) due to baryonic processes (Cui et al. 2016a). That said, in our worst-case scenario, we have two matched haloes with  $\sim 40$  per cent mass difference caused by a massive merging subhalo. In general cases, these different kinds of mismatching only happen for the dynamically unrelaxed clusters, not for the relaxed ones.

<sup>10</sup><http://tao.asvo.org.au>

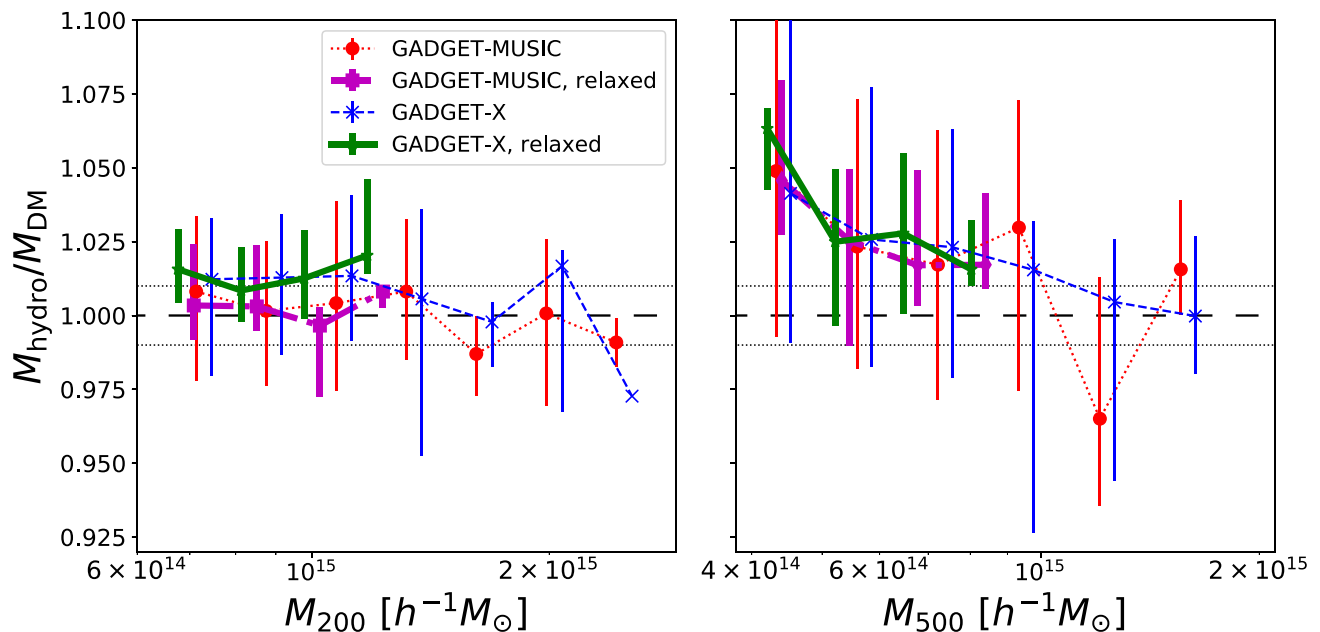


**Figure 1.** The distribution of galaxies within  $R_{200}$  of the most massive cluster within re-simulation region 1. The upper row shows the results from GADGET-MUSIC (left-hand side) and GADGET-X (right-hand side). The lower row shows the results from the SAMs GALACTICUS (left-hand side) and SAG (right-hand side). The projected DM density is shown in the background with a blue-red colour map. Galaxy colour is taken from their SDSS  $r$ ,  $g$ , and  $u$  band magnitude and the symbol size is proportional to stellar mass. The two circles mark the radii  $R_{200}$  (outer circle) and  $R_{500}$  (inner circle).

Accurate estimates of cluster masses are very important for constraining cosmological parameters and cosmological models (e.g. Bocquet et al. 2016; Sartoris et al. 2016). Therefore, we present here a quantitative comparison of the halo masses as found in the hydrodynamical simulations with their respective counterparts from the DM-only MDPL2 simulation (see Cui & Zhang 2017, for a review of the baryon effect). Fig. 2 shows the mass ratio of clusters in GADGET-MUSIC (red circle and lines) and GADGET-X (blue star and

lines) to their MDPL2 counterparts;  $M_{200}$  is shown in the left-hand side panel and  $M_{500}$  in the right-hand side panel.<sup>11</sup> In order to reduce any issues due to mismatching, we select a sample of dynamically

<sup>11</sup>The  $M_{500}$  sample was constructed by using AHF to find the largest halo contained within each of the 324 clusters of the mass-complete sample (and matching these as before).



**Figure 2.** The mass ratio between matched clusters at  $z = 0$  identified in the hydrodynamical simulations ( $M_{\text{hydro}}$ ) and in the corresponding cosmological DM-only run MDPL2 ( $M_{\text{DM}}$ ) for  $M_{200}$  (left-hand side panel) and  $M_{500}$  (right-hand side panel) as a function of  $M_{\text{DM}}$ . The complete sample used here is in thin lines, while the dynamically relaxed subsample is in thick lines. The median value for each mass bin is shown via the symbols (red dots for GADGET-MUSIC and blue stellar symbols for GADGET-X) with error bars indicating the 16th and 84th percentiles. The black horizontal long-dashed and dotted lines indicate equivalent mass and 1 per cent variation, respectively.

relaxed clusters (see below for details) from the complete sample and repeat the comparison. The mass ratio for  $M_{200}$  from both hydrodynamical simulations is very close to unity (with the median difference lying basically within 1 per cent), with a scatter less than  $\sim 5$  per cent ( $\sim 2.5$  per cent for the relaxed subsample). At the low-mass end, GADGET-X (for both samples) tends to have about 1 per cent higher mass than its MDPL2 counterpart. However, the  $M_{500}$  mass in both sets of hydrodynamical simulations tends to be several (up to 6) per cent higher than its DM-only counterpart below  $\sim 9 \times 10^{14} h^{-1} M_{\odot}$ . Above this halo mass the ratio drops to around 1 again. It is worth noting that for  $M_{500}$  there is a larger scatter of  $\sim 8$  per cent for the complete sample and  $\sim 4$  per cent for the relaxed subsample. We ascribe this larger mass change for  $M_{500}$  to baryonic processes that have a larger effect closer to the cluster’s centre and for the less massive haloes. The two simulation codes show similar results for  $M \gtrsim 10^{15} h^{-1} M_{\odot}$  at both overdensities, which means that the baryon physics has little influence on both  $M_{500}$  and  $M_{200}$  at this cluster mass range. For the  $M_{200}$  mass changes, this is in agreement with previous similar comparisons (e.g. Cui et al. 2012; Cui, Borgani & Murante 2014b; Cui et al. 2016b). For  $M_{500}$ , Cui et al. (2014b) reported a slight mass decrease when an AGN feedback is included and a slight mass increase without the AGN feedback. At this halo mass range,  $M_{500} > 10^{14.5} h^{-1} M_{\odot}$ , the difference between GADGET-X and Cui et al. (2014b) could be caused by either a sample effect (Cui et al. 2014b studied very few clusters) or due to the details of the baryonic model implemented in the simulation. We will explore this in detail in a follow-up paper (Cui et al. in preparation) which will also focus on cluster mass estimates based upon different observational methods applied to our simulation data.

### 3.1.2 Dynamical relaxation

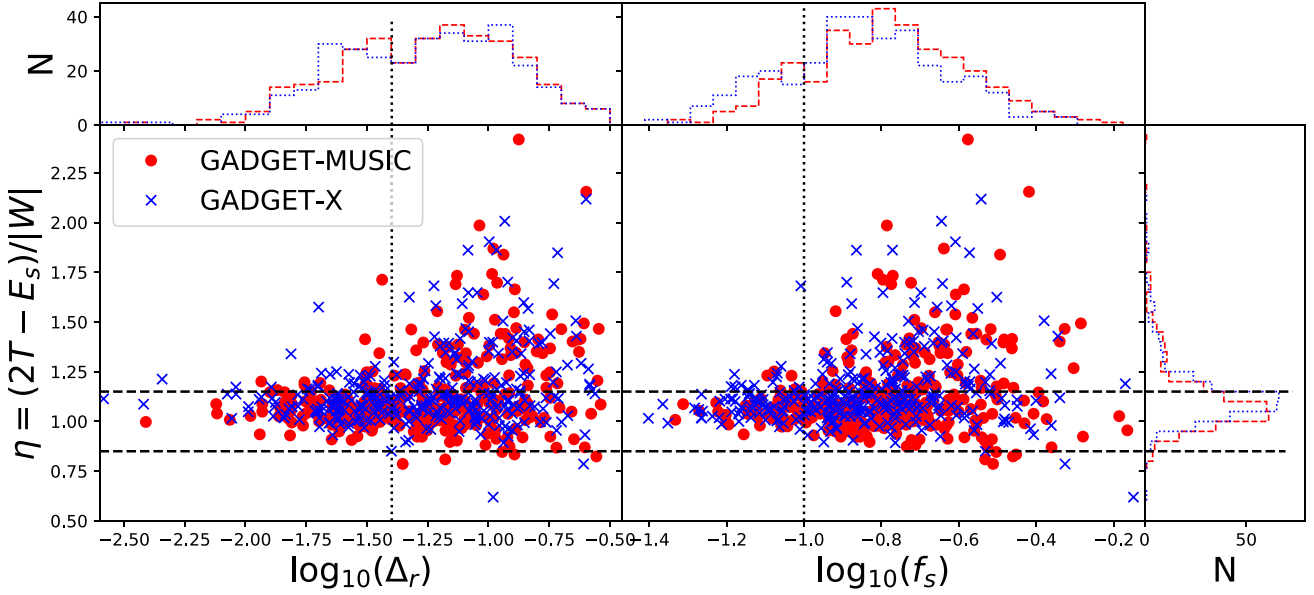
To determine the dynamical state of the cluster sample we study three indicators, following Cui et al. (2017), specifically:

- (i) the virial ratio  $\eta = (2T - E_s)/|W|$ , where  $T$  is the total kinetic energy,  $E_s$  is the energy from surface pressure, and  $W$  is the total potential energy,
- (ii) the centre-of-mass offset  $\Delta_r = |R_{\text{cm}} - R_c|/R_{200}$ , where  $R_{\text{cm}}$  is the centre-of-mass within a cluster radius of  $R_{200}$ ,  $R_c$  is the centre of the cluster corresponding to the maximum density peak of the halo. Using the position of the minimum of the gravitational potential would give a similar result as investigated by Cui et al. (2016a).
- (iii) the fraction of mass in subhaloes  $f_s = \sum M_{\text{sub}}/M_{200}$  where  $M_{\text{sub}}$  is the mass of each subhalo.

We adopt the following criteria to select dynamically relaxed clusters:  $0.85 < \eta < 1.15$ ,  $\Delta_r < 0.04$  and  $f_s < 0.1$ , which need to be satisfied at the same time (see, for instance, Neto et al. 2007; Knebe et al. 2008; Power, Knebe & Knollmann 2012). Note that we use here a slightly larger limit for  $f_s$  than in Cui et al. (2017). This is because (1)  $R_{200}$  is used instead of the virial radius,<sup>12</sup> and (2) this threshold for  $f_s$  gives a relaxation fraction ( $\sim 20$  per cent for both hydrodynamical simulations) comparable to observations (e.g. Mantz et al. 2015; Biffi et al. 2016).

In Fig. 3, we show the relations between these three parameters for the mass-complete sample:  $\Delta_r$  versus  $\eta$  in the left-hand

<sup>12</sup>Note that for the given cosmology  $R_{200} < R_{\text{vir}}$  and hence the  $M_{200}$  masses of the host haloes considered here will be about 25 per cent smaller than  $M_{\text{vir}}$ .



