Sembolini, Federico and Yepes, Gustavo and Pearce, Frazer R. and Knebe, Alexander and Kay, Scott T. and Power, Chris and Cui, Weiguang and Beck, Alexander M. and Borgani, Stefano and Dalla Vecchia, Claudio and Davé, Romeel and Elahi, Pascal Jahan and February, Sean and Huang, Shuiyao and Hobbs, Alex and Katz, Neal and Lau, Erwin and McCarthy, Ian G. and Murante, Giuseppe and Nagai, Daisuke and Nelson, Kaylea and Newton, Richard D. A. and Perret, Valentin and Puchwein, Ewald and Read, Justin I. and Saro, Alexandro and Schaye, Joop and Teyssier, Romain and Thacker, Robert J. (2016) nIFTy galaxy cluster simulations – I. Dark matter and non-radiative models. Monthly Notices of the Royal Astronomical Society, 457 (4). pp. 4063-4080. ISSN 1365-2966

For more information, please contact eprints@nottingham.ac.uk

nIFTy galaxy cluster simulations – I. Dark matter and non-radiative models

Federico Sembolini,[1,2,3]★ Gustavo Yepes,[1,2] Frazer R. Pearce,[4] Alexander Knebe,[1,2]
Scott T. Kay,[5] Chris Power,[6] Weiguang Cui,[6] Alexander M. Beck,[7,8,9]
Stefano Borgani,[10,11,12] Claudio Dalla Vecchia,[13,14] Romeel Davé,[15,16,17]
Pascal Jahan Elahi,[18] Sean February,[19] Shuiyao Huang,[20] Alex Hobbs,[21] Neal Katz,[20]
Erwin Lau,[22,23] Ian G. McCarthy,[24] Guiseppe Murante,[10] Daisuke Nagai,[22,23,25]
Kaylea Nelson,[23,25] Richard D. A. Newton,[5,6] Valentin Perret,[26] Ewald Puchwein,[27]
Justin I. Read,[28] Alexandro Saro,[7,29] Joop Schaye,[30] Romain Teyssier[26]
and Robert J. Thacker[31]

Affiliations are listed at the end of the paper

ABSTRACT

We have simulated the formation of a galaxy cluster in a Λ cold dark matter universe using 13 different codes modelling only gravity and non-radiative hydrodynamics (RAMSES, ART, AREPO, HYDRA and nine incarnations of GADGET). This range of codes includes particle-based, moving and fixed mesh codes as well as both Eulerian and Lagrangian fluid schemes. The various GADGET implementations span classic and modern smoothed particle hydrodynamics (SPH) schemes. The goal of this comparison is to assess the reliability of cosmological hydrodynamical simulations of clusters in the simplest astrophysically relevant case, that in which the gas is assumed to be non-radiative. We compare images of the cluster at $z = 0$, global properties such as mass and radial profiles of various dynamical and thermodynamical quantities. The underlying gravitational framework can be aligned very accurately for all the codes allowing a detailed investigation of the differences that develop due to the various gas physics implementations employed. As expected, the mesh-based codes RAMSES, ART and AREPO form extended entropy cores in the gas with rising central gas temperatures. Those codes employing classic SPH schemes show falling entropy profiles all the way into the very centre with correspondingly rising density profiles and central temperature inversions. We show that methods with modern SPH schemes that allow entropy mixing span the range between these two extremes and the latest SPH variants produce gas entropy profiles that are essentially indistinguishable from those obtained with grid-based methods.

Key words: methods: numerical – galaxies: haloes – cosmology: theory – dark matter.

1 INTRODUCTION

Galaxy clusters are the largest virialized objects in the Universe and, as such, provide both an important testbed for our theories of cosmological structure formation and a challenging laboratory within which to study the fundamental physical processes that drive galaxy formation and evolution. The overdense regions that contain clusters in the present-day Universe were the first to collapse and

virialize in the early Universe, and so our theories must predict their assembly history over a large fraction of the lifetime of the Universe (see Allen, Evrard & Mantz 2011; Kravtsov & Borgani 2012, for a recent review). At the same time, galaxies in the cores of clusters have orbited within the often violently growing cluster potential over many dynamical times, while the broader cluster galaxy population is continually replenished by the infall of galaxies from the field.

Computer simulations are now well established as a powerful and indispensable tool in the interpretation of astronomical observations (see for instance Borgani & Kravtsov 2011). Early *N*-body

simulations (White 1976; Fall 1978; Aarseth, Turner & Gott 1979), which included the gravitational effects of dark matter only, were vital in interpreting the results of galaxy redshift surveys and unveiling the large-scale structure of the Universe, and subsequently in resolving structure in the non-linear regime of dark matter halo formation. The focus of modern simulations has now shifted to modelling galaxy formation in a cosmological context (see Borgani & Kravtsov 2011 for a recent review), incorporating the key physical processes that drive galaxy formation – such as the cooling of a collisional gaseous component (e.g. Pearce et al. 2000; Muanwong et al. 2001; Davé, Katz & Weinberg 2002; Kay et al. 2004; Nagai, Kravtsov & Vikhlinin 2007; Wiersma, Schaye & Smith 2009), the birth of stars from cool overdense gas (e.g. Springel & Hernquist 2003; Schaye & Dalla Vecchia 2008), the growth of black holes (Di Matteo, Springel & Hernquist 2005) and the injection of energy into the interstellar medium by supernovae (e.g. Metzler & Evrard 1994; Borgani et al. 2004; Davé, Oppenheimer & Sivanandam 2008; Dalla Vecchia & Schaye 2012) and powerful active galactic nuclei (AGN) outflows (e.g. Thacker, Scannapieco & Couchman 2006; Sijacki et al. 2007, 2008; Puchwein, Sijacki & Springel 2008; Booth & Schaye 2009). These processes span an enormous dynamic range, both spatial and temporal, from the sub-pc scales of black hole growth to the accretion of gas on to Mpc scales from the cosmic web. Galaxy clusters offer an ideal testbed for the study of these processes and their complex interplay, precisely because their enormous size encompasses the range of scales involved. For this reason, the study of galaxy formation and evolution in cluster environments occupies a fundamental position in observational and numerical astrophysics.

This raises the important question of how reliably cosmological galaxy formation simulations recover the properties of galaxy clusters. In the classic Santa Barbara Cluster Comparison Project, Frenk et al. (1999, hereafter SB99) compared the bulk dark matter and gas properties of a galaxy cluster formed in a non-radiative cosmological hydrodynamical simulation run using 12 (then state-of-the-art) mesh- and particle-based (hereafter smoothed particle hydrodynamics, SPH) codes. They displayed visual representations of the cluster at $z = 0$ and 0.5 when a major merger event was ongoing, and compared several observable quantities such as enclosed mass, gas temperature and X-ray emission. The large scatter in essentially all quantities between the 12 models is evident from the plots. The origin of these discrepancies was partly the poor timing agreement between the methods due to an ambiguity in the start redshift as well as large differences in the effective numerical resolution that arose because the simulation challenged the computing resources available at the time. The key discrepancy has, however, stood the test of times: the divergence between mesh-based and SPH codes in the radial entropy profile of the gas, defined in SB99 as

$$S(R) = \log \left[T_{gas}(R)/\rho_{gas}(R)^{2/3} \right] , \qquad (1)$$

where R is the spherical radius with respect to the cluster centre of mass, T_{gas} is the gas temperature and ρ_{gas} is the gas density. Fig. 18 of SB99 showed some tentative indication that the entropy profiles of the grid-based codes (principally those of Bryan and Cen) displayed a constant entropy core whereas the entropy profiles of the particle-based SPH codes continued to fall all the way into the centre.

These results have been confirmed subsequently by several studies (Dolag et al. 2005; O'Shea et al. 2005; Voit, Kay & Bryan 2005; Wadsley, Veeravalli & Couchman 2008; Mitchell et al. 2009; Valdarnini 2012). For example, Wadsley et al. (2008) and Mitchell

et al. (2009) suggested that the discrepancy owes also to the artificial surface tension and the associated lack of multiphase fluid mixing in classic SPH, while similar conclusions have been reached by Sijacki et al. (2012) when comparing the moving mesh code AREPO of Springel (2010) with classic SPH, using P-GADGET3 with the entropy conserving SPH version of Springel & Hernquist (2002). Interestingly, in their recent study, Power, Read & Hobbs (2014) have shown that SPHS (Read & Hayfield 2012), an extension of SPH that improves the treatment of mixing by means of entropy dissipation, produces constant entropy cores in non-radiative galaxy cluster simulations that are consistent with the results of the adaptive mesh refinement (AMR) code RAMSES (Teyssier 2002). Both Wadsley et al. (2008) and Maier et al. (2009) report entropy cores in runs that incorporate subgrid models for turbulence. These results suggest that modern SPH codes can overcome the problems first reported in SB99 and subsequently in Agertz et al. (2007). It is worth noting that it is not obvious that mesh-based codes necessarily recover the correct form for the entropy profile. For example, Springel (2010) reports significant variation in the entropy profile for the same AMR code (ENZO) that is particularly sensitive to choice of refinement criteria. More generally, differences are apparent when comparing AMR results to that of the moving mesh code AREPO; Springel (2010) reports an entropy core that is significantly lower than that found in AMR codes (e.g. compare fig. 45 of Springel 2010 with fig. 18 of SB99 or fig. 5 of Voit et al. 2005).

In this work – emerging out of the 'nIFTy cosmology' workshop[1] – we revisit the idea of the Santa Barbara Cluster Comparison Project 15 yr later. We take five modern cosmological simulation codes (with one of them taken in eight different flavours, for a total of 12 different codes) and study the formation and evolution of a large galaxy cluster (with final mass $M_{200}^{crit} \simeq 1.1 \times 10^{15}$ M$_{\odot}$). First we perform blind dark matter-only simulations with the favoured parameter sets of each group. The results from these show the typical scatter that is expected for currently published works in this area. We then use a common parameter set to align our gravitational framework before continuing to study non-radiative gas simulations. This allows us to focus solely on the differences between the models that arise due to the different hydrodynamical implementations. This also permits us to cleanly study the formation (or not) of a gas entropy core.

The rest of this paper is organized as follows. in Section 2 we briefly describe the 12 methods used in this study and supply tables listing parameter choices. In Section 3 we describe how our initial conditions were generated and chosen. The main results are presented in Section 4, which discusses the dark matter-only simulations, and in Section 5, which presents the results from the non-radiative simulations. We discuss convergence and scatter among the different codes in Section 6. We summarize out results in Section 7. We present additional supporting material in Appendix A.

2 THE CODES

The numerical codes used for this project can be divided into two main groups: mesh-based and SPH codes. The mesh-based codes used in this work are RAMSES (Teyssier 2002), ART (Rudd, Zentner & Kravtsov 2008) and AREPO (Springel 2010): the first two use Eulerian hydrodynamics while AREPO uses a moving unstructured mesh and

---

[1] http://popia.ft.uam.es/nIFTyCosmology

can be considered almost Lagrangian. SPH codes use Lagrangian hydrodynamics: we study HYDRA (Couchman, Thomas & Pearce 1995) and nine different versions of GADGET (Springel 2005). Among the SPH codes we can distinguish classic and modern forms. We define modern variants as those which adopt an improved treatment of discontinuities.

The various codes employ different techniques to solve the evolution equations for a two-component fluid of dark matter and non-radiative gas coupled by gravity. To calculate gravitational forces, RAMSES and ART use AMR while AREPO and GADGET are based on the TREEPM method. HYDRA uses an AP^{3}M approach. Hydrodynamical forces are treated in the following ways: by means of SPH in GADGET and HYDRA, using a Voronoi mesh in AREPO, and using Eulerian AMR in RAMSES.

There follow short summaries detailing the specifics of each simulation code contributing to this comparison and the authors responsible for running each method. We focus on key differences and novel aspects. Generalized descriptions of the individual codes can be found in their respective methods papers, listed in Table 1. Table 2 summarizes the key numerical parameters used by each team.

Table 1. List of all the simulation codes participating in the nIFTy cluster comparison project, including reference and gravity algorithm adopted.

