

# The MillenniumTNG Project: the hydrodynamical full physics simulation and a first look at its galaxy clusters

Rüdiger Pakmor<sup>1</sup>,<sup>\*</sup> Volker Springel<sup>1</sup>, Jonathan P. Coles,<sup>2</sup> Thomas Guillet<sup>3</sup>, Christoph Pfrommer<sup>4</sup>, Sownak Bose<sup>5</sup>, Monica Barrera,<sup>1</sup> Ana Maria Delgado,<sup>6</sup> Fulvio Ferlito,<sup>1</sup> Carlos Frenk,<sup>5</sup> Boryana Hadzhiyska,<sup>6</sup> César Hernández-Aguayo<sup>1,7</sup>, Lars Hernquist,<sup>6</sup> Rahul Kannan<sup>6</sup> and Simon D. M. White<sup>1</sup>

<sup>1</sup>Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str 1, D-85748 Garching, Germany

<sup>2</sup>Swiss National Supercomputing Centre (CSCS), Lugano 6900, Switzerland

<sup>3</sup>Physics and Astronomy, University of Exeter, Exeter EX4 4QL, UK

<sup>4</sup>Leibniz-Institute für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany

<sup>5</sup>Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

<sup>6</sup>Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

<sup>7</sup>Excellence Cluster ORIGINS, Boltzmannstrasse 2, D-85748 Garching, Germany

Accepted 2022 December 6. Received 2022 December 6; in original form 2022 October 18

## ABSTRACT

Cosmological simulations are an important theoretical pillar for understanding non-linear structure formation in our Universe and for relating it to observations on large scales. In several papers, we introduce our MillenniumTNG (MTNG) project that provides a comprehensive set of high-resolution, large-volume simulations of cosmic structure formation aiming to better understand physical processes on large scales and to help interpret upcoming large-scale galaxy surveys. We here focus on the full physics box MTNG740 that computes a volume of  $740 \text{ Mpc}^3$  with a baryonic mass resolution of  $3.1 \times 10^7 M_{\odot}$  using AREPO with 80.6 billion cells and the IllustrisTNG galaxy formation model. We verify that the galaxy properties produced by MTNG740 are consistent with the TNG simulations, including more recent observations. We focus on galaxy clusters and analyse cluster scaling relations and radial profiles. We show that both are broadly consistent with various observational constraints. We demonstrate that the SZ-signal on a deep light-cone is consistent with Planck limits. Finally, we compare MTNG740 clusters with galaxy clusters found in Planck and the SDSS-8 RedMaPPer richness catalogue in observational space, finding very good agreement as well. However, *simultaneously* matching cluster masses, richness, and Compton- $y$  requires us to assume that the SZ mass estimates for Planck clusters are underestimated by 0.2 dex on average. Due to its unprecedented volume for a high-resolution hydrodynamical calculation, the MTNG740 simulation offers rich possibilities to study baryons in galaxies, galaxy clusters, and in large-scale structure, and in particular their impact on upcoming large cosmological surveys.

**Key words:** hydrodynamics – methods: numerical – galaxies: clusters: general.

## 1 INTRODUCTION

Recent large cosmological hydrodynamical simulations have been very successful in reproducing realistic galaxy populations on cosmological scales. These include Illustris (Vogelsberger et al. 2014a, b), Eagle (Schaye et al. 2015), HorizonAGN (Dubois et al. 2016), IllustrisTNG (Springel et al. 2018), Simba (Davé et al. 2019), NewHorizon (Dubois et al. 2021), Thesan (Kannan et al. 2022b), and Astrid (Bird et al. 2022; Ni et al. 2022). (For a recent overview of the corresponding modelling techniques, see Vogelsberger et al. 2020.) However, among these simulation boxes even TNG300 – as the largest of them – is still too small to properly study large-scale structure features such as baryonic acoustic oscillations or to

contain a representative sample of massive galaxy clusters. Larger cosmological simulations exist, e.g. Magneticum (Dolag, Komatsu & Sunyaev 2016) or Bahamas (McCarthy et al. 2017), but they can only afford a mass resolution that is too low to properly resolve individual galaxies, or have not reached  $z = 0$  yet (Astrid). Yet, it is crucial to understand baryonic physics on large scales, i.e. in galaxy clusters or in the cosmic web, to anchor the baryonic physics model with galaxies and at low redshift where we have many good observational constraints.