**Figure 3.** For the mass-complete sample, the left-hand side panel shows the relation between the virial ratio ( $\eta$ ) and the centre-of-mass offset ( $\Delta_r$ ). The right-hand side panel shows the relation between  $\eta$  and the subhalo mass fraction ( $f_s$ ). The top and right-hand subpanels show their corresponding histograms. Red filled circles (red dashed line for the histogram) show the clusters from the GADGET-MUSIC run, while the blue crosses (blue dotted line for the histogram) show the GADGET-X results. The two horizontal dashed lines show the selection limits for the  $\eta$  parameter, while the vertical dotted lines show the selection limits for  $\Delta_r$  and  $f_s$  (see text).

**Table 4.** The fraction of relaxed clusters. The first column shows the mass range. The second to fourth columns show the relaxation fractions from all three methods combined,  $\Delta_r$  plus  $f_s$ , and only  $f_s$  criterion, respectively. Each cell shows two values, of which the first one is the relaxation fraction for GADGET-MUSIC and the second value is for GADGET-X. Clusters with  $M_{200} < 6.42 \times 10^{14} h^{-1} M_\odot$  (mass bins above the dashed line) are taken from the comprehensive sample.

$M_{200} [10^{14} h^{-1} M_\odot]$	$\eta, \Delta_r \& f_s$	$\Delta_r \& f_s$	$f_s$
0.10–0.50	0.44 / 0.36	0.56 / 0.48	0.70 / 0.65
0.50–1.00	0.36 / 0.34	0.45 / 0.46	0.56 / 0.57
1.00–6.42	0.27 / 0.29	0.30 / 0.35	0.43 / 0.48
>6.42	0.15 / 0.17	0.16 / 0.21	0.17 / 0.23

side panel and  $f_s$  versus  $\eta$  in the right-hand side panel. The two hydro-runs show a similar distribution of relaxed clusters (shown for convenience at the top and to the right-hand side of the figure panels), in agreement with Cui et al. (2017). The histogram peak of the  $\eta$  parameter from GADGET-X has a slightly higher value than the peak from GADGET-MUSIC. This could be due to the AGN feedback, which releases additional energy into the kinetic component.

A quantitative analysis of the relaxation fraction within our comprehensive halo catalogue, for different mass bins and with different combinations of relaxation parameters is given in Table 4. The fraction of relaxed clusters shows a clear decreasing trend as halo mass increases. This is simply because the more massive the object is, the less likely it is to have reached dynamical relaxation by redshift  $z = 0$ . This can be traced back to the relation between formation time and halo mass (see Fig. 2 in Power et al. 2012, for instance). There is a very little change in relaxation fraction for the complete sample when different criteria are applied. There is a noticeable difference in the relaxed cluster fraction for the smallest mass bin, with the fraction for GADGET-X being significantly lower ( $\sim 8$  per cent) than that for GADGET-MUSIC when all three criteria are applied. This is due

to the AGN feedback in GADGET-X efficiently ceasing star formation in small objects and creating gas turbulence. The relaxation fractions for the mass-complete sample from both GADGET-MUSIC and GADGET-X show an obvious decrease. Contrary to the smallest mass bin, the relaxation fraction from GADGET-MUSIC seems lower than from GADGET-X. This overturn is simply because the mass fraction of substructures in GADGET-MUSIC is higher than GADGET-X, which dominates the relaxation fraction. In an upcoming paper we will provide a more detailed investigation of the evolution of the cluster dynamical state and the impact of input physics on various observational classification methods (Old et al. in preparation).

### 3.1.3 Concentration–Mass ( $c - M$ ) relation

Knowledge of the halo concentration,  $c$ , and mass,  $M$ , would specify the full evolution of a halo in the spherical collapse model (Bullock et al. 2001). The relation  $c - M$  between these two fundamental properties, alongside its standard deviation, are related to the variance in the assembly histories of DM haloes (e.g. Zhao et al. 2003a,b). Furthermore, the normalization and evolution of this relation also depend on the cosmological model (e.g. Dolag et al. 2004; Carlesi et al. 2012). However, there exists some tension between the observationally estimated relation and the theoretical prediction. This could result from not comparing like-with-like when contrasting baryonic simulations and observational results with carefully imposed selection criteria (see Rasia et al. 2013; Biviano et al. 2017, for example). Here, we only use our relaxed galaxy clusters from the mass-complete sample to investigate and compare this relation with the observational results.

The halo density profiles can be analytically described by an NFW profile (Navarro, Frenk & White 1997),

$$\frac{\rho(r)}{\rho_{\text{crit}}} = \frac{\delta_c}{(r/r_s)(1+r/r_s)^2}, \quad (1)$$

which is characterized by the two parameters,  $r_s$  and  $\delta_c$ . The concentration  $c_{200}$  is then given by  $R_{200}/r_s$ . We fit our simulated cluster density profiles, defined by equally spaced log-bins, to this functional form with both parameters free, but exclude the very central region in this process. Due to the presence of the BCG, the mass profile in the centre is much steeper than the total mass profile (Schaller et al. 2015b). As the edge of the BCG is not clearly defined, we adopt two different inner ‘exclusion’ radii during the fitting:  $0.05 R_{200}$ , as suggested by, for example, Schaller et al. (2015b), Cui et al. (2016b), and  $\sim 34 h^{-1}$  kpc following Biviano et al. (2017). We have verified that the NFW profile provides a good fit regardless of the adopted inner radii ( $34 h^{-1}$  kpc or  $0.05 R_{200}$ ). In both cases the difference between the fit and the original density profile is within 20 per cent at all radii.

In the left-hand side panel of Fig. 4 we show the  $c - M$  relation for our relaxed galaxy clusters and compare the relation with observational results coming from both X-ray and optical data obtained with different techniques (please refer to the figure caption and legend, respectively). For each of the two hydrodynamical simulation codes, we show results stemming from either truncation approach: circles for using the range  $[0.05 R_{200} - R_{200}]$  and stars for a fixed inner radius of  $34 h^{-1}$  kpc. We fit our  $c - M$  relation using the following analytical function:

$$\log_{10} c_{200} = \alpha - \beta \log_{10}(M_{200}/M_{\odot}). \quad (2)$$

The fitting parameters  $\alpha$  and  $\beta$  are listed in Table 5.

It is evident that the  $c - M$  relation from our hydrodynamical clusters is closer to the observational results from Merten et al. (2015), Okabe & Smith (2016), and Biviano et al. (2017) than those from Mantz et al. (2016) and Groener et al. (2016). The  $c - M$  relation from the GADGET-MUSIC run is slightly higher than from the GADGET-X run and it is in better agreement with observational results which have lower concentrations. It is obvious that the concentrations with a  $34 h^{-1}$  kpc inner cut-off are systematically higher than the ones with a  $0.05 R_{200}$  cut-off (see also Rasia et al. 2013, for similar results with different inner radii). Our fitted  $c - M$  relation from the GADGET-X clusters is much flatter than Schaller et al. (2015a), simply because their fit covers a much larger mass range, which is dominated by the lower mass objects. Furthermore, GADGET-X shows an increasing slope with  $\beta = -0.01$  when a fixed inner radius of  $34 h^{-1}$  kpc is taken. This can be understood because  $34 h^{-1}$  kpc corresponds to a smaller fraction of  $R_{200}$  for a massive cluster than for a less massive halo. Therefore it is not surprising to see a relatively high concentration for the most massive haloes when a fixed physical cut-off radius is applied.

In the right-hand side panel of Fig. 4, we investigate the baryon effects on the  $c - M$  relation by showing the relative change in concentration from DM-only simulated clusters to their equivalent in the two hydro-runs. The change on  $c - M$  relation due to baryons varies from  $\sim 25$  per cent (for both radii) for GADGET-X to about 1.5–2 times ( $0.05 R_{200} - 34 h^{-1}$  kpc) for GADGET-MUSIC. However, this ratio is much lower for the highest mass bin for GADGET-X with both inner radii (also for GADGET-MUSIC with the inner radius of  $34 h^{-1}$  kpc). The influence of baryons on the concentration is a little higher than in Rasia et al. (2013), which may be the result of both the different radius range used for profile fitting and differences in the baryonic model employed.

### 3.2 Baryon fractions

The formation of a galaxy cluster depends not only on gravity acting on cosmic scales but also on subresolution phenomena such as star

formation and various feedback mechanisms returning energy back to the intracluster gas. It is a process that involves interplay between dark and baryonic matter. One of the most important quantities to quantify the relation between DM and baryons is the baryonic mass fraction. It has therefore been intensively studied: on the theoretical side, mostly by means of hydrodynamical simulations (e.g. Planelles et al. 2013; Sembolini et al. 2013; Wu et al. 2015; Barnes et al. 2017b); on the observational side via multiwavelength observations (e.g. Laganá et al. 2013; Eckert et al. 2016; Chiu et al. 2017).

In Fig. 5, we show the gas and stellar mass fractions for the comprehensive sample from hydrodynamical simulations within  $R_{500}$ . The gas fraction for GADGET-X is larger than for GADGET-MUSIC at the massive end, and drops more quickly towards lower mass haloes. The gas fraction from GADGET-X shows a better agreement with the data of Gonzalez et al. (2013) at the massive end; both simulations are in line with the results from Zhang et al. (2011) due to its large scatter. The offset between the two hydro-runs is much larger (a factor of 2–3) for the stellar fraction. Again, GADGET-X shows a better agreement with the observational data points at the massive end. However, it has a flatter slope than the observational results, which is close to the GADGET-MUSIC result at  $M_{500} \lesssim 10^{13.5} h^{-1} M_{\odot}$ . This is possibly caused by the strong AGN feedback in GADGET-X. Essentially both hydrodynamic models have a stellar fraction versus mass slope that is inconsistent with the observational data.