Code name	Reference	Gravity model
RAMSES	Teyssier (2002)	AMR
ART	Rudd et al. (2008)	AMR
AREPO	Springel (2010)	TREEPM
GADGET:	Springel (2005)	TREEPM
G2-ANARCHY	Dalla Vecchia et al. (in preparation)	
G3-X-ART	Beck et al. (2016)	
G3-SPHS	Read & Hayfield (2012)	
G3-MAGNETICUM	Hirschmann et al. (2014)	
G3-PESPH	Huang et al. (in preparation)	
G3-MUSIC	Sembolini et al. (2013)	
G3-OWLS	Schaye et al. (2010)	
G2-X	Pike et al. (2014)	
HYDRA	Couchman et al. (1995)	AP^{3}M

Table 2. Key numerical parameters used for each run. Columns 2 and 3 list values for the Plummer-equivalent softening lengths for the dark matter particles in physical (not comoving) units (kpc), columns 4 and 5 list the same but for the gas particles (where present) and column 6 lists the number of FFT cells along each side of the box. We use the label 'Adp' when adaptive force resolution is used.

Code name	ϵ_{DM}	ϵ_{DM}^{max}	ϵ_{gas}	ϵ_{gas}^{max}	N_{FFT}
RAMSES	11.0	11.0	11.0	11.0	256
ART	Adp	Adp	Adp	Adp	256
AREPO	33.0	5.5	Adp	Adp	512
G2-ANARCHY	20.0	5.0	20.0	5.0	512
G3-X-ART	8.0	6.0	8.0	6.0	512
G3-SPHS	5.0	5.0	Adp	Adp	1024
G3-MAGNETICUM	11.25	3.75	3.75	3.75	256
G3-PESPH	5.0	5.0	5.0	5.0	256
G3-MUSIC	8.0	6.0	8.0	6.0	512
G3-OWLS	9.77	5.0	9.77	5.0	1024
G2-X	24.0	6.0	24.0	6.0	512
HYDRA	Adp	5.0	Adp	5.0	512

2.1 Mesh-based codes

RAMSES (Perret, Teyssier)

RAMSES is an AMR code. For fluid dynamics a directionally unsplit, second-order Godunov scheme with the HLLC Riemann solver is used. The *N*-body solver is an adaptive particle mesh code, for which the Poisson equation is solved using the multigrid technique. The grid is adaptively refined on a cell-by-cell basis, following a quasi-Lagrangian refinement strategy whereby a cell is refined into eight smaller new cells if its dark matter or baryonic mass grows by more than a factor of 8. Time integration is performed using an adaptive, level-by-level, time stepping strategy. Parallel computing is based on the MPI library, with a domain decomposition set by the Peano–Hilbert space filling curve.

ART (Nagai, Lau, Nelson)

ART (Adaptive Refinement Tree, ART) is an *N*-body plus hydrodynamics code with adaptive mesh refinement developed by Kravtsov, Klypin & Khokhlov (1997). The computational grid is refined based on the oct-tree structure: when a property of a given grid cell, e.g. density, exceeds certain predefined threshold, the cell is refined into eight smaller child cells. We solve the Poisson equation using FFT on the base grid, while for each refinement level we employ a multilevel relaxation method. The code solves the inviscid fluid dynamical equations using a second-order accurate Godunov method with piecewise-linear reconstructed boundary states and the exact Riemann solver of Colella & Glaz (1985). The version of the code (Rudd et al. 2008) used for this work has been parallelized for distributed machines using MPI and features a flexible time-stepping hierarchy.

AREPO (Puchwein)

AREPO uses a Godunov scheme on an unstructured moving Voronoi mesh. The mesh cells move (roughly) with the fluid. The main differences to Eulerian AMR codes consist in that AREPO is almost Lagrangian and it is Galilean invariant by construction; furthermore, AREPO has automatic refinement for hydrodynamics and gravity and uses a TREEPM gravity solver. The main difference to SPH codes is that the hydrodynamic equations are solved with a finite-volume Godunov scheme. In this work, we always use the total energy as a conserved quantity in the Godunov scheme rather than the entropy–energy formalism described in Springel (2010).

2.2 SPH codes

G2-ANARCHY (Dalla Vecchia)

G2-ANARCHY is an implementation of GADGET2 employing the pressure-entropy SPH formulation derived by Hopkins (2013). The artificial viscosity switch has been implemented following Cullen & Dehnen (2010), whose algorithm allows precise detection of shocks which avoids excessive viscosity in pure shear flows. G2-ANARCHY uses a purely numerical switch for entropy diffusion similar to the one of Price (2008), but without requiring any diffusion limiter. The kernel adopted is the C^2 function of Wendland (1995) with 100 neighbours, with the purpose of avoiding particle pairing (as suggested by Dehnen & Aly 2012). The time stepping adopted is described in Durier & Dalla Vecchia (2012).

G3-X-ART (Murante)

G3-X-ART is an advanced version of GADGET3 incorporating the Wendland kernel functions (Dehnen & Aly 2012) with 200/295 neighbours, artificial conductivity to promote fluid mixing following Price (2008) and Tricco & Price (2013), but with an additional

limiter for gravitationally induced pressure gradients, a time-step limiting particle wake-up scheme (Pakmor et al. 2012) and a high-order scheme for gradient computation (Price 2012), shock detection and artificial viscosity similar to Cullen & Dehnen (2010) and Hu et al. (2014). A companion paper (Beck et al. 2016) presents an improved hydrodynamical scheme and its performance in a wide range of test problems. In order to clearly illustrate the changes due to an improved SPH implementation within an identical code, this variant of GADGET has also been run with a cubic spline kernel, 64 neighbours, low-order viscosity and a shear flow limiter (Balsara 1995) and no conduction. Results from this version, G3-X-STD, are included on the radial plots below but omitted from the panel figures as the results are in all cases essentially indistinguishable from the other GADGET implementations of classic SPH.

G3-SPHS (Power)

G3-SPHS was developed to overcome the inability of classic SPH to resolve instabilities. It has been implemented in the GADGET3 code. The key differences with respect to standard GADGET3 are in the choice of kernel and in dissipation. Rather than the cubic spline kernel, G3-SPHS uses as an alternative either the HOCT kernel with 442 neighbours or the Wendland C4 kernel with 200 neighbours. A higher order dissipation switch detects, in advance, when particles are going to converge. If this happens, conservative dissipation is switched on for all advected fluid quantities. The dissipation is switched off again once particles are no longer converging. This ensures that all fluid quantities are single valued throughout the flow by construction.

G3-MAGNETICUM (Saro)

G3-MAGNETICUM is an advanced version of GADGET3. In this non-radiative version, a higher order kernel based on the bias-corrected, sixth-order Wendland kernel (Dehnen & Aly 2012) with 295 neighbours is included. The code also incorporates a low viscosity scheme to track turbulence as originally described in Dolag et al. (2005) with improvements following Beck et al. (2016), which uses a high-order scheme (Price 2012) to compute velocity gradients. Thermal conduction is modelled isotropically at 1/20th of the *Spitzer* rate (Dolag et al. 2004). At last, it uses a particle wake-up algorithm (Pakmor et al. 2012; Dolag, private communication) for a more accurate time integration.

G3-PESPH (February)

G3-PESPH is an implementation of GADGET3 with pressure-entropy SPH (Hopkins 2013) which features special galactic wind models. The SPH kernel is an HOCTS ($n = 5$) B-spline with 128 neighbours. The time-dependent artificial viscosity is that of Morris & Monaghan (1997).

G3-MUSIC (Yepes)

The original MUSIC runs (G3-MUSIC) were completed with the GADGET3 code, based on the entropy-conserving formulation of SPH (Springel & Hernquist 2002). GADGET3 employs a spline kernel (Monaghan & Lattanzio 1985) and parametrized artificial viscosity following the model described by Monaghan (1997).

G3-OWLS (McCarthy)

The G3-OWLS simulations were run using a version of GADGET3 which was significantly modified to include new subgrid physics for radiative cooling, star formation, stellar feedback, black hole growth and AGN feedback, developed as part of the OWLS/cosmo-OWLS projects (Schaye et al. 2010). Standard entropy-conserving SPH (Le Brun et al. 2014) was used with 48 neighbours.

G2-X (Kay)

G2-X is a modified version of the GADGET2 code, using the TREEPM gravity solver and standard entropy-conserving SPH. It includes a number of subgrid modules to model metal-dependent radiative cooling, star formation and feedback from supernovae and AGN.

HYDRA (Thacker)

HYDRA-OMP (Thacker & Couchman 2006) is a parallel implementation of the earlier serial HYDRA code (Couchman et al. 1995). Aside from the parallel nature of the code, HYDRA-OMP differs from the serial implementation by using a standard pair-wise artificial viscosity along with the Balsara modification for rotating flows. Otherwise, the SPH implementation is classic in nature, using 52 neighbours, and does not include more recently preferred kernels, terms for conduction or explicit shock tracking to modify the dissipation. It also uses a conservative time-stepping scheme that keeps all particles on the same synchronization.

Colour scheme

In all the radial plots below we adopt the following colour scheme to distinguish the various underlying code groupings amongst our implementations. The AMR codes RAMSES and ART and the moving-mesh code AREPO have black lines. For the nine GADGET codes we use three colours, blue, yellow and red. Blue (G2-ANARCHY, G3-X-ART and G3-SPHS) indicates modern SPH codes with artificial mixing and conduction, yellow (G3-MAGNETICUM and G3-PESPH) indicates modern SPH codes with conduction and red (G3-MUSIC, G3-X-STD, G3-OWLS and G2-X) indicates classic SPH implementations. HYDRA, also a classic SPH code, is shown in green.

2.3 Progress with modern SPH codes

Since the first development of SPH by Gingold & Monaghan (1977) and Lucy (1977) the accuracy and stability of SPH simulations have been greatly improved. In particular, much attention has been given to the treatment of discontinuities where artificial viscosity is necessary for a proper fluid sampling at shocks and to prevent particle interpenetration. The first spatially constant low-order formulations of artificial viscosity applied viscosity not only in shocks, but also in shearing flows and unshocked regions leading to an overdiffusion of kinetic energy. Modern formulations of artificial viscosity rely on improved shock detection methods and high-order gradient estimators to distinguish between shocked and unshocked or shearing regions (Morris & Monaghan 1997; Cullen & Dehnen 2010; Price 2012; Hu et al. 2014; Biffi & Valdarnini 2015). Most importantly, they preserve kinetic energy to a much higher degree and follow the growth of turbulence and hydrodynamical instabilities. The lack of turbulence growth within classic SPH implies that these older forms fail to treat any interfaces between different gas phases and their subsequent mixing adequately. This is sometimes expressed as a lack of diffusive terms or as an artificial spurious surface tension, as shown for instance by Agertz et al. (2007). However, it should be stressed that the actual mixing behaviour of the various gas phases within the intracluster medium is still not well known, a point we will revisit later.

Read, Hayfield & Agertz (2010) showed that the mixing problem in SPH arises from two causes: the force inaccuracy and the lack of entropy mixing. Artificial conduction (Price 2008; Valdarnini 2012) and pressure-entropy (Ritchie & Thomas 2001; Hopkins 2013; Saitoh & Makino 2013) formulations have been developed to overcome these issues. They act by providing for the transport of heat between particles or by adjusting the physical state slightly. However, in the presence of gravitationally induced pressure and

Table 3. SPH schemata used for the comparison runs. We list the employed kernel functions and number of neighbours (N_{SPH}) as well as the minimum smoothing length (h_{min}) in terms of the gravitational softening length. Furthermore, we provide in the last four columns information about the artificial viscosity and conductivity switches.

Code name	Hydrodyn. kernel	N_{SPH}	h_{min}	Art. visc.	Shear flow	Mixing	Limiter
G2-ANARCHY	Wendland C2	100 ± 3	0	Adaptive	Low	Artificial	Yes
G3-X-ART	Wendland C6	295 ± 10	0.1	Adaptive	High	Artificial	Yes
G3-SPHS	Wendland C4	200 ± 0.5	1.0	Adaptive	Low	Artificial	Yes
G3-MAGNETICUM	Wendland C6	295 ± 0.5	0.001	Adaptive	High	Physical	–
G3-PESPH	HOCTS B-spline	128 ± 2	0.1	Adaptive	Low	–	–
G3-MUSIC	Cubic spline	40 ± 3	0.1	Constant	Low	–	–
G3-X-STD	Cubic spline	64 ± 1	0.1	Constant	Low	–	–
G3-OWLS	Cubic spline	48 ± 1	0.01	Constant	Low	–	–
G2-X	Cubic spline	50 ± 3	1	Constant	Low	–	–
HYDRA	Cubic spline	$53 \pm 1\,5$	0.5	Constant	Low	–	–

temperature gradients, artificial conduction schemes can lead to excessive (unwanted) transport of heat, making the use of numerical limiters as well as correction terms highly desirable. Read & Hayfield (2012) showed that pressure-entropy SPH fails for strong shocks and/or if the density gradient is large. This was significantly improved by Hopkins (2013) who derived a conservative pressure-entropy SPH for the first time. However, the problem with large density gradients remained. Read & Hayfield (2012) proposed instead to use higher order switching, similar to Cullen & Dehnen (2010), but applied to all advected fluid quantities. This solved the mixing problem in SPH also for high density contrasts and opened the door to multimass SPH for the first time.