Moreover, even larger cosmological box simulations are needed as a basis for the interpretation of upcoming enormous galaxy surveys which aim to better constrain the cosmological parameters of the Universe to the per cent level (e.g. DES, eBOSS, DESI, or Euclid). However, such simulations with sufficiently large volumes can only follow dark matter (Potter, Stadel & Teyssier 2017; Angulo & Hahn 2022) and require significant post-processing to model actual

\* E-mail: [rpakmor@mpa-garching.mpg.de](mailto:rpakmor@mpa-garching.mpg.de)

observables, which for the most part are either exclusively based on baryons or are at least affected by baryonic physics.

A similar problem arises when simulating galaxy clusters. Their sparsity requires large simulation volumes in order to contain a significant number of massive clusters. To simulate them at sufficient resolution, zoom simulations are thus often employed that focus most resolution elements in a small region of interest centred around a single galaxy cluster. These zoom simulations (Bahé et al. 2017; Barnes et al. 2017; Cui et al. 2018) enable the study of the internal structure of galaxy clusters and galaxy cluster scaling relations without the need to follow large volumes at high resolution. It is difficult, however, to compose truly representative samples of clusters with this technique, and to properly model their cosmological foregrounds at the same time.

This study is one of the introductory papers of our new MillenniumTNG project in which we seek to make progress on pushing direct hydrodynamical simulations of galaxy formation to much larger volume than available thus far, and on linking these hydrodynamical results to still larger dark-matter-only simulations, thereby allowing the hydrodynamical results to be more reliably extrapolated to cosmological scales. Simultaneously, this offers a new opportunity to study large samples of hydrodynamically simulated galaxy clusters.

To this end, our work applies the IllustrisTNG state-of-the-art cosmological galaxy formation model in an unparalleled large simulation volume at a mass resolution that still allows us to reasonably describe the properties of individual galaxies. Our most ambitious hydrodynamic calculation is carried out in the  $500 h^{-1}$  Mpc periodic box size of the seminal Millennium simulation, thus motivating the name ‘MillenniumTNG’ we coined for the whole project. The nearly 15 times increase of the simulated volume compared to TNG300 at slightly lower mass resolution allows our hydrodynamic calculation to be more directly compared to upcoming large volume cosmological surveys, and importantly, it enables us to calibrate and improve approximate methods to predict galaxy catalogues based on dark-matter-only simulations. To facilitate the latter, MillenniumTNG additionally consists of a suite of dark-matter-only simulations computed in the same volume, with two ‘fixed-and-paired’ versions at each resolution that use a variance suppression technique that effectively boosts the statistical power of the simulated volume (Angulo & Pontzen 2016). Furthermore, we have computed yet much larger dark-matter-only models that also include massive neutrinos, with up to 1.1 trillion particles in a 3000 Mpc box, which is meant to propel the statistical power of our predictions into the regime probed by the upcoming surveys.