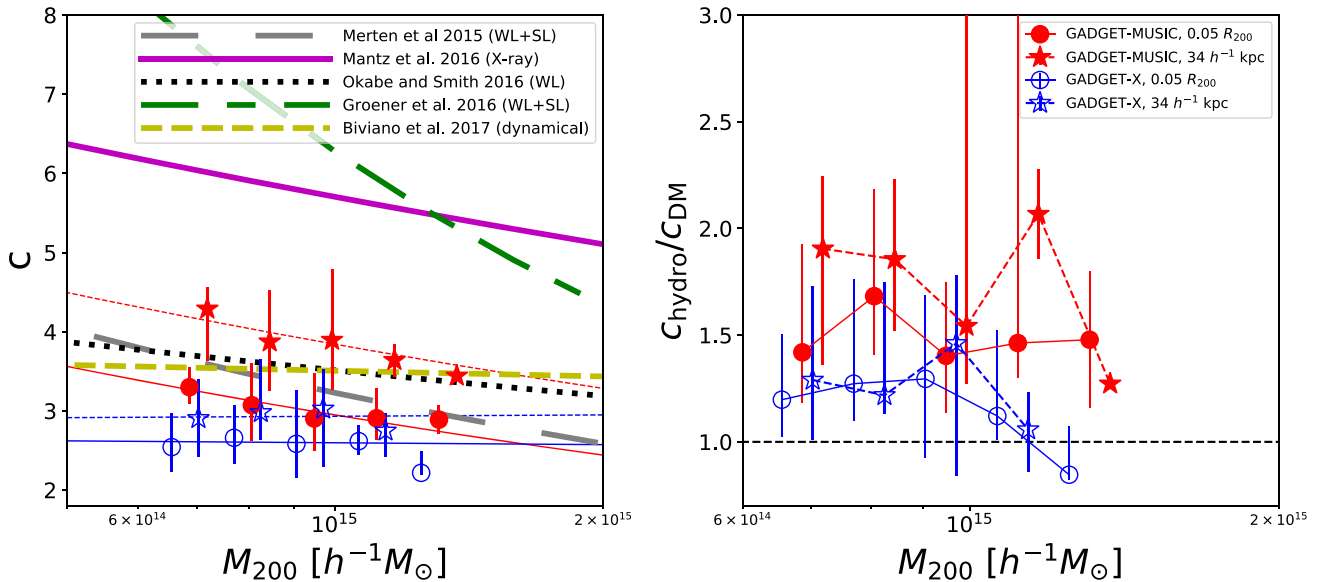
Previous comparisons of the stellar and gas mass fractions from full-physics hydrodynamical simulations with observations have shown that models without AGN feedback consistently have too low a gas fraction and too high a stellar fraction due to the overcooling problem (e.g. Planelles et al. 2013). This is also seen in Fig. 5 comparing the GADGET-MUSIC and the GADGET-X runs. Although GADGET-X tends to have a better agreement with the observational results, the AGN feedback implementation featured by this code is still not perfect: the most massive clusters at  $M_{500} \gtrsim 10^{15} h^{-1} M_{\odot}$  still have a stellar fraction that is a little too high; while intermediate and low-mass haloes ( $M_{500} \lesssim 10^{14} h^{-1} M_{\odot}$ ) have stellar fractions that are too low. Nevertheless, we note that the stellar mass fraction estimated from observations is not without issues: there is relative uncertainty about the contribution of the intracluster light (e.g. Zibetti et al. 2005; Gonzalez, Zaritsky & Zabludoff 2007; Puchwein et al. 2010; Cui et al. 2014a), which is included in Gonzalez et al. (2013) and Kravtsov et al. (2018), but not in Zhang et al. (2011). Another problem is the influence of the different initial mass functions adopted in observations to derive stellar mass from luminosities (see e.g. Chiu et al. 2017, for detailed discussions).

The difference in the stellar mass fractions shows the importance of the detailed prescription for baryon processes. Therefore, we are working on a follow-up paper (Rasia et al., in preparation) to investigate in detail the connection between the encapsulated physics and the resultant baryonic fractions, examining the difference between relaxed and unrelaxed clusters, between cool core and non-cool core clusters, as well as the redshift evolution of these fractions.

## 4 STELLAR AND GAS RELATIONS OF CLUSTERS

Scaling relations between the total cluster mass and observational quantities are derived in several multiwavelength studies. Commonly used observational probes include stellar luminosity, X-ray temperature, or the Comptonization parameter (e.g. Reiprich & Böhringer 2002; Lin, Mohr & Stanford 2004; Andersson et al. 2011), which normally show a self-similar relation to cluster mass.





**Figure 4.** Left-hand side panel: The concentration–halo mass relation for the relaxed galaxy clusters from the two hydrodynamical simulation runs compared with various observational results. As indicated in the legend, thick lines with different styles show the best-fitting results from recent observational data obtained with different methods (Merten et al. 2015; Groener, Goldberg & Sereno 2016; Mantz, Allen & Morris 2016; Okabe & Smith 2016; Biviano et al. 2017). Symbols show the median values with the 16th–84th percentile error bars from the hydrosimulations: circles and stars (red filled symbols for GADGET-MUSIC and blue open symbols for GADGET-X) for the concentration derived by fitting the density profile up to two inner radii ( $34 h^{-1}$  kpc and  $0.05 R_{200}$ , see text for details). The red (blue), thin solid and dashed lines are the best-fitting result to the concentration–mass relation of GADGET-MUSIC (GADGET-X) clusters. In the right-hand side panel of this figure, we represent the ratio of the concentration between the hydrodynamical simulation clusters and their match in the original MDPL2 DM-only simulation. Again, the symbols show the median values with the 16th–84th percentile error bars.

**Table 5.** The fitting parameters for the concentration–mass relation with fitting function:  $\log_{10} c_{200} = \alpha - \beta \log_{10} M_{200}/M_{\odot}$ . The first row shows the results with the inner radius set to  $0.05 R_{200}$ , while the second row is for a  $34 h^{-1}$  kpc inner radius. Each cell shows two values, of which the first one is for the fitting parameter  $\alpha$  and the second value is  $\beta$ .

Inner radius	GADGET-MUSIC $\alpha / \beta$	GADGET-X $\alpha / \beta$
$0.05 R_{200}$	4.60 / 0.27	0.62 / 0.013
$34 h^{-1}$ kpc	4.02 / 0.23	0.34 / -0.01

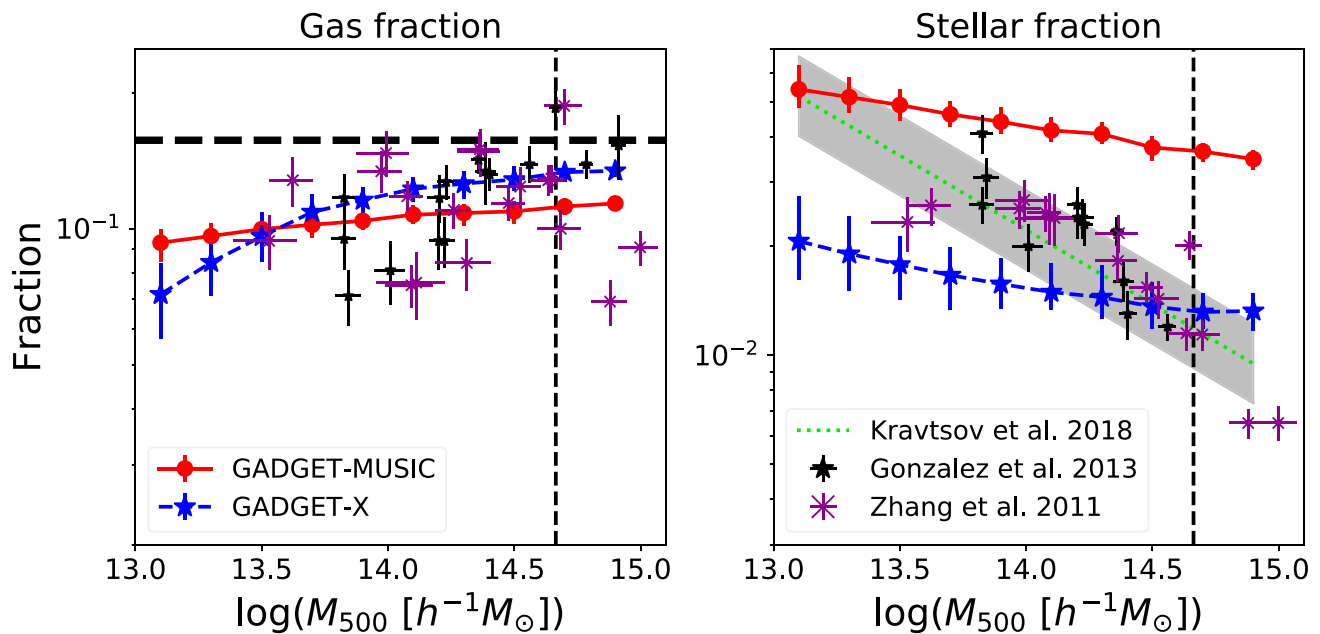
They are very powerful tools to derive total cluster masses from different observations. Before this can happen, they need to be accurately calibrated and their dispersions properly estimated. It is worth noting that the scaling relations derived from observations could be biased by sample selection which should have no influence on our mass-complete sample. In this section, we investigate the scaling relations found in our hydrodynamical simulations, and compare them with those from SAMs and observations.

## 4.1 Stellar relations

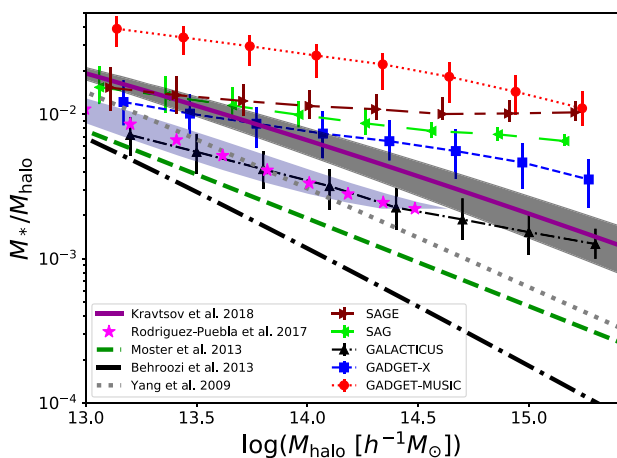
### 4.1.1 Stellar-to-halo mass relation

How galaxy properties relate to their host DM halo is an open question in astronomy. Therefore, a substantial effort has focused on establishing robust determinations of the galaxy–halo connection, commonly reported in the form of the stellar-to-halo mass relation, SHMR (Guo et al. 2010; Yang et al. 2012; Behroozi et al. 2013; Moster et al. 2013, and references therein). In Fig. 6, we compare our SHMR with results from the literature. It is worth noting here that the haloes from the comprehensive sample with mass below

the completeness limit constitute a biased sample, which are lying in a dense environment compared to observations. We only include central galaxies in the calculation as the haloes of satellites galaxies will have suffered tidal disruption. However, as the hydrodynamical simulations feature stars in the halo (which can be treated as intra-cluster light, hereafter ICL), we also include the mass of the ICL in the calculation for the SAMs, SAG, and SAGE. Therefore, the central galaxy here is BCG+ICL. In agreement with our previous findings in Figs 1 and 5, GADGET-MUSIC has the highest stellar-to-halo-mass fraction. SAGE, SAG, and GADGET-X are in the second family, which tend to agree with the observational result at the lower mass end, but deviate from them at the massive end. GALACTICUS, which does not have ICL included, is in better agreement with Rodríguez-Puebla et al. (2017) and Yang et al. (2009). Moreover, we confirm that SAGE also presents a better agreement with the observations if the ICL is excluded. We further note here that the BCG mass from Ragone-Figueroa et al. (2018; a similar cluster simulation based on GADGET-X) is in a good agreement with observational results after applying a cut in radius. In addition, Pillepich et al. (2018) also reported that the exact functional form and magnitude of the SHMR strongly depend on the definition of a central galaxy’s stellar mass. Therefore, the differences shown in this plot could be simply caused by the definition of the central galaxy. We further include the fitting result from Kravtsov et al. (2018), who claim to account for the stellar mass in the same way as the model results here, i.e. BCG mass plus ICL mass. It is interesting to see that their  $M_{\text{BCG}} - M_{\text{halo}}$  relation is much closer to the results from our models (except GADGET-MUSIC which is far from any observation results and GALACTICUS which does not include ICL), especially at  $M_{\text{halo}} \lesssim 10^{14} h^{-1} M_{\odot}$ . However, the offsets between the solid purple line and our model results (including GALACTICUS when compared with the observational results that do not include ICL) are still large for the most massive haloes.



**Figure 5.** The baryonic fractions from the two hydrodynamical simulations within  $R_{500}$ . Gas fractions are shown on the left-hand side panels, while stellar fractions are shown on the right-hand side panels. As shown in the legend on the top-left-hand side panel, hydrodynamical simulations are shown with red filled symbols (median value) with error bars (16th–84th percentile) for GADGET-MUSIC and blue stars with error bars for GADGET-X. Observational data points from Gonzalez et al. (2013) and Zhang et al. (2011) are shown as black stars and magenta cross symbols, respectively, while the lime dotted line shows the fitting result from Kravtsov, Vikhlinin & Meshcheryakov (2018) with the grey shaded scatter. The thick black horizontal dashed lines on the left-hand side panels indicate the cosmic baryon fraction ( $\Omega_b/\Omega_m$ ). The vertical dashed lines in the upper row show the mass limit for the complete sample.