Lastly, the commonly employed cubic spline kernel function can easily become unstable: this leads to a pairing instability, incorrect gradient estimators and poor fluid sampling. A change of kernel function is therefore highly recommended; Wendland kernels (Dehnen & Aly 2012) are now commonly used to retain fluid sampling. Table 3 provides an overview of the different SPH codes participating in our cluster comparison project and lists their numerical details.

3 THE SIMULATION

The cluster we have adopted for this project was drawn from the MUSIC-2 sample (Sembolini et al. 2013, 2014; Biffi et al. 2014) which consists of a mass-limited sample[2] of resimulated haloes selected from the MultiDark cosmological simulation (Prada et al. 2012). This simulation is dark matter only and contains 2048^3 (almost 9 billion) particles in a $(1\,h^{-1}\,\text{Gpc})^3$ cube. It was performed in 2010 using ART (Kravtsov et al. 1997) at the NASA Ames Research Centre. All the data from this simulation are accessible online through the *MultiDark Database*.[3] The run was done using the best-fitting cosmological parameters to WMAP7+BAO+SNI ($\Omega_M = 0.27$, $\Omega_b = 0.0469$, $\Omega_\Lambda = 0.73$, $\sigma_8 = 0.82$, $n = 0.95$, $h = 0.7$; Komatsu et al. 2011). This is the reference cosmological model used in the rest of the paper.

The MUSIC-2 cluster catalogue was originally constructed by selecting all the objects in the simulation box which are more massive than $10^{15}\,h^{-1}\,\text{M}_\odot$ at redshift $z = 0$. In total, 282 objects were found above this mass limit. A zooming technique, described in Klypin et al. (2001), was used to produce the initial conditions for

[2] Specifically, it is cluster 19 of MUSIC-2 data set; all the initial conditions of MUSIC clusters are available at http://music.ft.uam.es
[3] www.MultiDark.org

the resimulations. All particles within a sphere with a radius of $6\,h^{-1}$ Mpc around the centre of each selected object at $z = 0$ were found in a low-resolution version (256^3 particles) of the MultiDark volume. This set of particles was then mapped back to the initial conditions to identify the Lagrangian region corresponding to a $6\,h^{-1}$ Mpc radius sphere centred on the cluster centre of mass at $z = 0$. The initial conditions of the original simulations were generated on a finer mesh of size 4096^3. By doing so, the mass resolution of the resimulated objects was improved by a factor of 8 with respect to the original simulations. In the high-resolution region the mass resolution for the dark matter-only simulations corresponds to $m_{DM} = 1.09 \times 10^9\,h^{-1}\,\text{M}_\odot$. For the runs with gas physics, $m_{DM} = 9.01 \times 10^8\,h^{-1}\,\text{M}_\odot$ and $m_{gas} = 1.9 \times 10^8\,h^{-1}\,\text{M}_\odot$.

4 DARK MATTER RUN COMPARISON

We first completed a dark matter-only simulation, performed using the parameter values given in Table 2. These were chosen independently by each modelling group following their previous experience, so this is in essentially a blind comparison without iteration. A visual comparison of the smoothed density field centred on the cluster at $z = 0$ is given in Fig. 1. While it is clear that all the methods have followed the formation of the same object (with a significant improvement with respect to fig. 1 of SB99) the precise location of the major subhalo (which is the object located at 7 o'clock with respect to the centre of the cluster in Fig. 1) is not accurately recovered. The variance across this figure illustrates the typical range of outcomes from commonly used cosmological codes. A possible major cause of the discrepancy (for the GADGET based codes at least) is the size of the base-level particle mesh, at least for GADGET-based codes. Those implementations which employed a base mesh of 256^3 did not sufficiently resolve the interface region between this low-resolution mesh and the higher resolution region placed over the cluster: G3-MAGNETICUM, G3-PESPH and G2-X – the GADGET codes which adopted a base mesh of 256^3 – show the major subhalo positioned closer to R_{200}^{crit} (the subhalo centre is located at $\sim 0.9 R_{200}^{crit}$ from the cluster centre) with respect to the codes with a higher base mesh ($\sim 0.75 R_{200}^{crit}$ from the cluster centre). In order to test this effect, we repeated the G3-MUSIC run changing only the base-level mesh from 512^3 to 256^3 and keeping all the other parameters unaltered; as expected, we found that the major subhalo had sensibly moved closer to the virial ratio of the cluster with respect to the original run.

Improving the resolution of the base-level mesh to 512^3 alleviates therefore this problem and aligns the dark matter component well. We illustrate this in Appendix A by reproducing Fig. 1 with all the base-level meshes set to 512^3. For the non-radiative comparison
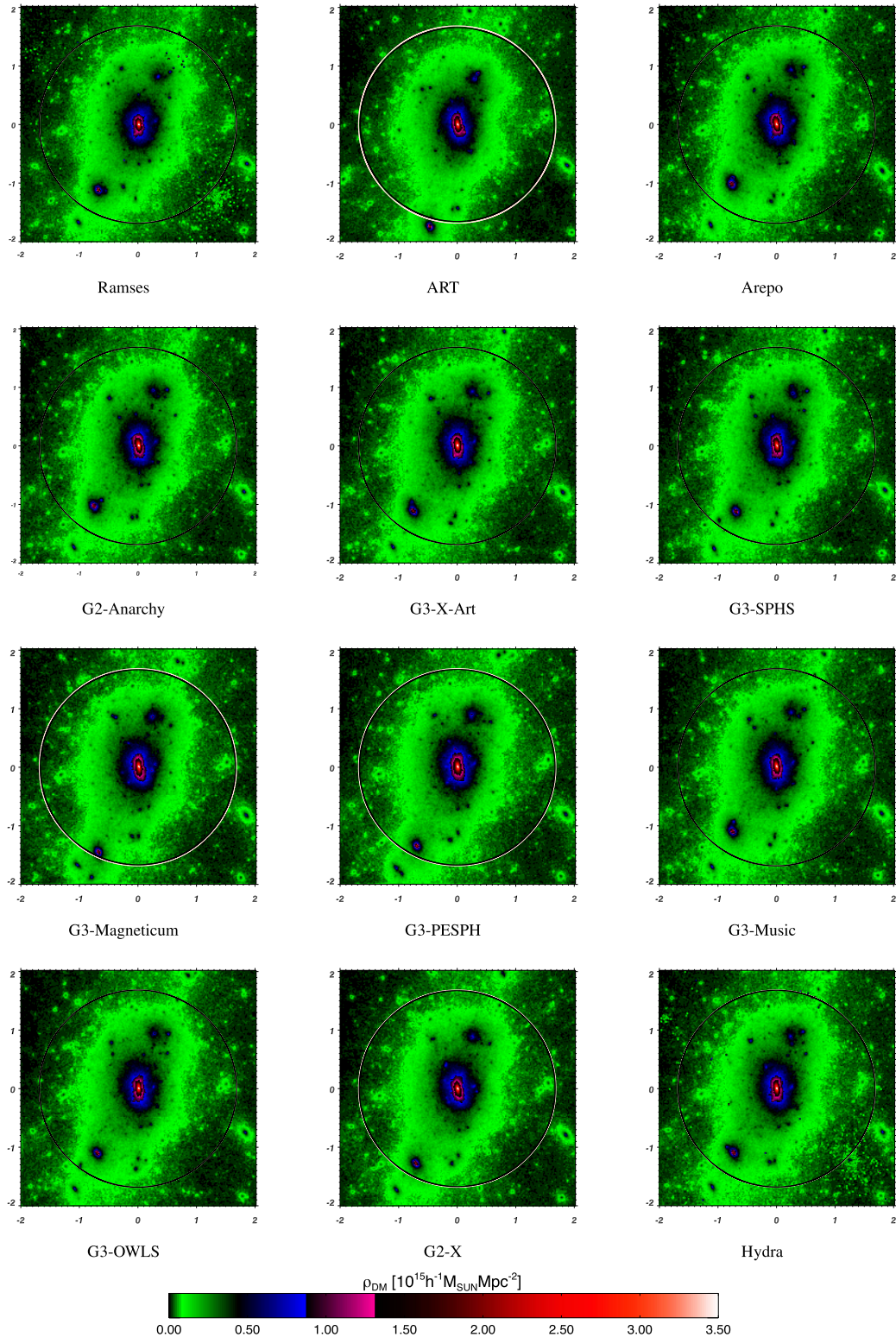
Figure 1. Projected density of the dark matter halo at $z = 0$ for each simulation as indicated. The box is $2\,h^{-1}$ Mpc on a side. The white circle indicates $M_{200}^{\rm crit}$ for the halo, the black circle shows the same but for the G3-MUSIC simulation.

described in the next section, we aligned all the codes to a common value of the size of the base level (512^3).

Fig. 2 displays the radially averaged projected dark matter density profiles for the 12 different non-aligned dark matter-only runs,

including also the residuals relative to the density profile of the reference G3-MUSIC simulation. All the simulations lie well within 10 per cent of each other at all radii with the HYDRA simulation being indistinguishable from the GADGET runs.
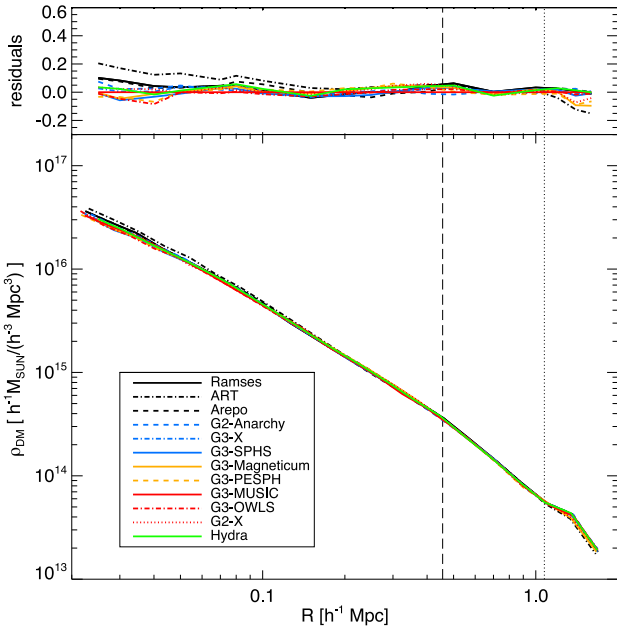
Figure 2. Radial density profiles for the dark matter-only simulations at $z = 0$ (bottom panel) and the difference between the radial density profiles of each dark matter-only simulation at $z = 0$ and the reference G3-MUSIC density profile (top panel). The vertical dashed line corresponds to R_{2500} and the vertical dotted line to R_{500} in the reference G3-MUSIC simulation. The radial profiles finish at R_{200}.
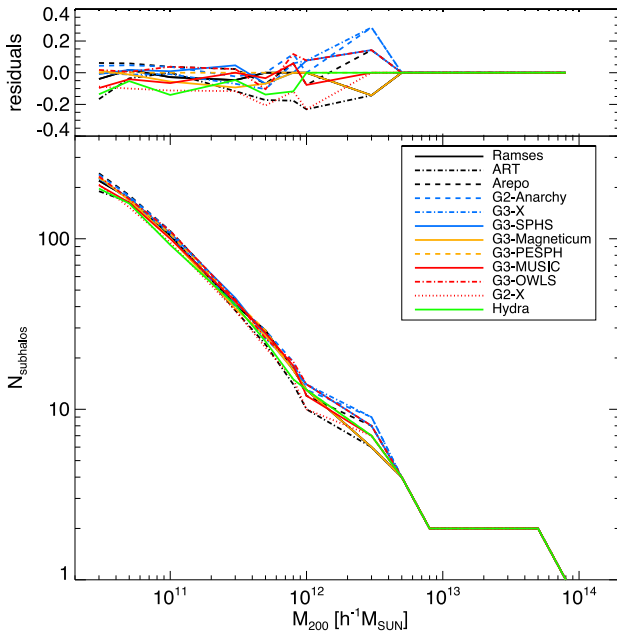
Figure 3. Subhalo mass function for the dark matter-only simulations at $z = 0$ (bottom panel) and difference between the subhalo mass function of each dark matter-only simulation at $z = 0$ and the reference G3-MUSIC subhalo count (top panel). The lines all overlay each other at $M_{200} > \times 10^{12} \, h^{-1} \, \mathrm{M}_\odot$.