We refer to a companion paper by Hernández-Aguayo et al. (2022) for full details on the simulation suite, its data products (including light-cone outputs and merger trees) and a study of basic matter and halo clustering statistics. In this paper, we focus on introducing the flagship hydrodynamical full physics simulation of the MillenniumTNG project, and on giving an initial characterization of the galaxy clusters in the simulation. This offers a first glimpse at the possibilities the project offers to understand current and future cosmological observations. In Barrera et al. (2022), we present a novel version of the L-Galaxies semi-analytic model of galaxy formation and its application to the light-cone outputs of the MTNG simulations. In further companion papers, Kannan et al. (2022a) analyse the properties of very high-redshift galaxies, while Bose et al. (2022) present a galaxy clustering study based on colour-selected galaxy samples. Hadzhiyska et al. (2022a, b) examine aspects of galaxy assembly bias, whereas Delgado et al. (in preparation) study intrinsic alignments and galaxy shapes. Contreras et al. (2022) introduce an inference technique to constrain the cosmological

parameters of MTNG from galaxy clustering. Finally, Ferlito et al. (in preparation) study weak gravitational lensing both in the dark matter and the full physics runs.

This paper is structured as follows. Section 2 introduces the MillenniumTNG full physics box and summarizes its galaxy formation model and model parameters. Section 3 compares global galaxy properties of the MTNG740 full physics box to the TNG simulations, as well as to updated observational data. Section 4 discusses galaxy clusters in MTNG740 and presents various galaxy cluster scaling relations and compares them to observations. Section 5 analyses different radial profiles of galaxy clusters and puts them in relation to observations. Section 6 compares the SZ signal of a galaxy cluster at  $z = 0.25$  computed from a snapshot time-slice with spherical and cylindrical apertures to the same observable but calculated using the simulation data on a deep full backwards light-cone. Section 7 considers galaxy and cluster observables to compare the galaxy clusters in MTNG740 in observational space with clusters observed with Planck and SDSS-8. Finally, Section 8 concludes with a summary of the paper and an outlook on applications of the MTNG simulations.

## 2 FULL PHYSICS SIMULATIONS IN MILLENNIUM-TNG

The MTNG physics model is based on the IllustrisTNG (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018, 2019b; Pillepich et al. 2018b, 2019; Springel et al. 2018) galaxy formation model (Weinberger et al. 2017; Pillepich et al. 2018a) that has been shown to produce a fairly realistic galaxy population on cosmological scales. It includes primordial and metal line cooling (Vogelsberger et al. 2013), an explicit sub-grid model for the interstellar medium and star formation (Springel & Hernquist 2003), mass return from stars and metal enrichment of the interstellar medium by core-collapse supernovae, thermonuclear supernovae, and AGB stars, an effective model for galactic winds (Pillepich et al. 2018a), and a model for the creation and growth of supermassive black holes as well as their feedback as active galactic nuclei (AGNs; Weinberger et al. 2017).

We keep the physics model of IllustrisTNG with all its parameters unchanged, subject to only a few minor modifications. We fixed all the small issues found during and after the IllustrisTNG simulations have been run and that are documented in Nelson et al. (2019a). Notably, we corrected the unintentional abrupt start of the UV background radiation at the epoch of cosmic reionization. We needed to remove magnetic fields and individual metal species from the model to reduce the memory requirements of the simulation. Instead of following individual metal species we reverted to evolving only a single scalar field that tracks the total metallicity of cells and star particles. Based on small test runs, we do not expect this change to affect the simulation results for galaxies in any significant way. Nevertheless, removing the magnetic fields does however change the model slightly (Pillepich et al. 2018a), and we discuss related effects in more detail in Section 3. Since magnetic fields are omitted, we switched to using an exact Riemann solver for hydrodynamics instead of using an approximate one. In contrast to IllustrisTNG, we did not include any passive tracer particles, again for reasons of memory consumption.