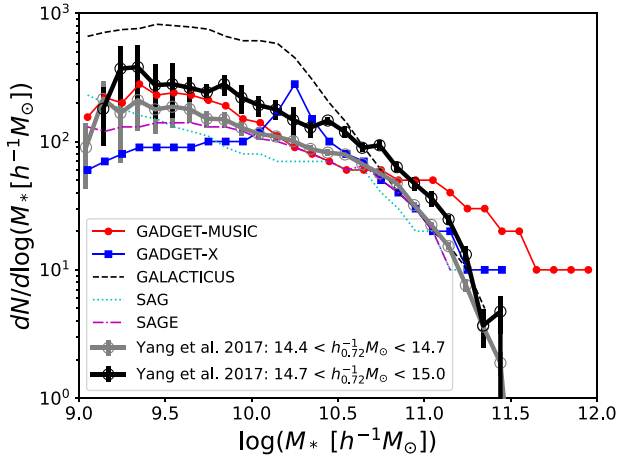
**Figure 6.** The stellar-to-halo mass relation for central galaxies in the complete sample. As indicated in the legend, observational results are shown as thick lines [Yang, Mo & van den Bosch (2009), grey dotted line, Behroozi et al. (2013), dot-dashed black line and Moster, Naab & White (2013), green dashed line] with the latest results from Rodríguez-Puebla et al. (2017) shown as magenta stars with the light shaded area and Kravtsov et al. (2018) as a solid purple line with the dark shaded region. Our hydrodynamical simulation and SAM results are shown in different symbols (median value) with error bars (16th–84th percentile): GADGET-MUSIC with red solid circles and dotted line; GADGET-X with blue solid squares and dashed line; GALACTICUS with black filled triangles and dash-dotted line, SAG with lime triangles and long dashed line and SAGE with maroon triangles and long-short dashed line.

This means that the quenching of star formation in these massive clusters is still problematic for the models investigated here.

In order to check for the properties and influence of the ICL, for example the fraction, the evolution and the connection to the SHMR, we will perform a detailed investigation for both SAMs and the hydrodynamical simulations through carefully separating BCG from ICL, and present the results in a follow-up work (Cañas et al. in preparation).

#### 4.1.2 Stellar mass function for satellite galaxies

Though the satellite-galaxy stellar-mass function is not a scaling relation, we briefly switch focus from central galaxies to satellite galaxies and present the result in this subsection. We only use the mass-complete sample for this investigation and limit our satellite galaxies to objects within  $R_{200}$  as per the observational sample. We show the stellar mass function – median averaged over all clusters – in Fig. 7. As indicated in the legend, different style thin lines represent different versions of the simulations and SAMs, while observational results from Yang et al. (2018) at two different cluster mass bins are highlighted as thick lines. Note that the complete cluster sample is used here without further binning in halo mass, because its mass limit is basically comparable with Yang’s most massive mass bin. The lower mass bin from Yang’s catalogue is presented here to aid the comparison. The horizontal extensions to the red and blue curves are artefacts of the median values. Compared to the observational results, GADGET-MUSIC has more massive satellite galaxies with masses  $M_* > 10^{11.5} h^{-1} M_\odot$ . GADGET-X shows a slightly reduced number of satellite galaxies towards the low-mass end. GALACTICUS features the opposite trend. These deviations from the actual observations can be understood as an overabundance of massive satellite galaxies in GADGET-MUSIC due to the lack of AGN feedback; too few low-mass satellite galaxies in GADGET-X can be caused by either a resolution issue (note that galaxies of



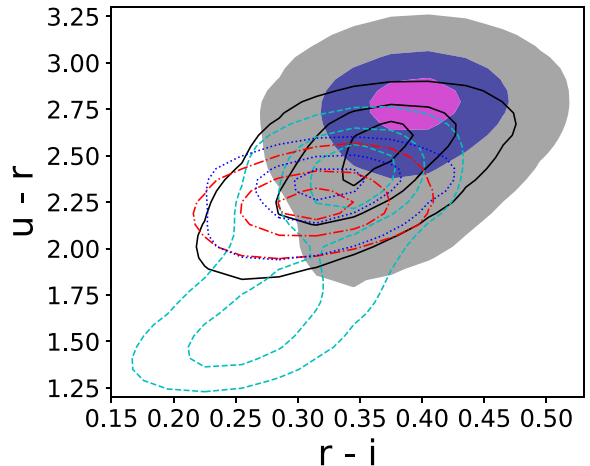
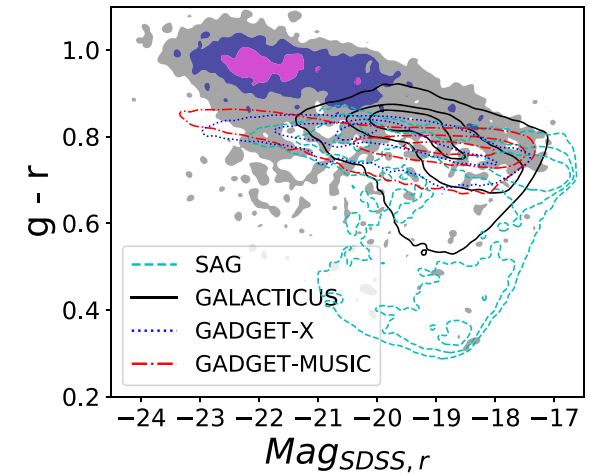
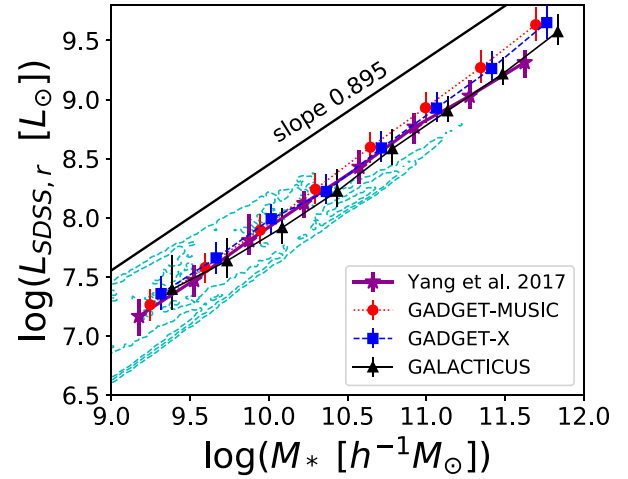
**Figure 7.** The median stellar mass function of satellite galaxies within the mass-complete cluster sample. *GADGET-MUSIC* is shown with a red line with circle symbols and *GADGET-X* with a blue line with square symbols. The three SAMs are presented by different lines: *GALACTICUS* as a black dashed line, *SAG* as a cyan dotted line, and *SAGE* as a magenta dot dashed line. They are compared with observational results from Yang et al. (2018), which are shown in thick black for halo mass range  $[10^{14.7} - 10^{15} h_{0.72}^{-1} M_{\odot}]$  and thick grey for halo mass range  $[10^{14.4} - 10^{14.7} h_{0.72}^{-1} M_{\odot}]$ , both lines include error bars.

$M_* \approx 10^{10} h^{-1} M_{\odot}$  only contain a few hundreds of stellar particles due to the poor simulation resolution) or the striped/heated gas due to the Wendland kernel and feedback; too many low-mass satellite galaxies in *GALACTICUS* is because of a surplus of orphan galaxies (see Table 2 in Knebe et al. 2018). *SAG* and *SAGE* seem not to suffer from this problem due to their different treatment of the orphan galaxy population. We refer to Pujol et al. (2017) for a detailed comparison of the orphan galaxies between different SAMs. However, we note that the scatter across models seen here is at the level found in previous comparisons of theoretically modelled galaxy stellar mass functions of galaxies (Knebe et al. 2015, 2018).

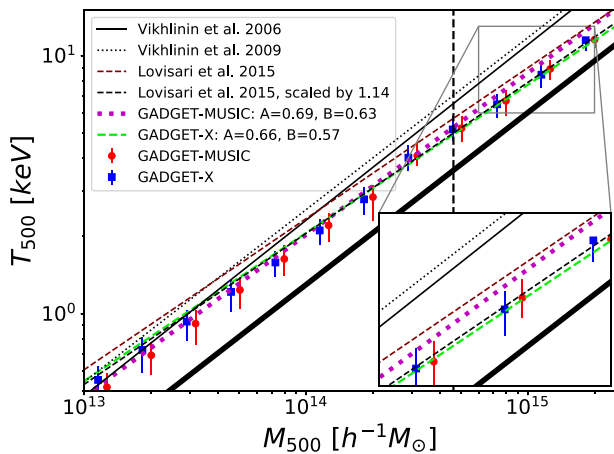
#### 4.1.3 Optical scaling relations

We continue to investigate the correlations between luminosity/magnitude, stellar mass, and colours by comparing our modelled galaxies to the observational results from Yang et al. (2018). We again only use the galaxies from our mass-complete sample here. For a fair comparison to our theoretical data, we apply the same mass cut ( $M_{200} \geq 6.42 \times 10^{14} h^{-1} M_{\odot}$ ) to the group catalogue of Yang et al. (2018) and use all the satellites and central galaxies with  $M_* > 10^9 h^{-1} M_{\odot}$  in these selected groups (the same criteria also applied to our complete sample). The results can be viewed in Fig. 8 where the top panel shows the luminosity–stellar mass relation (based upon the SDSS-*r* band), the middle panel presents the  $g - r$  colour–magnitude (at SDSS-*r* band) relation, and the bottom panel shows the colour–colour relation with  $u - r$  versus  $r - i$ . Note that the *SAGE* model does not provide luminosities ab initio and has hence been excluded from this plot. Similar to Fig. 5, the contours are drawn at the same percentile density levels (16th, 50th, and 84th) after a normalized 2D binning with the observational results shown as different colour-filled areas.

In the top panel, we recover a very tight correlation between luminosity and stellar mass with little variation between obser-



**Figure 8.** Top panel: the luminosity–stellar mass relation for all the galaxies inside the clusters (using the SDSS-*r* band). As indicated in the legend, different symbols (median value) with error bars (16th – 84th percentile) are for different models and for the observational result from Yang et al. (2018), while the result from *SAG* is presented in cyan contours. The top sloping black line (shifted up by 0.5 dex) shows the slope 0.895 which fits both the models and the observational result. Middle panel: the colour–magnitude relation for the galaxies inside the clusters. Bottom panel: the colour–colour relation for galaxies inside the clusters. The legend in the middle panel distinguishes the colours for the models with different line styles for both middle and bottom panels with the colour map is again from Yang et al. (2018).