The subhalo mass function at $z = 0$ (Fig. 3) is recovered with only small differences between the models (the differences are always below 20 per cent at all masses). Subhaloes were identified using AHF (Gill, Knebe & Gibson 2004; Knollmann & Knebe 2009; freely available from http://popia.ft.uam.es/AHF). The number of

subhaloes is essentially identical except for tiny mass differences which are driven by the small divergences in radial location that were identified above. These code-to-code variations lead to differences in the mass associated with each subhalo as the threshold that defines where a subhalo is separated from the background halo varies. As expected, subhaloes closer to the centre of the main halo than the equivalent subhalo in one of the other models have a lower recovered mass.

Comparison of the dark matter distribution and of its radial density profile at $z = 1$ yields results similar to those described above. We conclude that the typical range of chosen parameters for cosmological simulation codes found in the literature introduces a variation of around 10 per cent in the density profile of collapsed objects. This scatter can be reduced to less than 5 per cent by aligning the numerical accuracy of our calculations (see Fig. A1 in Appendix A). Although this is not essential for many applications, we choose to do this in the remainder of this paper so that the underlying dark matter framework agrees closely, allowing us to focus on differences resulting from the various hydrodynamical schemes. Given these concerns we adopt for the non-radiative runs a common set of parameters, including a base-level mesh fixed to 512^3 (the full set of aligned parameters is listed in Table A1).

5 NON-RADIATIVE RUN COMPARISON

We now proceed to include a gas phase into our calculations. We repeat the simulation of the same cluster as used in Section 4 including gas which we do not allow to radiate energy. If radiation were allowed many additional physical processes are required to produce a realistic simulation. While such full physics models exist (indeed we have already completed them) their study is beyond the scope of this initial project and forms the basis of a subsequent paper. Fig. 4 shows some of the global properties of the selected cluster as calculated by the different codes used in this work: radius, mass, mass-weighted temperature, gas fraction, the velocity dispersion of the dark matter and axial ratios. All these quantities have been calculated at an aperture radius corresponding to R_{200}^{crit}, the radius enclosing an overdensity of 200 relative to the critical density, defined as

$$\rho_c(z) = \frac{3 H_0^2 E(z)^2}{8 \pi G}, \qquad (2)$$

where H_0 is the present value of the Hubble constant and G is the universal gravitational constant. Using the same definition we refer in the text to R_{2500}^{crit} and R_{500}^{crit} as the radii enclosing an overdensity of 2500 and 500 relative to $\rho_c(z)$. It is interesting to note that all the codes were able to recover the same values for the different properties of the halo with a scatter smaller than 1 per cent for mass, radius, axial ratio and dark matter velocity dispersion and around 2 per cent for baryon fraction and gas temperature. The same properties were estimated with a scatter between 5 and 10 per cent in SB99, with differences of up to 30 per cent between the maximum and minimum values: in our comparison the same difference is always below 5 per cent (except for the gas fraction where we register a disagreement of 8 per cent between the maximum and minimum values). Of course SB99 used a wider range of codes and the discrepancies were also partially due to the poor agreement between the underlying dark matter frameworks which we have carefully aligned first.

Thumbnail images of the gas density distribution for each of the methods at $z = 0$ are given in Fig. 5. We see a dramatic variation
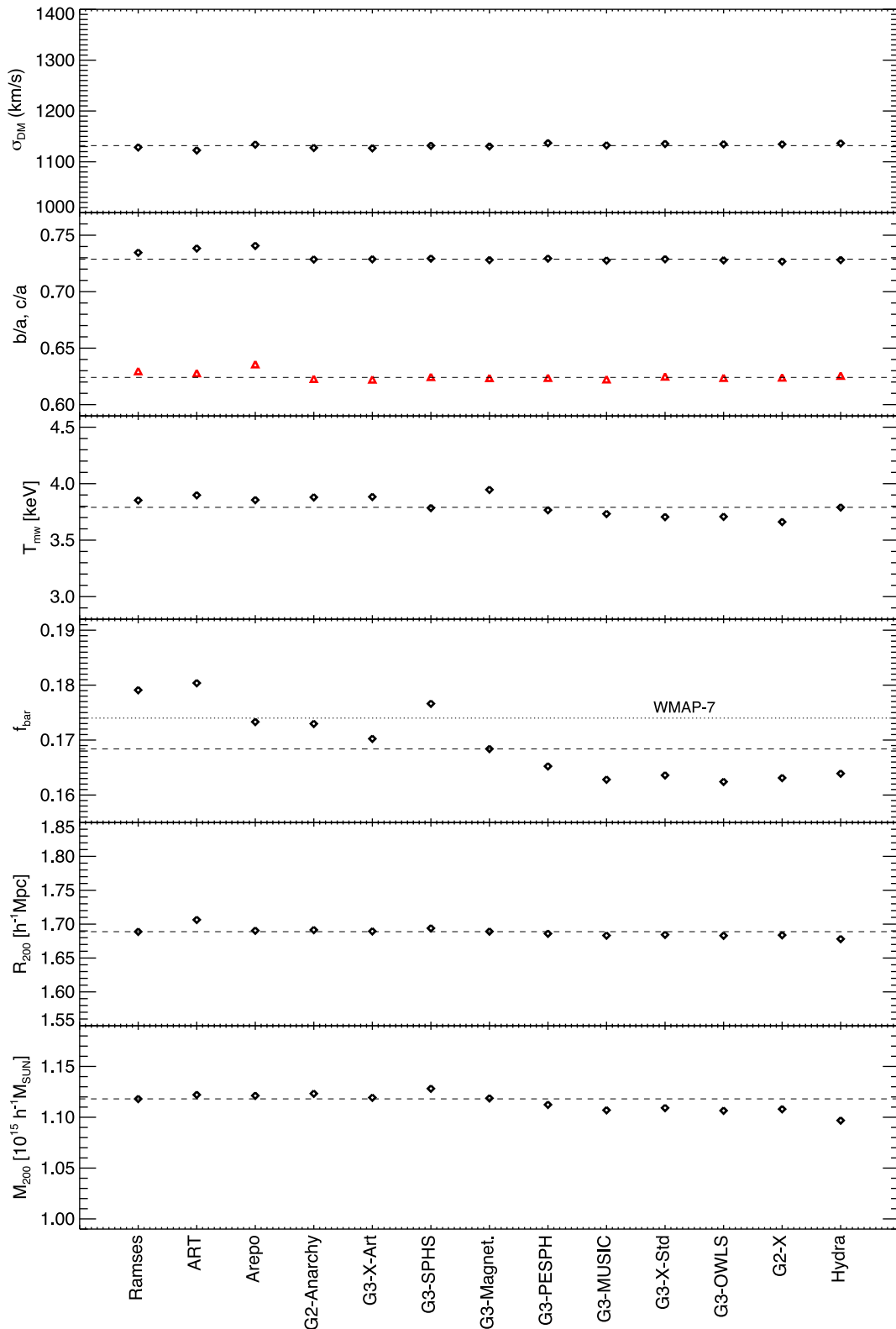
Figure 4. Global properties of the cluster calculated by the different codes. All quantities are computed within R_{200}^{crit}. From top panel to bottom panel: (1) the one-dimensional velocity dispersion of the dark matter; (2) the axial ratio (b/a in black, c/a in red); (3) the mass-weighted temperature; (4) the gas fraction (the dotted line represents the value of the cosmic ratio from *WMAP*7 (Komatsu et al. 2011); (5) the radius and (6) the total cluster mass. The dashed lines represent the median value for each one of the plotted quantities.
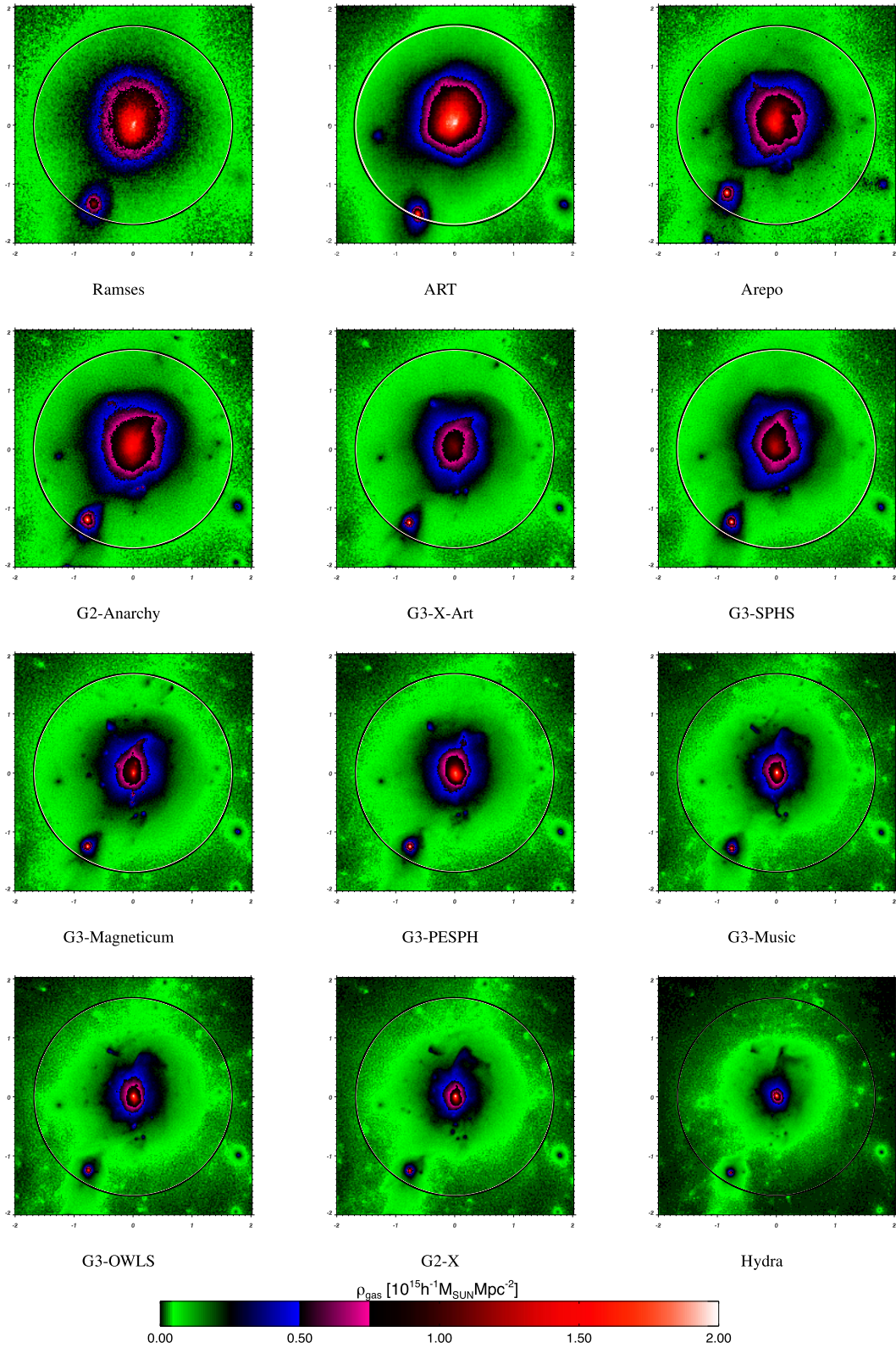
Figure 5. Projected density of the gas halo at $z = 0$ for each simulation as indicated. The box is 2 h^{-1} Mpc on a side. The white circle indicates M_{200}^{crit} for the halo, the black circle shows the same but for the G3-MUSIC simulation.
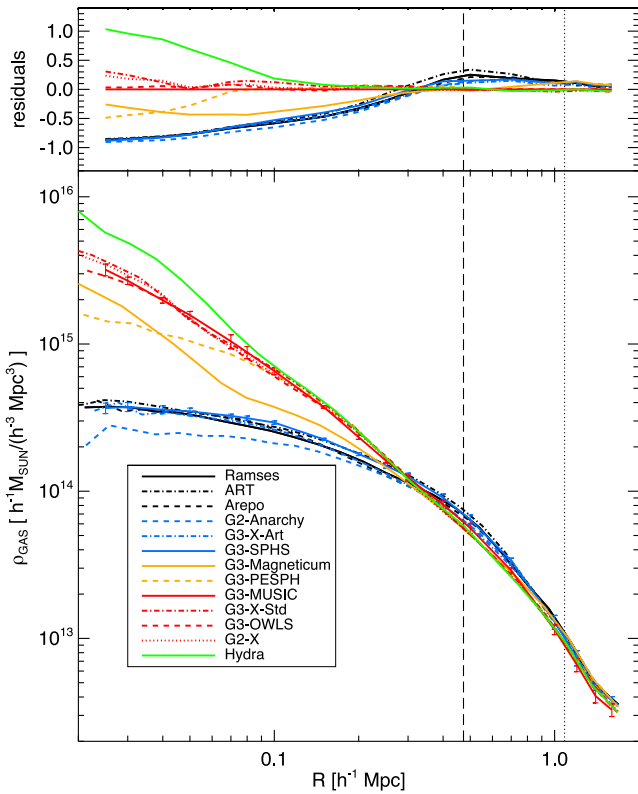
Figure 6. Radial gas density profiles at $z = 0$ for each simulation as indicated (bottom panel) and difference in radial gas density profiles at $z = 0$ between each simulation and the reference G3-MUSIC simulation. The vertical dashed line corresponds to R_{2500} and the dotted line to R_{500} of the reference G3-MUSIC values. The error bars on G3-SPHS (black) and G3-MUSIC (red) are calculated from the scatter between snapshots averaged over the final 0.27 Gyr. The data are cut-off when the radius goes below the softening scale of the code at the inside and at R_{200} at the outside.

in the central concentration of the gas, with some methods having significantly larger extended nuclear gas regions.

This trend is born out by the radial gas density profiles given in Fig. 6. We see that the radial gas density contains a significant core for RAMSES, ART and AREPO when compared to the classic SPH schemes employed by some SPH codes such as HYDRA and G2-X. SPH codes that implement artificial diffusion can produce results that are close to those of RAMSES and AREPO. In between these two extremes the various SPH implementations can produce a range of core sizes in the radial gas density profile.

Similarly to all subsequent radial plots the differences in the gas density compared to the fiducial G3-MUSIC simulation are shown in the top panel of Fig. 6. At $z = 0$ the lowest central densities are an order of magnitude lower than in the G3-MUSIC simulation while the highest central densities are around two to three times larger, i.e. the variation in the central gas density across our runs is well over an order of magnitude. The scatter becomes more moderate when considering the outer region of the cluster, not exceeding 20 per cent at radii larger than $R_{2500}^{\rm crit}$.

We next show the radial temperature profiles for all the simulations in Fig. 7. We use the mass-weighted temperature, defined as

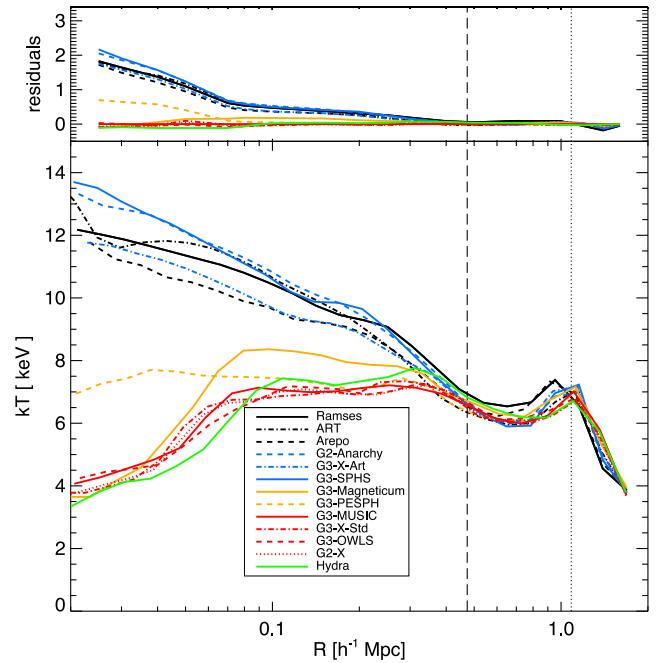$$T_{\rm mw} = \frac{\sum_i T_i m_i}{\sum_i m_i}, \qquad (3)$$

Figure 7. Radial temperature profile at $z = 0$ for each simulation as indicated and difference between each simulation and the reference G3-MUSIC simulation. The vertical dashed line corresponds to R_{2500} and the vertical dotted line to R_{500} of the reference G3-MUSIC values. The lines terminate at R_{200} for each model.

where m_i and T_i are the mass and the electronic temperature of the gas particles, which are assumed to be in thermodynamic equilibrium. The central temperatures vary by more than a factor of 3, with a group of methods displaying a central temperature inversion with inner temperatures around half the peak value of 7–8 keV. In contrast, some codes display a monotonically rising temperature profile as the radius falls with a peak temperature up to 14 keV at the very centre. This group of codes consists of those with the most extended cores in the radial gas density profiles.

At radii larger than $R_{2500}^{\rm crit}$ the scatter is significantly more moderate, and the residuals appear to be a factor of 2 smaller than in fig. 17 of SB99 in the same cluster regions.

Finally we show the radial gas entropy profiles for all the codes in Fig. 8. We adopt the definition of entropy commonly used in the literature, particularly with reference to X-ray observations:

$$S(R) = \frac{T_{\rm gas}(R)}{n_{\rm e}^{2/3}(R)}, \qquad (4)$$

where $n_{\rm e}$ is the number density of free electrons. This clearly displays the now traditional split between grid-based codes and classic SPH methods, such as HYDRA and G2-X, which show a falling inner entropy as the radius is decreased all the way into the very centre. This is completely consistent with the inner temperature inversion and high central gas density. Conversely, the grid-based codes such as RAMSES, ART and AREPO display the well-known flat inner entropy cores that result from rising inner temperature profiles and extended gas densities. However, we see that in between these extremes we have a range of entropy profiles that depend on the specific SPH implementation employed. We note that modern, sophisticated SPH codes which employ mixing are now capable of recovering entropy profiles that lie anywhere between the coreless classic SPH
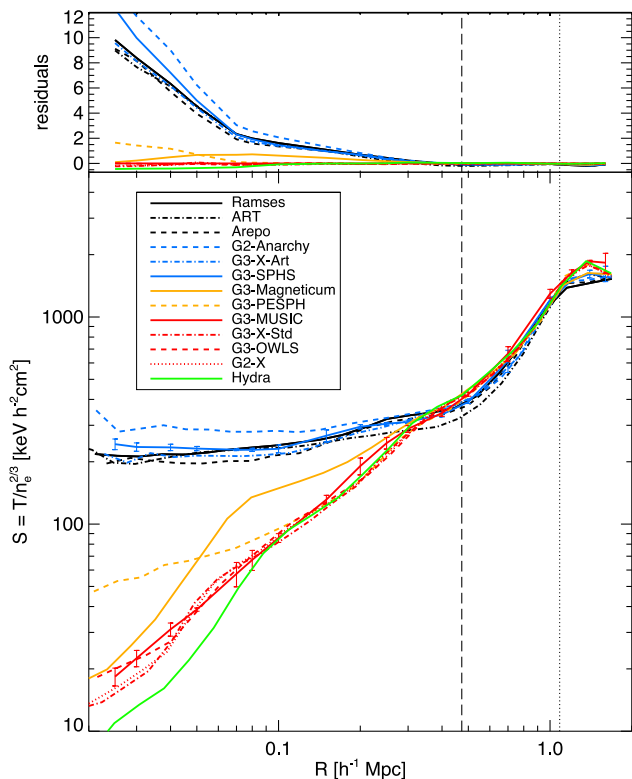
Figure 8. Radial entropy profile at $z = 0$ (bottom panel) for each simulation as indicated and difference between each simulation and the reference G3-MUSIC simulation (top panel). The dashed line corresponds to R_{2500} and the dotted line to R_{500} of the reference G3-MUSIC values. The error bars on G3-SPHS (blue) and G3-MUSIC (red) are calculated from the scatter between snapshots averaged over the last 0.27 Gyr.

schemes and the cored profiles of the grid-based approaches depending upon the precise nature of the scheme and the amount of mixing employed. We highlight that modern SPH codes such as G3-SPHS, G2-ANARCHY and G3-X-ART are able to recover the same flat entropy core observed for grid-based codes, with a scatter smaller than 20 per cent, even in the inner cluster regions. G3-PESPH and G3-MAGNETICUM, which have an artificial viscosity switch but different artificial conductivity with respect to the other modern SPH codes, show an intermediate behaviour between classic and modern SPH codes.

5.1 Other quantities in the non-radiative simulations

It is important to note that the differences in radial gas density, temperature and entropy evidenced above are not driven by code issues such as poor thermalization or large-scale flows. In Fig. 9 we show the ratio of gas thermal, U, to kinetic energy, K (relative to the centre of momentum of the cluster), at $z = 0$:

$$\eta = \frac{2K}{|U|}. \tag{5}$$

All the methods agree closely on the value of η as a function of halo radius and none displays any evidence of poor thermalization. ART is the only code showing some moderate discrepancy from the others. The scatter is always below 20 per cent.

Given our radial dark matter density and gas density profiles we can also calculate the radial gas fraction for all the methods.
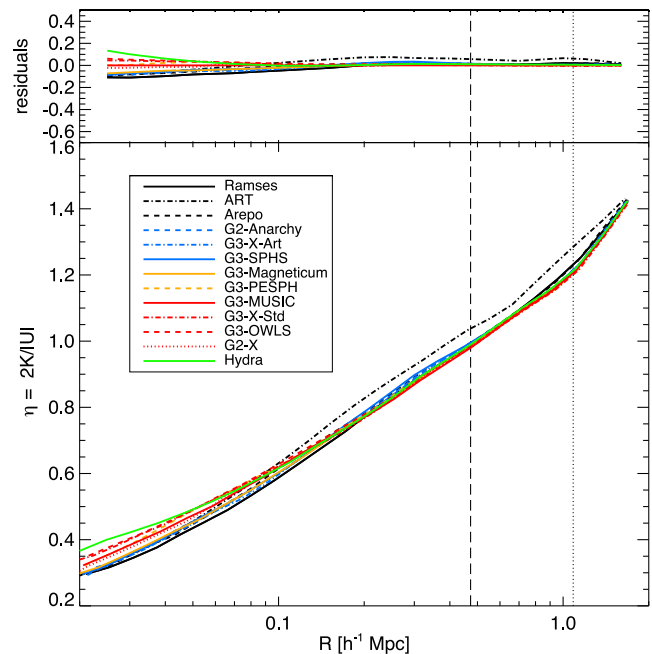
Figure 9. The ratio of kinetic to thermal energy in the gas, η, measured radially at $z = 0$ for each simulation as indicated (bottom) and difference between each simulation and the reference G3-MUSIC simulation (top). The dashed line corresponds to R_{2500} and the dotted line to R_{500} of the reference G3-MUSIC values.
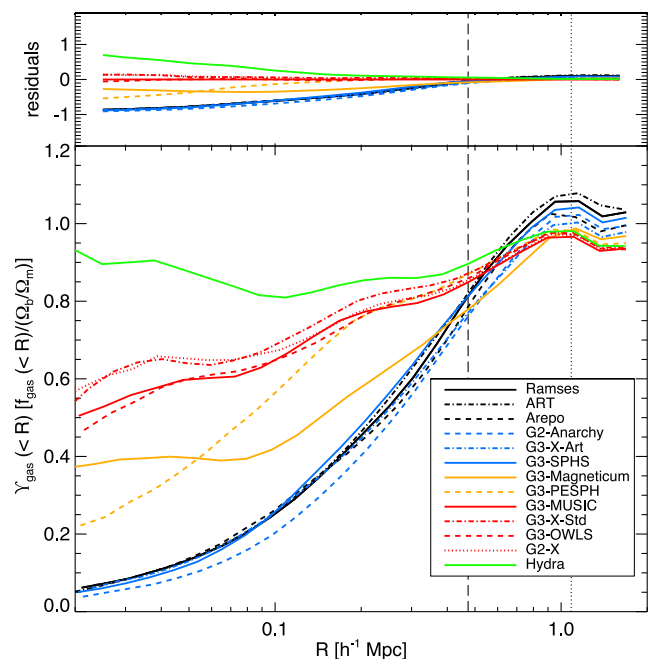

Figure 10. Radial gas fraction at $z = 0$ relative to the cosmic value for each simulation as indicated (bottom) and difference between each simulation and the reference G3-MUSIC simulation (top). The dashed vertical line corresponds to R_{2500} and dotted vertical line to R_{500} of the reference G3-MUSIC values.