**Figure 9.** The temperature–mass relation for the clusters from the two hydrodynamical simulations. Red filled circles (blue filled squares) with error bars (16th – 84th percentile) are for GADGET-MUSIC (for GADGET-X). The solid and dotted black lines show the observational results from Vikhlinin et al. (2006) and Vikhlinin et al. (2009), respectively. The maroon dashed line shows the fitting result from Lovisari, Reiprich & Schellenberger (2015; scaled by 1.14 as a black dashed line). Our fitting results from GADGET-MUSIC and GADGET-X are presented by magenta dotted and lime dashed lines, respectively. The thick solid black line shows the self-similar relation  $T_{500} \propto M_{500}^{2/3}$  predicted from non-radiative simulations.

vation and the models (excluding SAG). GADGET-MUSIC, GADGET-X, GALACTICUS, and Yang’s observational results are binned only in stellar mass and presented by symbols with error bars indicating the 16th–84th percentile. While SAG, which tends to have a larger spread in luminosity, is shown with cyan contours. Moreover, we fit the  $M_*$ –luminosity relation for the models (excluding SAG) and the observational result with a linear function  $f(x) = ax$ . We find that all the models give a consistent result with a slope of 0.895, which is shown by the solid black line shifted up by 0.5 dex in the top panel. In the colour–magnitude relation, both hydrodynamical simulations and SAMs show values of  $\sim 0.1 - 0.2$  below the  $g - r$  colour of the observations. There are very few galaxies with a  $g - r$  colour less than 0.7 in the observational results compared to the SAMs. This indicates that the SAMs – as applied to a full cosmological simulation here – fail to reduce their star-forming galaxies sufficiently in the cluster environment. The hydrodynamical simulations also have problems in ceasing star formation, especially for the brightest galaxies. For the colour–colour plot presented in the bottom panel, the results from the two hydrodynamical simulations are in agreement with the two SAMs. Although they all show a noticeable overlap with the observational results, the peaks for the four models are slightly shifted to smaller values in both colours compared to the observations.

## 4.2 Gas scaling relations

For the gas scaling relations, we now use our comprehensive sample of objects, but restrict our analysis to the hydrodynamical simulations for which we have immediate access to multiple gas properties. We confine the analysis to  $M_{500}$  by reselecting all gas particles within  $R_{500}$  to facilitate direct comparison to the observational results.

We first investigate the temperature–mass ( $T - M$ ) relation. The gas temperature is computed using the mass-weighted temperature formula  $T = \sum_i T_i m_i / \sum_i m_i$ , where  $T_i$  and  $m_i$  are the temperature and mass of a gas particle, respectively. In Fig. 9, we show the

**Table 6.** The fitted parameters for the  $T_{500} - M_{500}$  relation with fitting function:  $T_{500} = 10^A (M_{500} / 6 \times 10^{14} M_\odot)^B$ , see equation (3) for details.

Simulation	A	B
GADGET-MUSIC	$0.688 \pm 0.011$	$0.627 \pm 0.007$
GADGET-X	$0.663 \pm 0.012$	$0.574 \pm 0.008$

relation between the mass-weighted gas temperature and  $M_{500}$ . We apply a simple linear fitting function in logarithm space to fit the data from all the samples:

$$T_{500} = 10^A \left( \frac{M_{500}}{6 \times 10^{14} M_\odot} \right)^B. \quad (3)$$

We especially note here that we exclude the  $h$  in the normalization mass of the fitting equation (3).

Since, as discussed above, our comprehensive cluster sample is not complete at the low-mass end, data points below our completeness threshold are weighted according to their completeness during the fitting. As the comprehensive sample forms a mass-incomplete set of haloes, they may conceivably be a biased data set. Such a bias could in principle arise due to their physical proximity to a larger halo but how to accurately quantify such a bias, if it exists, is unclear. Best-fitting curves are shown as a magenta dotted line for GADGET-MUSIC and a green dashed line for GADGET-X; the parameters are summarized both in the legend and Table 6. Since the low-mass data has less weight and there are few clusters in the high mass range, it is not surprising to see that the fitting lines are offset from the symbols which show the median values in each mass bin.

The best-fitting parameters are slightly different between the two hydrodynamical simulations: GADGET-MUSIC has a steeper slope close to the self-similar relation with  $B = 2/3$  (Kaiser 1986, also predicted by the non-radiative simulations, see Bryan & Norman 1998; Thomas et al. 2001 for example) compared to GADGET-X. This is mainly caused by the low temperature of the clusters with small halo mass. Compared to the results from Vikhlinin et al. (2006, 2009), there is a good agreement at low halo mass with our simulations. However, there is a clear offset between our simulation result and their results for massive haloes. This could be caused by the hydrostatic method used in observations which can underestimate the total mass due to a non-thermal pressure component. This bias has been corrected in Lovisari et al. (2015), which, although it is still above our best-fitting lines, is closer to our data for the most massive haloes (closer to GADGET-MUSIC than to GADGET-X). In addition, their result is also slightly higher than our simulation results at low halo mass. This is because of the spectrum-weighted temperature adopted in Lovisari et al. (2015), which is about 14 per cent higher than the mass-weighted temperature (Vikhlinin et al. 2006; Biffi et al. 2014). We follow Biffi et al. (2014) by correcting for this difference by scaling down the fitting function from Lovisari et al. (2015) by a factor of 1.14 (black dashed line in Fig. 9). This produces a very good match to the fitting result from GADGET-X. It is worth noting that the self-similar relation does not provide a good fit to our data (see also Truong et al. 2018). Lastly, Truong et al. (2018) reported lower temperatures than observed resulting in a normalization shift of about 10 per cent for the  $T - M$  relation for their AGN model. Similarly, Henden et al. (2018) also found such a difference with zoomed-in cluster simulations. However, they claimed this is most likely caused by the underestimated total mass due to the biased X-ray hydrostatic mass than a lower temperature in their simulation.

The Sunyaev-Zel'dovich (SZ) effect (Sunyaev & Zeldovich 1970) – which is the diffusion of cosmic microwave background photons within a hot plasma (normally inside galaxy clusters) due to inverse Compton scattering – provides a unique view of a galaxy cluster. Therefore, it has become one of the most powerful cosmological tools used to study the ICM, as well as the nature of the DM and dark energy components of the Universe. Numerous works have been devoted to investigate and understand this effect, both observationally (e.g. Staniszewski et al. 2009; Marriage et al. 2011; Planck Collaboration XXXII 2015) and theoretically by means of cosmological simulations (e.g. da Silva et al. 2000; Sembolini et al. 2013; Le Brun, McCarthy & Melin 2015; Dolag, Komatsu & Sunyaev 2016).

The thermal SZ signal is characterized by the dimensionless Compton  $y$ -parameter, which is defined as

$$y = \frac{\sigma_T k_B}{m_e c^2} \int n_e T_e dl, \quad (4)$$

here  $\sigma_T$  is the Thomson cross-section,  $k_B$  the Boltzmann constant,  $c$  the speed of light,  $m_e$  the electron rest-mass,  $n_e$  the electron number density, and  $T_e$  the electron temperature. The integration is done along the observer's line of sight. In the hydrodynamical simulations, the electron number density,  $n_e$ , for one gas particle can be represented as  $n_e = N_e/dV = N_e/dA/dl$ , here  $N_e$  is the number of electrons in the gas particle,  $dV$  is its spatial volume which is broken down into  $dA$  (the projected area), and  $dl$  (the line-of-sight distance). Therefore, the integration can be represented by the summation (Sembolini et al. 2013; Le Brun et al. 2015):

$$y = \frac{\sigma_T k_B}{m_e c^2 dA} \sum_i T_i N_{e,i} W(r, h_i). \quad (5)$$

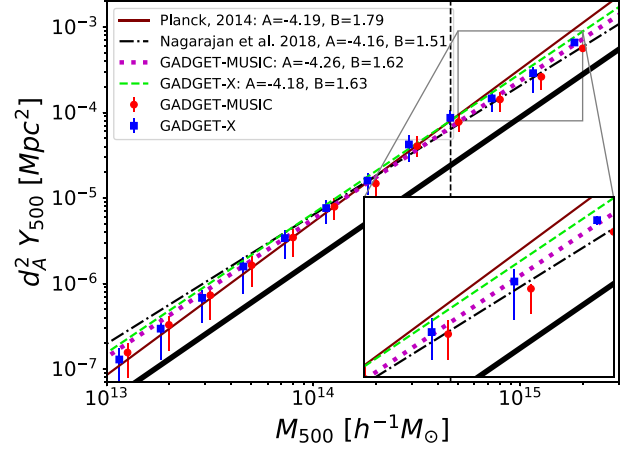
Here we applied the same SPH smoothing kernel  $W(r, h_i)$  as the hydrodynamical simulation to smear the  $y$  signal from each gas particle to the projected image pixels where  $h_i$  is the gas smoothing length from the simulations. It is worth noting that the number of electrons per gas particle is metallicity dependent:  $N_i = \frac{N_e m_i (1-Z-Y_{\text{He}})}{\mu m_p}$ , where  $N_e$  is the number of ionized electrons per hydrogen particle,  $m_i$  the mass of the gas particle,  $Z$  the metallicity of the gas particle,  $Y_{\text{He}}$  the helium mass fraction of the gas particle,  $\mu$  the mean molecular weight, and  $m_p$  the proton mass.<sup>13</sup>

The integrated Comptonization parameter  $Y$  over an aperture inside  $R$  is given by

$$Y = \int y d\Omega = \sum_{i \in R} y_i, \quad (6)$$

where  $\Omega$  is a solid angle, which can be expressed as an aperture of radius  $R$ . In observations, this  $Y$  parameter is normally re-expressed as  $d_A(z)^2 E(z) Y$ , where  $d_A(z)$  is the angular diameter distance and  $E(z) = H(z)/H_0 = \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}$  gives the redshift evolution of the Hubble parameter,  $H(z)$ , in a flat  $\Lambda$ CDM universe. Here we are only presenting clusters at redshift  $z = 0$ , for which  $E(z) = 1$ . In the subsequent analysis, we focus on  $Y_{500}$  within an aperture of  $R_{500}$ . Moreover, we only present projected results in the  $x$ - $y$  plane here. Since we have a large number of samples, the projection effect should have a negligible impact on our results.

In Fig. 10, we show the scaling between  $Y_{500}$  and  $M_{500}$ . Similar to Fig. 9, symbols with error bars are calculated from our comprehensive sample by binning in mass. We refer to the legend in Fig. 10 for



**Figure 10.** The  $Y_{500} - M_{500}$  relation. Similar to Fig. 9, red circles (median value) with error bars (16th–84th percentile) are for GADGET-MUSIC while blue squares with error bars are for GADGET-X. The thin maroon line comes from the *Planck* observation (Planck Collaboration XX 2014) and the dash-dotted line is the fitted result from Nagarajan et al. (2018) with cluster mass estimated by the weak-lensing method. While the black dotted and lime dashed lines show our fitting results for GADGET-MUSIC and GADGET-X, respectively. The lower thick black line shows the self-similar relation  $Y_{500} \propto M_{500}^{5/3}$ .

**Table 7.** The fitted parameters for the  $Y_{500} - M_{500}$  relation. See equation (7) for details.

Simulation	A	B
GADGET-MUSIC	$-4.26 \pm 0.07$	$1.62 \pm 0.31$
GADGET-X	$-4.18 \pm 0.07$	$1.63 \pm 0.29$

further details. Here, we adopt a similar functional form as used for the  $T - M$  relation to fit the data from our comprehensive sample:

$$d_A^2 Y_{500} = 10^A \left( \frac{M_{500}}{6 \times 10^{14} M_\odot} \right)^B. \quad (7)$$

The best-fitting parameters from Planck Collaboration XX (2014) are  $A = -4.19$  and  $B = 1.79$ , which relies on mass estimates from a mass-proxy relation due to Kravtsov, Vikhlinin & Nagai (2006). The fitting result from Nagarajan et al. (2018) which used the weak lensing mass of the APEX-SZ clusters, is shown as a purple dash-dotted line with  $A = -4.16$  and  $B = 1.51$ . We fit our simulation data to the same function and present the results in Fig. 10 for GADGET-MUSIC as a black dotted line and for GADGET-X as a green dashed line. The value of the best-fitting parameters are shown in both the figure legend and Table 7. Compared to the best-fitting Planck relation, our simulation results have a slightly flatter slope. However, comparing to the result from Nagarajan et al. (2018) who used a more precise mass estimation method, both GADGET-X and GADGET-MUSIC are slightly above (similar offsets as comparing with the Planck result) the purple line at the high mass end. On the contrary, the Planck (APEX-SZ) fitting line is under (above) the simulation results at the low-mass end ( $M_{500} < 10^{13.5} h^{-1} M_\odot$ ). In addition, GADGET-X only shows a marginally higher amplitude than GADGET-MUSIC, especially at the high-mass end of the relation. Both are also in agreement with the self-similar relation with  $B = 5/3$  (e.g. Bonamente et al. 2008). This means that the scaling between  $M_{500}$  and  $Y_{500}$  is almost independent of the gas physics and is the more robust relation, which is in agreement with Planelles et al.