In Fig. 10 we show the radial profiles of the depletion factor Υ, defined as

$$\Upsilon = \frac{M_{\text{gas}}(< R)}{M(< R)} \left(\frac{\Omega_{\text{b}}}{\Omega_{\text{m}}} \right)^{-1}. \tag{6}$$
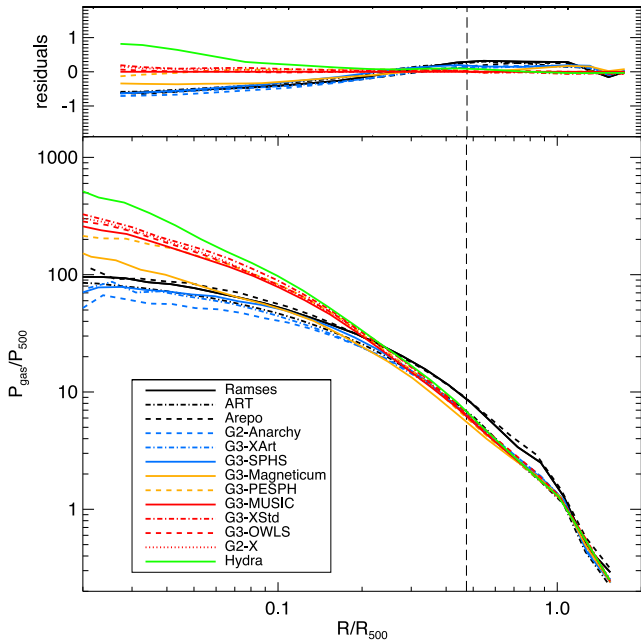
Figure 11. Radial gas pressure at $z = 0$ measured in each simulation as indicated (bottom panel) and difference between each simulation and the reference G3-MUSIC simulation (top panel). The pressure, as well as the radius, has been normalized to the corresponding value at R_{500} for each code. The dashed vertical line corresponds to R_{2500} of the reference G3-MUSIC value.

Figure 12. Radial X-ray luminosity profiles at $z = 0$ for each simulation as indicated (bottom panel) and difference between each simulation and the reference G3-MUSIC simulation (top panel). The dashed vertical line corresponds to R_{2500} and the dotted vertical line to R_{500} of the reference G3-MUSIC values.

The results reflect the existence of methods that produce very centrally concentrated gas and methods yielding an extended core with much higher average entropy. For those codes with an extended entropy core the gas fraction falls significantly with decreasing radius and can reach values as low as 5 per cent of the universal baryon fraction for these non-radiative simulations. Within 100 h^{-1} kpc it can fall below 25 per cent of the cosmic gas fraction in both grid-based and modern SPH codes. The differences in the radial gas fraction reflect those already evidenced in the gas density profiles and warn against using a universal calibration of the baryon depletion at $R_{2500}^{\rm crit}$ based on simulations in cosmological applications of the cluster baryon fraction, especially when using only non-radiative gas. We expect these results to be very different (and closer to observational results) when radiative physics is included. The scatter in the outer regions of the cluster appears to be more limited (less than 20 per cent) with respect to the results shown in fig. 13 of SB99, where differences of up to 50 per cent between the different codes where registered even close to the virial radius.

We can combine our measurements of the gas density and temperature to produce Fig. 11, the gas pressure profiles and Fig. 12, the X-ray emission profiles. We define the pressure as $P = \rho_{\rm gas}T$ and we normalize the profiles to the value of P_{500} (the value of the pressure as calculated at $R_{500}^{\rm crit}$) in order to be consistent with the definition of universal pressure profile introduced by Arnaud et al. (2010).

The X-ray luminosity profile is defined as $4\pi R^3 \mathcal{L}_{\rm X}$ and we approximate the X-ray luminosity density as $\mathcal{L}_{\rm X} \propto \rho_{\rm gas}^2 T^{1/2}$. The variation in the gas density and temperature produce very different pressure and X-ray emission profiles.

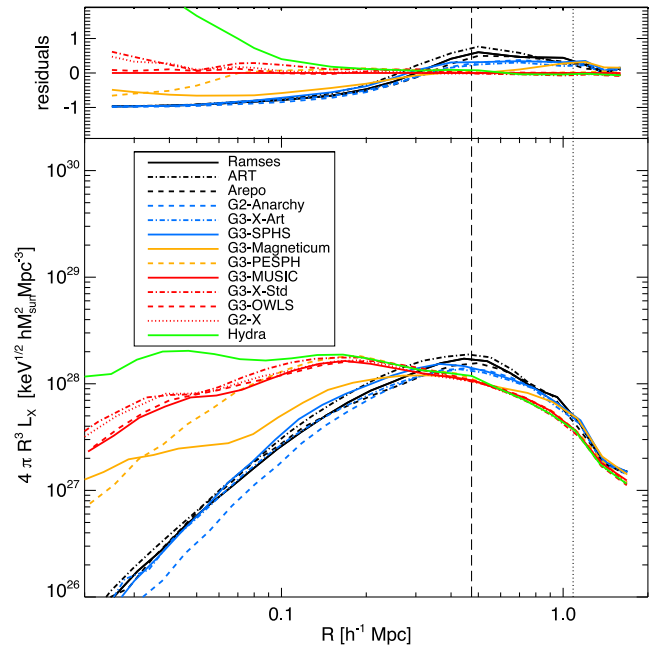As expected, the pressure profiles continue to rise all the way into the centre for all codes (as the central gas is close to hydrostatic equilibrium in all cases). Because of the very high central density, the central X-ray emission for HYDRA is over two orders of magnitude higher than that found by the grid-based codes and some modern SPH methods which form extended cores.

6 CONVERGENCE AND SCATTER BETWEEN SIMULATIONS

6.1 Dark matter-only runs

We began with an essentially blinded simulation of the same dark matter cluster. By doing this we discovered that the major cause of the discrepancy (for the GADGET-based codes at least) between the methods is the size of the base-level particle mesh. Between this and the refined region that covers the cluster there is a resolution jump, typically to an effective resolution of at least 2048³. A base mesh of 256³ is not sufficient to resolve the interface region between the low-resolution region (i.e. the base level) and the high-resolution region (i.e. the refined level) placed over the cluster; we found that a base-level of at least 512³ is required to ensure that the dark matter component is well aligned across the different codes. We then resimulated the cluster with an agreed set of matching parameters. For the *N*-body only simulations we find excellent agreement between the density profiles of the main halo and the statistics of their subhaloes (Figs 2 and 3), once the input parameters of the various runs are matched (see Appendix A). The numerous implementations of GADGET, RAMSES and HYDRA agree remarkably well, although the GADGET agreement is perhaps not surprising given that the gravity solver is in each case built on the same foundation.

6.2 Variation amongst classic SPH methods

Despite the excellent agreement between the dark matter-only runs, the classic SPH simulations also show scatter in their central gas density, temperature and entropy that is significantly larger than the theoretical error on each (as calculated from the scatter between snapshots averaged over the final 0.27 Gyr; see the error bars marked on the plots shown in Figs 6 and 8). This is likely because in classic SPH entropy mixing is extremely low and so the final entropy profile forms a historical record reflecting the result of all the mergers that formed the cluster. The resultant radial profile is well sorted in entropy as low entropy material sinks and high entropy material rises. Small differences between the codes in their ability to resolve the entropy jumps due to accretion shocks are therefore amplified by the final time (Power et al. 2014). Nevertheless, a similar scatter is observed also for the gas density profiles of modern SPH codes, suggesting that the lack of mixing may not be the only cause of the discrepancies. We recall that the behaviour of low entropy particles in the real ICM, which depends on how the different phases mix, is largely unknown.

6.3 Variation between modern SPH methods and grid-based schemes

Of considerable interest are the differences between the modern SPH flavours (G3-X-ART; G3-SPHS; G2-ANARCHY; G3-PESPH; G3-MAGNETICUM) and the grid-based schemes. There is excellent agreement between the dark matter distributions across the runs, which allows us to isolate the effect of the hydrodynamics solver (Fig. A2). This agreement for the dark matter is somewhat overemphasized by the plots because all the GADGET methods and AREPO use the same underlying gravity solver. However, RAMSES and HYDRA are not related to the GADGET family and still look remarkably similar for the dark matter (see Fig. A1 for example).

For the radial gas density profiles shown in Fig. 6, RAMSES, AREPO and G3-X-ART are in excellent agreement with one another and G3-SPHS and ART only show some moderate discrepancies, while G2-ANARCHY seems to be an outlier. In Fig. 7 the temperature profiles for AREPO and G3-X-ART are very closely matched but look to have an offset from G3-SPHS and G2-ANARCHY, which are very close to one other. RAMSES produces a result that is intermediate between these two pairs. ART produces a temperature lower than the other two grid-based codes outside of the cluster centre. In Fig. 8 the radial entropy profiles of G3-SPHS, G3-X-ART and RAMSES are again very close to one another while G2-ANARCHY is somewhat higher and AREPO somewhat lower. ART show a lower entropy in the region around R_{500}. The origin of these differences is yet to be explained, although we note that it cannot be attributed to different merger phases and must result from the hydrodynamics solver. In the case of G2-ANARCHY a possible cause of discrepancy may be the choice of the kernel (using a C4 kernel with 200 neighbours gives slightly different values for the central entropy). Power et al. (2014) showed that at the resolution of the simulations in this paper, G3-SPHS is numerically converged. It would be interesting to see if the remaining differences between G3-SPHS, G3-X-ART, AREPO, RAMSES, ART and G2-ANARCHY remain with increasing resolution. We defer such a resolution study to future work, whose intent will be to narrow down these numerical uncertainties.

The AREPO simulations use the total energy as a conserved quantity in the Godunov scheme, which is the usual choice in finite volume codes and the same choice as that used in the most recent AREPO studies (as e.g. in Vogelsberger et al. 2014). As discussed

in Springel (2010), using an entropy-energy formalism results in smaller entropy cores and higher central gas densities somewhat closer to classic SPH results (similar results are also shown by GIZMO; Hopkins 2015). Following most recent AREPO studies, we have not employed the latter method here due to concerns regarding the artificial suppression of weak shocks and the potentially less accurate energy conservation.

Finally, we note that the results of G3-PESPH are very different from those of the other modern SPH flavours (with the exception of G3-MAGNETICUM), and are more similar to those of classic SPH simulations. A key difference is that this version of G3-PESPH does not include any explicit conductivity or mixing, while all the other modern variants do. Hu et al. (2014) showed that G3-PESPH performs much better than previous versions of SPH for surface instabilities by greatly mitigating surface tension problems, but in areas of very strong shocks ($M \sim 1000$) adding artificial conduction provides a better match to analytic solutions. Insights into the behaviour of G3-PESPH may be gained by considering the SPH method presented in Read et al. (2010), and the earlier multiphase method by Ritchie & Thomas (2001). As pointed out by Read & Hayfield (2012), the Richie & Thomas method only works well for relatively small entropy contrasts between different fluid phases. As the entropy contrast becomes very large, the admixture of low and high entropy particles within the smoothing kernel creates large pressure fluctuations that prevent mixing, as in classic SPH. This was recognized also by Ritchie & Thomas (2001) who proposed scaling the neighbour number with the entropy contrast to combat this. However, this rapidly becomes prohibitively expensive in real-world applications, which led Read & Hayfield (2012) to abandon such methods in favour of dissipation switching, as proposed by Price (2008); such switching is common to G3-X-ART, G3-SPHS and G2-ANARCHY and helps to explain their similarity. The discrepancy between G3-PESPH and the other modern SPH flavours may also reflect the treatment of artificial viscosity. They all adopt the artificial viscosity model suggested by Cullen & Dehnen (2010) which can produce larger entropy gains in shocks while the authors of G3-PESPH do not use this because it also seems to add substantial entropy into very diffuse intergalactic gas that may be spurious (Katz, private communication). In short, it is unclear how much artificial conductivity and/or mixing is appropriate in SPH, or even whether the mesh-based codes are providing the correct solution that all SPH codes should be trying to emulate. Nonetheless, consistency with mesh codes appears to require modern SPH using conduction, mixing and a higher order dissipation switch.