<sup>13</sup>The analysis pipeline for this calculation is publicly available as a python package from <https://github.com/weiguangcui/pymysz>.

(2017) and Truong et al. (2018), for example. It is worth noting that neither observations used mass  $M_{500} < 10^{14} h^{-1} M_{\odot}$  to do the fitting. It is interesting to see that this scaling relation extends down to mass  $M_{500} = 10^{13} h^{-1} M_{\odot}$  for our models.

## 5 CONCLUSIONS

In this paper we introduce THE THREE HUNDRED project, i.e. a data base of more than 300 synthetic galaxy clusters with mass  $M_{200} > 6 \times 10^{14} h^{-1} M_{\odot}$ . The clusters have been individually modelled in a cosmological volume of side length  $1 h^{-1}$  Gpc with all the relevant baryonic physics (including AGN feedback) using the ‘modern’ SPH code GADGET-X (Beck et al. 2016). The large re-simulation regions of radius  $15 h^{-1}$  Mpc – centred on the 324 most massive galaxy clusters as found in the parent DM-only MDPL2 simulation – contain many additional objects, in total about 5500 objects with a mass  $M_{200} > 10^{13} h^{-1} M_{\odot}$ . This suite of massive galaxy clusters therefore not only allows to study the formation and evolution of a mass-complete sample, but also carefully investigate their environments and the preprocessing of material entering the galaxy cluster.

This introductory paper focuses on presenting the galaxy clusters by primarily studying their redshift  $z = 0$  properties and comparing them to observational data. This serves as a validation of the public data. Additionally, we do have at our disposal the same suite of clusters, but simulated with a ‘classical’ SPH technique and without AGN feedback (i.e. the GADGET-MUSIC code, Sembolini et al. 2013). This forms a comparison benchmark, demonstrating the differences that choices surrounding physical prescriptions can make. We further presented – where appropriate – the results as obtained via three distinct SAMs (GALACTICUS, SAG, and SAGE) that were applied to the underlying DM-only MDPL2 simulation. A comparison between full physics simulations and semi-analytic models of galaxy formation on this scale or with this number of objects adds to the existing efforts of gauging the relevance of various physical processes and its numerical modelling. In subsequent papers we will apply a more elaborate analysis including redshift evolution and formation processes.

We find that our clusters are in reasonable agreement with observations and summarize our main findings as follows:

(i) The cluster mass difference between the hydrodynamical simulations and their DM-only counterpart is very small for  $M_{200}$ , with about 5 per cent scatter. However,  $M_{500}$  is about 2–6 per cent higher in the hydrodynamical simulation than their MDPL2 counterparts at  $4 \times 10^{14} \lesssim M_{500} \lesssim 10^{15}$ , with a large scatter of about 10 per cent. Using the dynamically relaxed sample reduces the scatter in half, but does not change the systematic differences.

(ii) The dynamically relaxed cluster sample has a  $c - M$  relation which appears to be flat for GADGET-X across the considered mass range. The concentrations for GADGET-MUSIC are generally larger (factor of  $\approx 1.3$ ) and in better agreement with observations. In both models the concentrations of the hydrodynamically modelled clusters are larger than those of their DM-only counterparts; for GADGET-MUSIC this applies to the full mass range whereas for GADGET-X concentrations appear unaffected by the inclusion of baryon physics beyond  $10^{15} h^{-1} M_{\odot}$ .

(iii) GADGET-X shows baryonic fractions at  $M_{500} \gtrsim 10^{14} h^{-1} M_{\odot}$  that are generally in agreement with observations, while GADGET-MUSIC forms too many stars due to the lack of AGN feedback. SAG has the highest gas fraction and the lowest stellar fraction in haloes.

SAGE and GALACTICUS share similar gas fractions and stellar fractions (slightly higher in SAGE than GALACTICUS).

(iv) Besides GALACTICUS, all the models included in this study do not produce an SHMR that is consistent with observations. This could be caused by the inclusion of the ICL. Even comparing with the observational result from Kravtsov et al. (2018), which has ICL included, the BCGs in our modelled clusters ( $M_{\text{halo}} \gtrsim 10^{14.5}$ ) are still massive.

(v) For the stellar mass function of the satellite galaxies, GADGET-MUSIC overproduces the number of massive satellites. At lower stellar mass, GALACTICUS (GADGET-X) has more (less) satellites than the observations.

(vi) The hydro runs and GALACTICUS show a linear (with a slope of 0.895) luminosity–mass relation which is very consistent with the observational result. All the models fail to represent the peak position from observations for the colour–magnitude and colour–colour contour.

(vii) For the gas scaling relations, both GADGET-X and GADGET-MUSIC are generally in agreement with the observational temperature–mass and  $Y_{500}$ –mass relations. The fitting for the hydrodynamical simulations extends to  $10^{13} h^{-1} M_{\odot}$ , which shows the power of the scaling relation. The small difference between the two simulations indicates that baryonic processes only have a weak influence on these relations (see also Hahn et al. 2017).

In addition to the publication of the simulations and halo catalogues, we plan to make publicly available a multiwavelength mock observation database (Cui et al., in preparation) which will include observational mock images from radio/SZ, optical bands to X-rays of all our simulated clusters at different redshifts. We will also provide gravitational lensing images and investigate the lensing efficiency in a follow-up paper (Vega-Ferrero et al. in preparation).

We close with the concluding remark that our theoretically modelled galaxies and galaxy clusters generally present similar results and matches to observations – at least on certain scales of interest. However, we do see deviations in multiple aspects between these models and the observations, especially for the massive central galaxy (BCG+ICL). To understand the disagreements and to connect them with the input subgrid baryonic models, we need to (i) extend the comparisons to even smaller scales than the ones presented here, (ii) consistently derive quantities by mimicking observations more quantitatively, and (iii) track the impact of these baryonic models over a wider range of redshifts. Eventually, as our cluster sample contains different physical implementations of various baryonic processes from both hydrodynamic and SAM modelling, this will allow us to investigate, understand, and pin down the differences between our results and connect them back with the underlying physics. Several such follow-up works are already underway and will be presented separately from this introductory paper (e.g. Vega-Martínez et al. in preparation; Li et al. in preparation). Further, in a companion paper (Mostoghiu et al. 2018), we investigate the density profile of these clusters together with its evolution. And in Wang et al. (2018) the analysis is extended to the comprehensive sample of haloes in the re-simulation regions, investigating how the environment affects their properties and, in particular, the star formation rate. Furthermore, disentangling the BCG from ICL (Cañas et al., in prep.) will help us to understand the too massive central galaxy problem in detail.

## ACKNOWLEDGEMENTS

The work has received financial support from the European Union's Horizon 2020 Research and Innovation programme under the Marie Skłodowska-Curie grant agreement number 734374, i.e. the LACEGAL project.<sup>14</sup> The workshop where this work has been finished was sponsored in part by the Higgs Centre for Theoretical Physics at the University of Edinburgh.

The authors would like to thank The Red Española de Supercomputación for granting us computing time at the MareNostrum Supercomputer of the BSC-CNS where most of the cluster simulations have been performed. The MDPL2 simulation has been performed at LRZ Munich within the project pr87yi. The CosmoSim database (<https://www.cosmosim.org>) is a service by the Leibniz Institute for Astrophysics Potsdam (AIP). Part of the computations with GADGET-x have also been performed at the 'Leibniz-Rechenzentrum' with CPU time assigned to the Project 'pr83li'.

This work has made extensive use of the PYTHON packages – IPYTHON with its Jupyter notebook (Pérez & Granger 2007), NUMPY (van der Walt, Colbert & Varoquaux 2011), and SCIPY (Oliphant 2007; Millman & Aivazis 2011). All the figures in this paper are plotted using the python MATPLOTLIB package (Hunter 2007). This research has made use of NASA's Astrophysics Data System and the arXiv preprint server.

WC, AK, GY, and RM are supported by the *Ministerio de Economía y Competitividad* and the *Fondo Europeo de Desarrollo Regional* (MINECO/FEDER, UE) in Spain through grant AYA2015-63810-P. WC further thanks TaiLai Cui (崔泰莱) for all the joys. AK is also supported by the Spanish Red Consolider MultiDark FPA2017-90566-REDC and further thanks Krog for making the days counts. CP acknowledges the Australia Research Council (ARC) Centre of Excellence (CoE) ASTRO 3D through project number CE170100013. PJE is supported by the ARC CoE ASTRO 3D through project number CE170100013. SB acknowledged financial support from PRIN-MIUR grant 2015W7KAWC, the agreement ASI-INAF n.2017-14-H.0, the INFN INDARK grant, the EU H2020 Research and Innovation Programme under the ExaNeSt project (Grant Agreement No. 671553). ER acknowledges the ExaNeSt and Euro Exa projects, funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 671553 and No. 754337 and financial contribution from the agreement ASI-INAF n.2017-14-H.0. DS' fellowship is funded by the *Spanish Ministry of Economy and Competitiveness* (MINECO) under the 2014 *Severo Ochoa* Predoctoral Training Programme. J.V-F acknowledges the hospitality of the Physics & Astronomy Department at the University of Pennsylvania for hosting him during the preparation of this work. YW is supported by the national science foundation of China (No. 11643005). XY is supported by the National Key Basic Research Program of China (No. 2015CB857002), national science foundation of China (No. 11233005, 11621303). JTA acknowledges support from a postgraduate award from STFC. SAC acknowledges funding from *Consejo Nacional de Investigaciones Científicas y Técnicas* (CONICET, PIP-0387), *Agencia Nacional de Promoción Científica y Tecnológica* (ANPCyT, PICT-2013-0317), and *Universidad Nacional de La Plata*(G11-124), Argentina. CVM acknowledges CONICET, Argentina, for their supporting fellowships. ASB, GC, and MDP are supported by Sapienza University of Rome-Progetti di Ricerca Anno 2016. ASB also acknowledges funding from Sapienza Uni-

versità di Roma under minor grant Progetti per Avvio alla Ricerca Anno 2017, prot. AR11715C82402BC7. RC is supported by the MERAC foundation postdoctoral grant awarded to Claudia Lagos and by the Consejo Nacional de Ciencia y Tecnología CONACYT CVU 520137 Scholar 290609 Overseas Scholarship 438594. SE acknowledges financial contribution from the contracts NARO15 ASI-INAF I/037/12/0, ASI 2015-046-R.0, and ASI-INAF n.2017-14-H.0. SEN is a member of the Carrera del Investigador Científico of CONICET. SP is supported by the Fundamental Research Program of Presidium of the RAS #28. JS acknowledges support from the "Centre National d'études spatiales" (CNES) postdoctoral fellowship program as well as from the "l'Oréal-UNESCO pour les femmes et la Science" fellowship program.