7 SUMMARY AND CONCLUSIONS

We have investigated the performance of 13 modern astrophysical simulation codes – RAMSES, ART, AREPO, HYDRA and nine versions of GADGET with different SPH implementations – by carrying out cosmological zoom simulations of a single massive galaxy cluster. Our goal was to assess the consistency of the different codes in reproducing the spatial and thermodynamical structure of the dark matter and non-radiative gas in a galaxy cluster.

As our initial step, we ran dark matter-only versions of the simulations with each simulation team using their preferred set of numerical parameters (e.g. time step accuracy, gravitational softening, dimension of the particle mesh), and examined the spherically averaged mass density profile and the spatial distribution of substructures. We found good consistency between the density profiles recovered by the codes at approximately the 10 per cent level, while there were small variations in the positions of

substructures. When these simulations were re-run with an agreed common set of numerical parameters, we found that these small variations could be suppressed (essentially entirely, in the case of the GADGET codes).

By adopting this common parameter set, we were able to isolate those differences between the results of the hydrodynamical simulations that arise only from the choice of hydrodynamical solver, rather than from the complex interplay of the hydrodynamical and gravity solvers. As expected from the wide range of hydrodynamical implementations covered we found that the resulting gas density profiles varied substantially amongst the codes. Our key findings can be summarized as follows.

(i) Some codes, essentially the oldest, with classic SPH implementations, exhibit continually falling inner entropy profiles, without any evidence of an entropy core. This is because these codes, particularly HYDRA, were carefully designed to be entropy conserving with very low levels of mixing. This lack of mixing preserves low-entropy gas particles at the centres of all objects, including subhaloes, which survive until late times. As the cluster relaxes, these particles sink to the centre of the radial density profile, decreasing the central entropy.

(ii) In contrast, the grid-based codes RAMSES, ART and AREPO produce extended cores with a large constant entropy core. In these mesh-based codes mixing of entropy arises as a consequence of the numerical diffusion associated with the Riemann solver: they naturally mix entropy between gas elements, essentially eliminating the very low entropy material.

(iii) Modern SPH codes such as G2-ANARCHY, G3-SPHS and G3-X-ART, which have dissipative switches and new kernels, can bridge the gap between the classic SPH codes and grid-based codes, producing entropy cores that are indistinguishable from those of the grid-based schemes.

Our results confirm that the discrepancies between grid-based codes and SPH codes in describing the radial entropy profile of simulated clusters, identified by the Santa Barbara Cluster Comparison Project presented in SB99, can be overcome by modern SPH codes. Importantly, all the codes employed in this work succeed in recovering the global properties and most of the radial profiles of a simulated large galaxy cluster with much greater accuracy and significantly smaller scatter than those presented in SB99; this highlights the enormous strides in the development of astrophysical hydrodynamical simulation codes over the last decade.

This work constitutes the first in a series of papers in which we examine in detail the predictions of modern astrophysical hydrodynamical simulation codes. The next paper in this series will focus on simulations of the same galaxy cluster, now modelled with a variety of galaxy formation processes including cooling, star formation, supernovae and feedback from AGN. This will allow us to establish how radiative processes affect the entropy cores of simulated clusters. Subsequent papers will look at the recovery of cluster properties such as X-ray temperature and Sunyaev–Zel'dovich profiles, gravitational lensing and cluster outskirts and hydrostatic-mass bias; all of which will add to our understanding of how consistently the results of different codes can inform our understanding of galaxy cluster properties.

---

[4] http://music.ft.uam.es

This research has made use of NASA's Astrophysics Data System (ADS) and the arXiv preprint server.

REFERENCES

Aarseth S. J., Turner E. L., Gott J. R., III, 1979, ApJ, 228, 664

Agertz O. et al., 2007, MNRAS, 380, 963

Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409

Arnaud M., Pratt G. W., Piffaretti R., Böhringer H., Croston J. H., Pointecouteau E., 2010, A&A, 517, A92

Balsara D. S., 1995, J. Comput. Phys., 121, 357

Beck A. M. et al., 2016, MNRAS, 455, 2110

Biffi V., Valdarnini R., 2015, MNRAS, 446, 2802

Biffi V., Sembolini F., De Petris M., Valdarnini R., Yepes G., Gottlöber S., 2014, MNRAS, 439, 588

Booth C. M., Schaye J., 2009, MNRAS, 398, 53

Borgani S., Kravtsov A., 2011, Adv. Sci. Lett., 4, 204

Borgani S. et al., 2004, MNRAS, 348, 1078

Colella P., Glaz H. M., 1985, J. Comput. Phys., 59, 264

Couchman H. M. P., Thomas P. A., Pearce F. R., 1995, ApJ, 452, 797

Cullen L., Dehnen W., 2010, MNRAS, 408, 669

Dalla Vecchia C., Schaye J., 2012, MNRAS, 426, 140

Davé R., Katz N., Weinberg D. H., 2002, ApJ, 579, 23

Davé R., Oppenheimer B. D., Sivanandam S., 2008, MNRAS, 391, 110

Dehnen W., Aly H., 2012, MNRAS, 425, 1068

Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604

Dolag K., Jubelgas M., Springel V., Borgani S., Rasia E., 2004, ApJ, 606, L97

Dolag K., Vazza F., Brunetti G., Tormen G., 2005, MNRAS, 364, 753

Durier F., Dalla Vecchia C., 2012, MNRAS, 419, 465

Fall S. M., 1978, MNRAS, 185, 165

Frenk C. S. et al., 1999, ApJ, 525, 554 (SB99)

Gill S. P. D., Knebe A., Gibson B. K., 2004, MNRAS, 351, 399

Gingold R. A., Monaghan J. J., 1977, MNRAS, 181, 375

Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, MNRAS, 442, 2304

Hopkins P. F., 2013, MNRAS, 428, 2840

Hopkins P. F., 2015, MNRAS, 450, 53

Hu C.-Y., Naab T., Walch S., Moster B. P., Oser L., 2014, MNRAS, 443, 1173

Kay S. T., Thomas P. A., Jenkins A., Pearce F. R., 2004, MNRAS, 355, 1091

Klypin A., Kravtsov A. V., Bullock J. S., Primack J. R., 2001, ApJ, 554, 903

Knollmann S. R., Knebe A., 2009, ApJS, 182, 608

Komatsu E. et al., 2011, ApJS, 192, 18

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

Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, ApJS, 111, 73

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

Lucy L. B., 1977, AJ, 82, 1013

Maier A., Iapichino L., Schmidt W., Niemeyer J. C., 2009, ApJ, 707, 40

Metzler C. A., Evrard A. E., 1994, ApJ, 437, 564

Mitchell N. L., McCarthy I. G., Bower R. G., Theuns T., Crain R. A., 2009, MNRAS, 395, 180

Monaghan J. J., 1997, J. Comput. Phys., 136, 298

Monaghan J. J., Lattanzio J. C., 1985, A&A, 149, 135

Morris J. P., Monaghan J. J., 1997, J. Comput. Phys., 136, 41

Muanwong O., Thomas P. A., Kay S. T., Pearce F. R., Couchman H. M. P., 2001, ApJ, 552, L27

Nagai D., Kravtsov A. V., Vikhlinin A., 2007, ApJ, 668, 1

O'Shea B. W., Nagamine K., Springel V., Hernquist L., Norman M. L., 2005, ApJS, 160, 1

Pakmor R., Edelmann P., Röpke F. K., Hillebrandt W., 2012, MNRAS, 424, 2222

Pearce F. R., Thomas P. A., Couchman H. M. P., Edge A. C., 2000, MNRAS, 317, 1029

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

Power C., Read J. I., Hobbs A., 2014, MNRAS, 440, 3243

Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018

Price D. J., 2008, J. Comput. Phys., 227, 10040

Price D. J., 2012, J. Comput. Phys., 231, 759

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

Read J. I., Hayfield T., 2012, MNRAS, 422, 3037

Read J. I., Hayfield T., Agertz O., 2010, MNRAS, 405, 1513

Ritchie B. W., Thomas P. A., 2001, MNRAS, 323, 743

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

Saitoh T. R., Makino J., 2013, ApJ, 768, 44

Schaye J., Dalla Vecchia C., 2008, MNRAS, 383, 1210

Schaye J. et al., 2010, MNRAS, 402, 1536

Sembolini F., Yepes G., De Petris M., Gottlöber S., Lamagna L., Comis B., 2013, MNRAS, 429, 323

Sembolini F., De Petris M., Yepes G., Foschi E., Lamagna L., Gottlöber S., 2014, MNRAS, 440, 3520

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

Sijacki D., Pfrommer C., Springel V., Enßlin T. A., 2008, MNRAS, 387, 1403

Sijacki D., Vogelsberger M., Kereš D., Springel V., Hernquist L., 2012, MNRAS, 424, 2999

Springel V., 2005, MNRAS, 364, 1105

Springel V., 2010, MNRAS, 401, 791

Springel V., Hernquist L., 2002, MNRAS, 333, 649

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

Teyssier R., 2002, A&A, 385, 337

Thacker R. J., Couchman H. M. P., 2006, Comput. Phys. Commun., 174, 540

Thacker R. J., Scannapieco E., Couchman H. M. P., 2006, ApJ, 653, 86

Tricco T. S., Price D. J., 2013, MNRAS, 436, 2810

Valdarnini R., 2012, A&A, 546, A45

Vogelsberger M. et al., 2014, Nature, 509, 177

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

Wadsley J. W., Veeravalli G., Couchman H. M. P., 2008, MNRAS, 387, 427

Wendland H., 1995, Adv. Comput. Math., 4, 389

White S. D. M., 1976, MNRAS, 177, 717

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

APPENDIX A: DARK MATTER ALIGNMENT

In order to perform a clean comparison of the various gas physics implementations we carefully aligned the underlying gravitational framework for each model. While Fig. 1 illustrates the range of outcomes that result in the dark matter-only runs from a blind comparison using individual parameter choices, we can choose a common parameter set for those quantities that control the accuracy of the gravitational forces. For instance, for GADGET, Table A1 gives the parameter choices made independently by each group. The final row lists the common parameter set adopted for the non-radiative comparison. The improvement in the alignment between the different methods is shown in Fig. A1, which displays the dark matter radial profiles of the dark matter-only simulations re-run using the common parameters listed in Table A1. After fixing the base-level mesh to the common value of 512^3, the scatter in the profiles among different codes is below 10 per cent. We tested that increasing the resolution of the base-level mesh further to 1024^3 did not result in further movement of this substructure.

Table A1. Numerical parameters used for the AREPO and GADGET runs: accuracy of the time step (ETIA), time step displacement factor (MRDF), minimum (MINT) and maximum (MAXT) time step, tolerance on the force accuracy (ETFA), accuracy of the tree algorithm (ETT), Courant factor (CFAC), double precision (DP, DF) and the number of FFT cells along each side of the box (N_{FFT}).