## REFERENCES

- Andersson K. et al., 2011, *ApJ*, 738, 48  
 Arthur J. et al., 2017, *MNRAS*, 464, 2027  
 Bahé Y. M. et al., 2017, *MNRAS*, 470, 4186  
 Barnes D. J., Kay S. T., Henson M. A., McCarthy I. G., Schaye J., Jenkins A., 2017a, *MNRAS*, 465, 213  
 Barnes D. J. et al., 2017b, *MNRAS*, 471, 1088  
 Beck A. M. et al., 2016, *MNRAS*, 455, 2110  
 Behroozi P. et al., 2015, *MNRAS*, 454, 3020  
 Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, *ApJ*, 762, 109  
 Benitez-Llambay A., 2015, py-sphviewer: Py-SPHViewer v1.0.0  
 Benson A. J., 2012, *New Astron.*, 17, 175  
 Bernyk M. et al., 2016, *ApJS*, 223, 9  
 Biffi V., Sembolini F., De Petris M., Valdarnini R., Yepes G., Gottlöber S., 2014, *MNRAS*, 439, 588  
 Biffi V. et al., 2016, *ApJ*, 827, 112  
 Biviano A. et al., 2017, *A&A*, 607, A81  
 Bocquet S., Saro A., Dolag K., Mohr J. J., 2016, *MNRAS*, 456, 2361  
 Bonamente M., Joy M., LaRoque S. J., Carlstrom J. E., Nagai D., Marrone D. P., 2008, *ApJ*, 675, 106  
 Bryan G. L., Norman M. L., 1998, *ApJ*, 495, 80  
 Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, *MNRAS*, 321, 559  
 Carlesi E., Knebe A., Yepes G., Gottlöber S., Jiménez J. B., Maroto A. L., 2012, *MNRAS*, 424, 699  
 Chabrier G., 2003, *PASP*, 115, 763  
 Chiu I. et al., 2018, *MNRAS*, 478, 3072  
 Cora S. A. et al., 2018, *MNRAS*, 479, 2  
 Croton D. J. et al., 2016, *ApJS*, 222, 22  
 Cui W., Zhang Y., 2017, Trends in Modern Cosmology, IntechOpen, London, UK, p. 22, 978-953-51-5426-6.  
 Cui W., Springel V., Yang X., De Lucia G., Borgani S., 2011, *MNRAS*, 416, 2997  
 Cui W., Borgani S., Dolag K., Murante G., Tornatore L., 2012, *MNRAS*, 423, 2279  
 Cui W. et al., 2014a, *MNRAS*, 437, 816  
 Cui W., Borgani S., Murante G., 2014b, *MNRAS*, 441, 1769  
 Cui W. et al., 2016a, *MNRAS*, 456, 2566  
 Cui W. et al., 2016b, *MNRAS*, 458, 4052  
 Cui W., Power C., Borgani S., Knebe A., Lewis G. F., Murante G., Poole G. B., 2017, *MNRAS*, 464, 2502  
 da Silva A. C., Barbosa D., Liddle A. R., Thomas P. A., 2000, *MNRAS*, 317, 37  
 Devriendt J. E. G., Guiderdoni B., Sadat R., 1999, *A&A*, 350, 381  
 Dolag K., Bartelmann M., Perrotta F., Baccigalupi C., Moscardini L., Meneghetti M., Tormen G., 2004, *A&A*, 416, 853  
 Dolag K., Komatsu E., Sunyaev R., 2016, *MNRAS*, 463, 1797  
 Eckert D. et al., 2016, *A&A*, 592, A12  
 Elahi P. J. et al., 2016, *MNRAS*, 458, 1096  
 Evrard A. E., Metzler C. A., Navarro J. F., 1996, *ApJ*, 469, 494  
 Gonzalez A. H., Zaritsky D., Zabludoff A. I., 2007, *ApJ*, 666, 147

<sup>14</sup>[https://cordis.europa.eu/project/rcn/207630\\_en.html](https://cordis.europa.eu/project/rcn/207630_en.html)

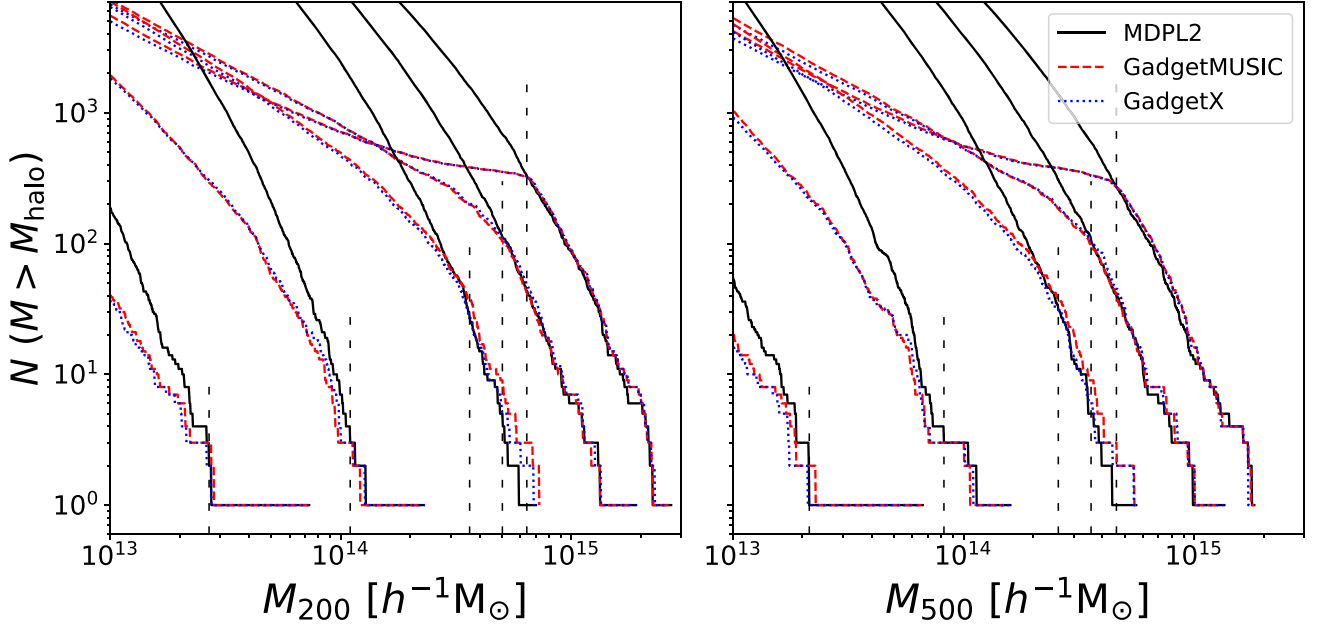
- Gonzalez A. H., Sivanandam S., Zabludoff A. I., Zaritsky D., 2013, *ApJ*, 778, 14
- Groener A. M., Goldberg D. M., Sereno M., 2016, *MNRAS*, 455, 892
- Guo Q., White S., Li C., Boylan-Kolchin M., 2010, *MNRAS*, 404, 1111
- Guo Q. et al., 2016, *MNRAS*, 461, 3457
- Haardt F., Madau P., 1996, *ApJ*, 461, 20
- Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, Clusters of Galaxies and the High Redshift Universe Observed in X-rays, XXIst Moriond astrophysics meeting, CEA, Savoie, France
- Hahn O., Martizzi D., Wu H.-Y., Evrard A. E., Teyssier R., Wechsler R. H., 2017, *MNRAS*, 470, 166
- Henden N. A., Puchwein E., Shen S., Sijacki D., 2018, *MNRAS*, 479, 5385
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Kaiser N., 1986, *MNRAS*, 222, 323
- Kennicutt R. C., Jr, 1998, in Gilmore G., Howell D., eds, ASP Conf. Ser., Vol. 142, The Stellar Initial Mass Function (38th Herstmonceux Conference), Astron. Soc. Pac., San Francisco. p. 1
- Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, *MNRAS*, 457, 4340
- Knebe A., Draganova N., Power C., Yepes G., Hoffman Y., Gottlöber S., Gibson B. K., 2008, *MNRAS*, 386, L52
- Knebe A. et al., 2011, *MNRAS*, 415, 2293
- Knebe A. et al., 2013, *MNRAS*, 435, 1618
- Knebe A. et al., 2015, *MNRAS*, 451, 4029
- Knebe A. et al., 2018, *MNRAS*, 475, 2936
- Knollmann S. R., Knebe A., 2009, *ApJS*, 182, 608
- Kravtsov A. V., Vikhlinin A., Nagai D., 2006, *ApJ*, 650, 128
- Kravtsov A. V., Vikhlinin A. A., Meshcheryakov A. V., 2018, *Astron. Lett.*, 44, 8
- Laganá T. F., Martinet N., Durret F., Lima Neto G. B., Maughan B., Zhang Y.-Y., 2013, *A&A*, 555, A66
- 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., Melin J.-B., 2015, *MNRAS*, 451, 3868
- Lin Y.-T., Mohr J. J., Stanford S. A., 2004, *ApJ*, 610, 745
- Lovisari L., Reiprich T. H., Schellenberger G., 2015, *A&A*, 573, A118
- Mantz A. B., Allen S. W., Morris R. G., Schmidt R. W., von der Linden A., Urban O., 2015, *MNRAS*, 449, 199
- Mantz A. B., Allen S. W., Morris R. G., 2016, *MNRAS*, 462, 681
- Marriage T. A. et al., 2011, *ApJ*, 737, 61
- McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, *MNRAS*, 465, 2936
- Merten J. et al., 2015, *ApJ*, 806, 4
- Millman K. J., Aivazis M., 2011, *Comput. Sci. Eng.*, 13, 9
- Monaco P., Benson A. J., De Lucia G., Fontanot F., Borgani S., Boylan-Kolchin M., 2014, *MNRAS*, 441, 2058
- Moster B. P., Naab T., White S. D. M., 2013, *MNRAS*, 428, 3121
- Mostoghiu R., Knebe A., Cui W., Pearce F. R., Yepes G., Power C., Dave R., Arth A., 2018, *MNRAS*, in review
- Nagarajan A. et al., 2018, *MNRAS* tmp, 1811, preprint ([arXiv:1804.03671](https://arxiv.org/abs/1804.03671))
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Neto A. F. et al., 2007, *MNRAS*, 381, 1450
- Okabe N., Smith G. P., 2016, *MNRAS*, 461, 3794
- Oliphant T. E., 2007, *Comput. Sci. Eng.*, 9, 10
- Pérez F., Granger B. E., 2007, *Comput. Sci. Eng.*, 9, 21
- Pillepich A. et al., 2018, *MNRAS*, 475, 648
- Planck Collaboration XX, 2014, *A&A*, 571, A20
- Planck Collaboration XXXII, 2015, *A&A*, 581, A14
- Planck Collaboration XIII, 2016, *A&A*, 594, A13
- Planelles S., Borgani S., Dolag K., Ettori S., Fabjan D., Murante G., Tornatore L., 2013, *MNRAS*, 431, 1487
- Planelles S. et al., 2017, *MNRAS*, 467, 3827
- Power C., Knebe A., Knollmann S. R., 2012, *MNRAS*, 419, 1576
- Puchwein E., Springel V., Sijacki D., Dolag K., 2010, *MNRAS*, 406, 936
- Pujol A. et al., 2017, *MNRAS*, 469, 749
- Ragone-Figueroa C., Granato G. L., Murante G., Borgani S., Cui W., 2013, *MNRAS*, 436, 1750
- Ragone-Figueroa C., Granato G. L., Ferraro M. E., Murante G., Biffi V., Borgani S., Planelles S., Rasia E., 2018, *MNRAS*, 479, 1125
- Rasia E., Borgani S., Ettori S., Mazzotta P., Meneghetti M., 2013, *ApJ*, 776, 39
- Reiprich T. H., Böhringer H., 2002, *ApJ*, 567, 716
- Rodríguez-Puebla A., Primack J. R., Avila-Reese V., Faber S. M., 2017, *MNRAS*, 470, 651
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Saro A., De Lucia G., Borgani S., Dolag K., 2010, *MNRAS*, 406, 729
- Sartoris B. et al., 2016, *MNRAS*, 459, 1764
- Schaller M. et al., 2015a, *MNRAS*, 451, 1247
- Schaller M. et al., 2015b, *MNRAS*, 452, 343
- Sembolini F., Yepes G., De Petris M., Gottlöber S., Lamagna L., Comis B., 2013, *MNRAS*, 429, 323
- Sembolini F. et al., 2016a, *MNRAS*, 457, 4063
- Sembolini F. et al., 2016b, *MNRAS*, 459, 2973
- Springel V., 2005, *MNRAS*, 364, 1105
- Springel V., Hernquist L., 2003, *MNRAS*, 339, 289
- Staniszewski Z. et al., 2009, *ApJ*, 701, 32
- Steinborn L. K., Dolag K., Hirschmann M., Prieto M. A., Remus R.-S., 2015, *MNRAS*, 448, 1504
- Sunyaev R. A., Zeldovich Y. B., 1970, *Ap&SS*, 7, 3
- Thomas P. A., Muanwong O., Pearce F. R., Couchman H. M. P., Edge A. C., Jenkins A., Onuora L., 2001, *MNRAS*, 324, 450
- Tornatore L., Borgani S., Dolag K., Matteucci F., 2007, *MNRAS*, 382, 1050
- Truong N. et al., 2018, *MNRAS*, 474, 4089
- van der Walt S., Colbert S. C., Varoquaux G., 2011, *Comput. Sci. Eng.*, 13, 22
- Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., 2006, *ApJ*, 640, 691
- Vikhlinin A. et al., 2009, *ApJ*, 692, 1033
- Wang Y. et al., 2018, *ApJ*, in review
- Wiersma R. P. C., Schaye J., Smith B. D., 2009, *MNRAS*, 393, 99
- Wu H.-Y., Evrard A. E., Hahn O., Martizzi D., Teyssier R., Wechsler R. H., 2015, *MNRAS*, 452, 1982
- Yang X., Mo H. J., van den Bosch F. C., 2009, *ApJ*, 695, 900
- Yang X., Mo H. J., van den Bosch F. C., Zhang Y., Han J., 2012, *ApJ*, 752, 41
- Yang X. et al., 2018, *ApJ*, 860, 30
- Yepes G., Kates R., Khokhlov A., Klypin A., 1997, *MNRAS*, 284, 235
- Zhang Y.-Y., Laganá T. F., Pierini D., Puchwein E., Schneider P., Reiprich T. H., 2011, *A&A*, 535, A78
- Zhao D. H., Mo H. J., Jing Y. P., Börner G., 2003a, *MNRAS*, 339, 12
- Zhao D. H., Jing Y. P., Mo H. J., Börner G., 2003b, *ApJ*, 597, L9
- Zibetti S., White S. D. M., Schneider D. P., Brinkmann J., 2005, *MNRAS*, 358, 949