Code name	ETIA	MRDF	MINT	MAXT	ETFA	ETT	CFAC	DP	DF	N_{FFT}
AREPO	0.025	0.125	0	0.01	0.0025	0.6	0.3	Y	Y	512
G2-ANARCHY	0.01	0.125	0.0	0.01	0.025	0.3	0.3	Y	Y	512
G3-X-ART	0.01	0.5	0	0.01	0.01	0.45	0.15	N	N	512
G3-SPHS	0.03	0.5	0	0.02	0.005	0.5	0.4	N	N	1024
G3-MAGNETICUM	0.05	0.25	0	0.05	0.005	0.45	0.15	Y	Y	256
G3-PESPH	0.05	0.25	10^{-7}	0.05	0.005	0.4	0.15	Y	Y	256
G3-MUSIC	0.01	0.5	0	0.01	0.01	0.4	0.15	Y	Y	512
G3-OWLS	0.025	0.25	10^{-10}	0.025	0.005	0.6	0.15	Y	N	1024
G2-X	0.02	0.25	10^{-7}	0.025	0.0025	0.3	0.15	Y	Y	256
Common parameter set	0.01	0.125	0.0	0.01	0.025	0.3	0.15	Y	Y	512

The largest reduction of scatter is seen in the outskirts, where the location of most massive subhalo around the 7 o'clock position is well matched just inside R_{200}. Despite this improvement, the ART profile shows deviation of order 10 per cent due to the positional offset of another massive subhalo around the 12 o'clock position. As we advance the ART run by ∼20 per cent of the dynamical time of the cluster at R_{200}, we are able to better match the position of this subhalo, indicating that this is due to timing offset. Further investigation is needed, and we plan to do so in our follow-up work using multiple haloes with varying dynamical states.

For this common choice the gravitational evolution of all the simulations is also essentially indistinguishable in the non-radiative case, as illustrated by Fig. A2, which gives visual images of the dark matter distribution in the non-radiative runs. Fig. A3 shows the radial density distribution of the dark matter in the non-radiative case, and the difference relative to the G3-MUSIC simulation. For HYDRA, the central gas density is so high that it steepens the central dark matter distribution relative to the other codes. As a consequence of adding the gas, the scatter in the radial dark matter density profile is higher than that seen in the dark matter-only runs. There is also a marked split between those codes that produce a core in the radial gas distribution and those that do not. This is in the sense expected and demonstrates the expected back-reaction in the dark matter potential. As in the case of dark matter-only aligned simulations, the difference in the profiles of ART is due to mismatch of the main subhalo in the 12 o'clock position caused by timing offset, which we will investigate in detail in another upcoming paper.
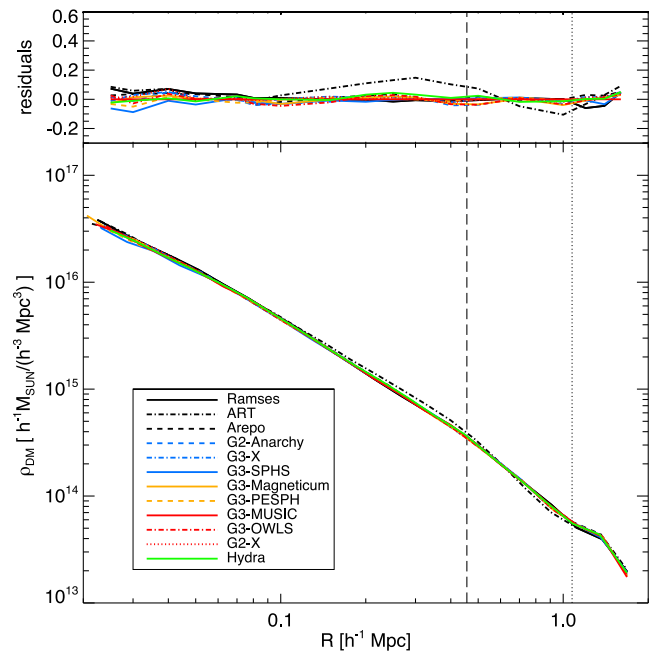

Figure A1. Radial density profiles for the dark matter-only simulations with aligned parameters at $z = 0$ (bottom panel) and difference between the radial density profiles of dark matter-only simulation at $z = 0$ and the reference G3-MUSIC density profile (top panel). The vertical dashed line corresponds to R_{2500} and the vertical dotted line to R_{500} of the reference G3-MUSIC values.
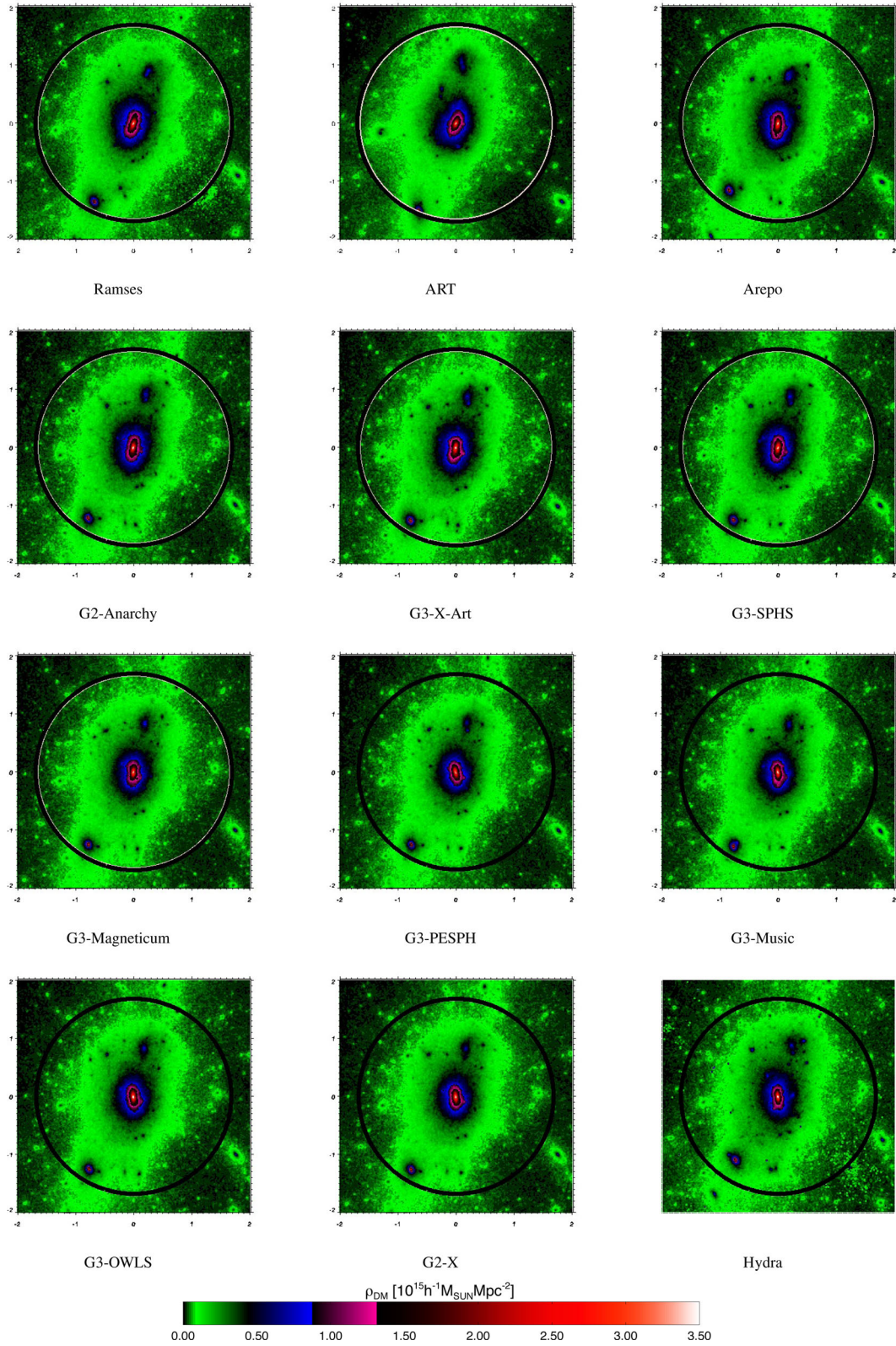
Figure A2. Projected density of the dark matter halo in the non-radiative simulations at $z = 0$ for each method as indicated. The box is $2\ h^{-1}$ Mpc on a side. The white circle indicates M_{crit}^{200} for the halo, the black circle the same but for the G3-MUSIC simulation.
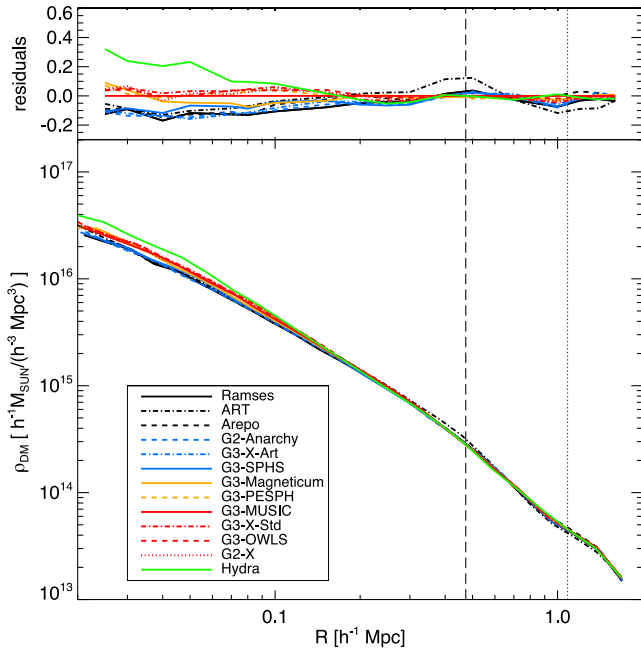
Figure A3. Radial dark matter density profiles for the non-radiative simulations at $z = 0$ (bottom panel) and difference between the radial density profiles of each non-radiative simulation at $z = 0$ and the reference G3-MUSIC density profile (top panel). The vertical dashed line corresponds to R_{2500} and the vertical dotted line to R_{500} of the reference G3-MUSIC values.

[1] *Departamento de Física Teórica, Módulo 8, Facultad de Ciencias, Universidad Autónoma de Madrid, E-28049 Madrid, Spain*

[2] *Astro-UAM, UAM, Unidad Asociada CSIC*

[3] *Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 5, I-00185 Roma, Italy*

[4] *School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, UK*

[5] *Jodrell Bank Centre for Astrophysics, School of Physics and Astronomy, The University of Manchester, Manchester M13 9PL, UK*

[6] *International Centre for Radio Astronomy Research, University of Western Australia, 35 Stirling Highway, Crawley, Western Australia 6009, Australia*

[7] *University Observatory Munich, Scheinerstr. 1, D-81679 Munich, Germany*

[8] *Max Planck Institute for Extraterrestrial Physics, Giessenbachstr. 1, D-85748 Garching, Germany*

[9] *Max Planck Institute for Astrophysics, Karl-Schwarzschild-Str. 1, D-85741 Garching, Germany*

[10] *Astronomy Unit, Department of Physics, University of Trieste, via G.B. Tiepolo 11, I-34143 Trieste, Italy*

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

[12] *INFN – Sezione di Trieste, via Valerio 2, I-34127 Trieste, Italy*

[13] *Instituto de Astrofísica de Canarias, C/ Vía Láctea s/n, E-38205 La Laguna, Tenerife, Spain*

[14] *Departamento de Astrofísica, Universidad de La Laguna, Av. del Astrofísico Franciso Sánchez s/n, E-38206 La Laguna, Tenerife, Spain*

[15] *Physics Department, University of Western Cape, Bellville, Cape Town 7535, South Africa*

[16] *South African Astronomical Observatory, PO Box 9, Observatory, Cape Town 7935, South Africa*

[17] *African Institute of Mathematical Sciences, Muizenberg, Cape Town 7945, South Africa*

[18] *Sydney Institute for Astronomy, A28 School of Physics, The University of Sydney, NSW 2006, Australia*

[19] *Center for High Performance Computing, CSIR Campus, 15 Lower Hope Street, Rosebank, Cape Town 7701, South Africa*

[20] *Astronomy Department, University of Massachusetts, Amherst, MA 01003, USA*

[21] *Institute for Astronomy, Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 16, CH-8093 Zurich, Switzerland*

[22] *Department of Physics, Yale University, New Haven, CT 06520, USA*

[23] *Yale Center for Astronomy and Astrophysics, Yale University, New Haven, CT 06520, USA*

[24] *Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK*

[25] *Department of Astronomy, Yale University, New Haven, CT 06511, USA*

[26] *Centre for Theoretical Astrophysics and Cosmology, Institute for Computational Science, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland*

[27] *Institute of Astronomy and Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*

[28] *Department of Physics, University of Surrey, Guildford, Surrey GU2 7XH, UK*

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

[30] *Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*

[31] *Department of Astronomy and Physics, Saint Mary's University, 923 Robie Street, Halifax, Nova Scotia B3H 3C3, Canada*

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