## APPENDIX: EVOLUTION OF THE HALO MASS FUNCTION

Our 30  $h^{-1}$  Mpc diameter re-simulated regions contain many more objects in addition to the central clusters. While there are lots of haloes in the region that surrounds the central cluster there would be many, many more similar haloes in the full volume. It is therefore important to understand the completeness of our comprehensive sample. Here, completeness refers to the total number of haloes above a given mass within a certain cosmological volume. The mass-complete sample in our hydrodynamic simulations is given by  $N_{\text{hydro}}(> M_X) \geq N_{\text{MDPL2}}(> M_X)$ . Here  $N$  is the total number of haloes above a certain mass  $M_X$  with  $X$  is the chosen mass overdensity e.g. 200. i.e. this is the mass above which our sample contains every cluster in the full volume. Below this mass some haloes have not been captured by our re-simulation procedure.

In Fig. A1, we show the cumulative halo mass functions for the two mass definitions  $M_{200}$  (left-hand side panel) and  $M_{500}$  (right-hand side panel) as derived from MDPL2 (solid black lines), GADGET-





**Figure A1.** The cumulative halo mass function from different simulation runs for  $M_{200}$  on the left-hand side panel and  $M_{500}$  on the right-hand side panel. Different colour and line styles represent different simulations: solid black lines are for the DM-only MDPL2, red dashed lines are for GADGET-MUSIC, and blue dotted lines are for GADGET-X. From left to right, we show the halo mass function at redshifts;  $z = 4.0, 2.3, 1.0, 0.5,$  and  $0.0$ , respectively. The dashed vertical lines indicate the mass to which we are complete (i.e. our simulation data set contains all the haloes above this mass in the full simulation volume). Table A1 lists the exact values.

**Table A1.** The mass-complete sample of the Three Hundred cluster catalogues at different redshifts. The first column shows the redshift. The second is the  $M_{200}$  mass limit and the third column gives the values for  $M_{500}$ .

Redshift	$M_{200}$ [ $10^{14} h^{-1} M_{\odot}$ ]	$M_{500}$ [ $10^{14} h^{-1} M_{\odot}$ ]
0.0	6.42	4.60
0.5	5.02	3.57
1.0	3.62	2.57
2.3	1.10	0.82
4.0	0.27	0.21

MUSIC (red dashed lines), and GADGET-X (blue dot-dash lines). There are five families of lines inside each panel, which, from left to right, show the results at  $z = 4.0, 2.3, 1.0, 0.5,$  and  $0.0$ . The mass function of the full halo catalogue from the MDPL2 is used here as a reference line. The vertical dashed lines indicate the mass down to which our sample is complete, determined by the crossing point between the GADGET-X and MDPL2 lines. The mass limit will slightly decrease at some redshifts if GADGET-MUSIC were to be used instead of GADGET-X. This is caused by the baryon effects, as GADGET-MUSIC forms more stars. In order to make sure that the complete sample is chosen to be conservative, we use GADGET-X which returns a higher mass limit. We especially note here that the complete sample is based on the MDPL2 halo mass function. This matching ignores any baryon effects on the halo mass function. However, this could only affect a small number of them near the mass limitation (see Fig. 2 for the mass difference). The precise values for these limits for our mass-complete sample are presented in Table A1.

Below the mass-complete limits the completeness fraction, which will be used later to weight the fitting of the scaling relations, is calculated by the ratio of these lines. It is interesting to note that even at  $z = 1$  the number of clusters in the complete sample has

fallen dramatically. This is because there is significant shuffling in the rank order of the most massive objects in the sample. The set of the largest objects at  $z = 4$  bears little relation to the largest objects at  $z = 0$  and one set does not evolve uniquely into the other. Conversely, the largest objects identified at  $z = 0$  are not all the largest objects at higher redshift and modelling them alone does not produce a large mass-complete sample at earlier times. We further note here that there is only a few mass-complete clusters at  $z \geq 2.3$ . The mass limits are more useful for indicating the boundary of the incomplete sample than for selecting the complete sample for statistical studies.

<sup>1</sup>Departamento de Física Teórica, Módulo 15 Universidad Autónoma de Madrid, E-28049 Madrid, Spain

<sup>2</sup>Centro de Investigación Avanzada en Física Fundamental (CIAFF), Universidad Autónoma de Madrid, E-28049 Madrid, Spain

<sup>3</sup>International Centre for Radio Astronomy Research, The University of Western Australia, 35 Stirling Highway, Crawley, Western Australia 6009, Australia

<sup>4</sup>School of Physics & Astronomy, University of Nottingham, Nottingham NG7 2RD, UK

<sup>5</sup>Institute for Astronomy, School of Physics & Astronomy, The University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ, UK

<sup>6</sup>University Observatory Munich, Scheinerstraße 1, D-81679 Munich, Germany

<sup>7</sup>Max-Planck-Institute for Extraterrestrial Physics, Giessenbachstrasse 1, D-85748 Garching, Germany

<sup>8</sup>Dipartimento di Fisica, Sezione di Astronomia, Università di Trieste, via Tiepolo 11, I-34143 Trieste, Italy

<sup>9</sup>INAF - Osservatorio Astronomico di Trieste, via Tiepolo 11, I-34143 Trieste, Italy

<sup>10</sup>INFN - Sezione di Trieste, via Valerio 2, I-34127 Trieste, Italy

<sup>11</sup>Max-Planck Institute for Astrophysics, Karl-Schwarzschild-Strabetae 1, D-85741 Garching, Germany

<sup>12</sup>*Instituto de Física Teórica, (UAM/CSIC), Universidad Autónoma de Madrid, Cantoblanco, E-28049 Madrid, Spain*

<sup>13</sup>*IFCA, Instituto de Física de Cantabria (UC-CSIC), Av. de Los Castros s/n, E-39005 Santander, Spain*

<sup>14</sup>*School of Physics and Astronomy, Sun Yat-sen University, 519082 Zhuhai, China*

<sup>15</sup>*Department of Astronomy, Shanghai Key Laboratory for Particle Physics and Cosmology, Shanghai Jiao Tong University, Shanghai 200240, China*

<sup>16</sup>*IFSA Collaborative Innovation Center, and Tsung-Dao Lee Institute, Shanghai Jiao Tong University, Shanghai 200240, China*

<sup>17</sup>*Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA 91101, USA*

<sup>18</sup>*Instituto de Astrofísica de La Plata (CCT La Plata, CONICET, UNLP), Paseo del Bosque s/n, B1900FWA, La Plata, Argentina*

<sup>19</sup>*Facultad de Ciencias Astronómicas y Geofísicas, Universidad Nacional de La Plata, Paseo del Bosque s/n, B1900FWA, La Plata, Argentina*

<sup>20</sup>*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia*

<sup>21</sup>*Department of Physics, Sapienza Università di Roma, p.le Aldo Moro 5, I-00185 Rome, Italy*

<sup>22</sup>*Dipartimento di Fisica, Università di Roma Tor Vergata, via della Ricerca Scientifica 1, I-00133 Roma, Italy*

<sup>23</sup>*South African Astronomical Observatory, PO Box 9, Observatory, Cape Town 7935, South Africa*

<sup>24</sup>*Department of Physics and Astronomy, University of the Western Cape, Cape Town 7535, South Africa*

<sup>25</sup>*INFN - Sezione di Roma, P.le A. Moro 2, I-00185 Roma, Italy*

<sup>26</sup>*Dipartimento di Fisica - Università degli Studi di Torino, Via Pietro Giuria, 1, I-10125 Torino, Italy*

<sup>27</sup>*INAF, Osservatorio di Astrofisica e Scienza dello Spazio, via Pietro Gobetti 93/3, I-40129 Bologna, Italy*

<sup>28</sup>*INFN, Sezione di Bologna, viale Berti Pichat 6/2, I-40127 Bologna, Italy*

<sup>29</sup>*Leibniz-Institut für Astrophysik, D-14482 Potsdam, Germany*

<sup>30</sup>*Instituto de Astronomía y Física del Espacio (IAFE, CONICET-UBA), CC 67, Suc. 28, 1428, Buenos Aires, Argentina*

<sup>31</sup>*Facultad de Ciencias Exactas y Naturales (FCEyN), Universidad de Buenos Aires (UBA), 1053, Buenos Aires, Argentina*

<sup>32</sup>*Department of Astronomy & Astrophysics, University of Toronto, Toronto, Canada*

<sup>33</sup>*Astro Space Centre of Lebedev Physical Institute, Russian Academy of Sciences, Profsoyuznaja 84/32, 117997 Moscow, Russia*

<sup>34</sup>*Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230, Saint-Genis-Laval, France*

<sup>35</sup>*CNRS and UPMC Univ. Paris 06, UMR 7095, Institut d'Astrophysique de Paris, 98 bis Boulevard Arago, F-75014 Paris, France*

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.