Despite keeping the physics model almost identical to IllustrisTNG, we implemented a large number of technical changes and improvements to the AREPO code (Springel 2010; Pakmor et al. 2016; Weinberger, Springel & Pakmor 2020) in order to make the MTNG740 full physics box fit into the memory of the supercomputer available to us. For example, for this purpose, we

now use shared memory on compute nodes via MPI-3 to store identical data only once per node rather than once per compute core. This includes, most importantly, information about the domain decomposition and the top-level tree that is shared between all MPI ranks, as well as various data tables such as stellar yields or photometric tables for stellar population synthesis models. We now also exploit the shared memory on nodes to improve the efficiency of various global MPI operations, for example, by computing collective results first on each node, and then exchanging them only between nodes rather than between all cores individually. Similarly, we have replaced the global domain decomposition of the code with a hierarchical strategy that first subdivides the simulated volume among all nodes, and then cuts down the pieces further within each node to distribute the work among the available compute cores. Despite the use of well over  $10^5$  MPI ranks, this allows us to always efficiently find a domain decomposition with a well-balanced computational load and a maximum memory overhead of around 10 per cent. Finally, we have reordered some operations in the gravitational tree algorithm, allowing the front-loading of a communication step. In this way, we can now almost completely avoid wait times originating from different numbers of imported and exported resolution elements on a local MPI task (Springel 2010).

We have included as much post-processing as possible already on the fly during the simulation. This includes running the friends-of-friends (FoF) group finder and a novel SUBFIND-HBT substructure finder (Springel et al. 2021) adopted from the GADGET4 into our moving-mesh code AREPO. It also involves computing merger trees and matter power spectra already while the simulation is run. Finally, we output resolution elements when they intersect the past backwards light-cone of a fiducial observer, i.e. light emitted from them will be seen by the fiducial observer exactly at  $z = 0$ . The box is periodically replicated if the light-cone extends beyond the size of the box. We use five different light-cone geometries as described in detail in Hernández-Aguayo et al. (2022).

In the following text, we focus almost exclusively on the hydrodynamical full physics box of the MillenniumTNG project that we refer to as ‘MTNG740’ following the naming convention of the IllustrisTNG project. This calculation follows structure formation in a cubic periodic box with side length  $500 h^{-1}$  Mpc (740 Mpc), using  $4320^3$  dark matter particles of mass  $1.7 \times 10^8 M_\odot$  and initially  $4320^3$  gas cells with an initial mass of  $3.1 \times 10^7 M_\odot$ , which is the targeted baryonic mass resolution. This makes MTNG740 15 times larger than TNG300 by volume at a mass resolution that is only 2.8 times worse. The minimum gravitational softening length for gas cells is set to  $\epsilon_{\text{gas, min}} = 370$  pc (it changes with the cell size), while the gravitational softening length of dark matter and stars is set to  $\epsilon_{\text{DM, *}} = 3.7$  kpc.

MTNG uses the Planck 2016 cosmology (Planck Collaboration XIII 2016a) in identical form as IllustrisTNG in order to facilitate direct comparisons with TNG, i.e.  $\Omega_0 = 0.3089$ ,  $\Omega_{\text{nb}} = 0.0486$ ,  $\Omega_\Lambda = 0.6911$ ,  $\sigma_8 = 0.8159$ ,  $n_s = 0.9667$ , and  $h = 0.6774$ . We generate the initial conditions at  $z = 63$  with second-order Lagrangian perturbation theory using GADGET4 (Springel et al. 2021). We apply the fixed-and-paired variance suppression technique (Angulo & Pontzen 2016), although we can only afford to run one realization of the pair, which gives however already a good part of the benefit of this approach. For the matching dark matter simulations, we have simulated both pairs, however.

The MTNG740 simulation was executed on 122 880 cores on the SuperMUC-NG machine at the Leibniz Computing Center, for a total wallclock time of 57 d. It consumed  $1.7 \times 10^8$  core-hours and had

a total memory requirement of 180 TB. The total data output of the simulation amounts to 1.1 PB.

In the following text, we refer to the radius that encloses an average density of 500 times the critical density of the universe as  $R_{500c}$ , and for brevity, we shall usually drop the subscript c in the rest of the paper. We do the same for the analogous quantities  $M_{500}$ ,  $M_{200}$ , and  $R_{200}$ .

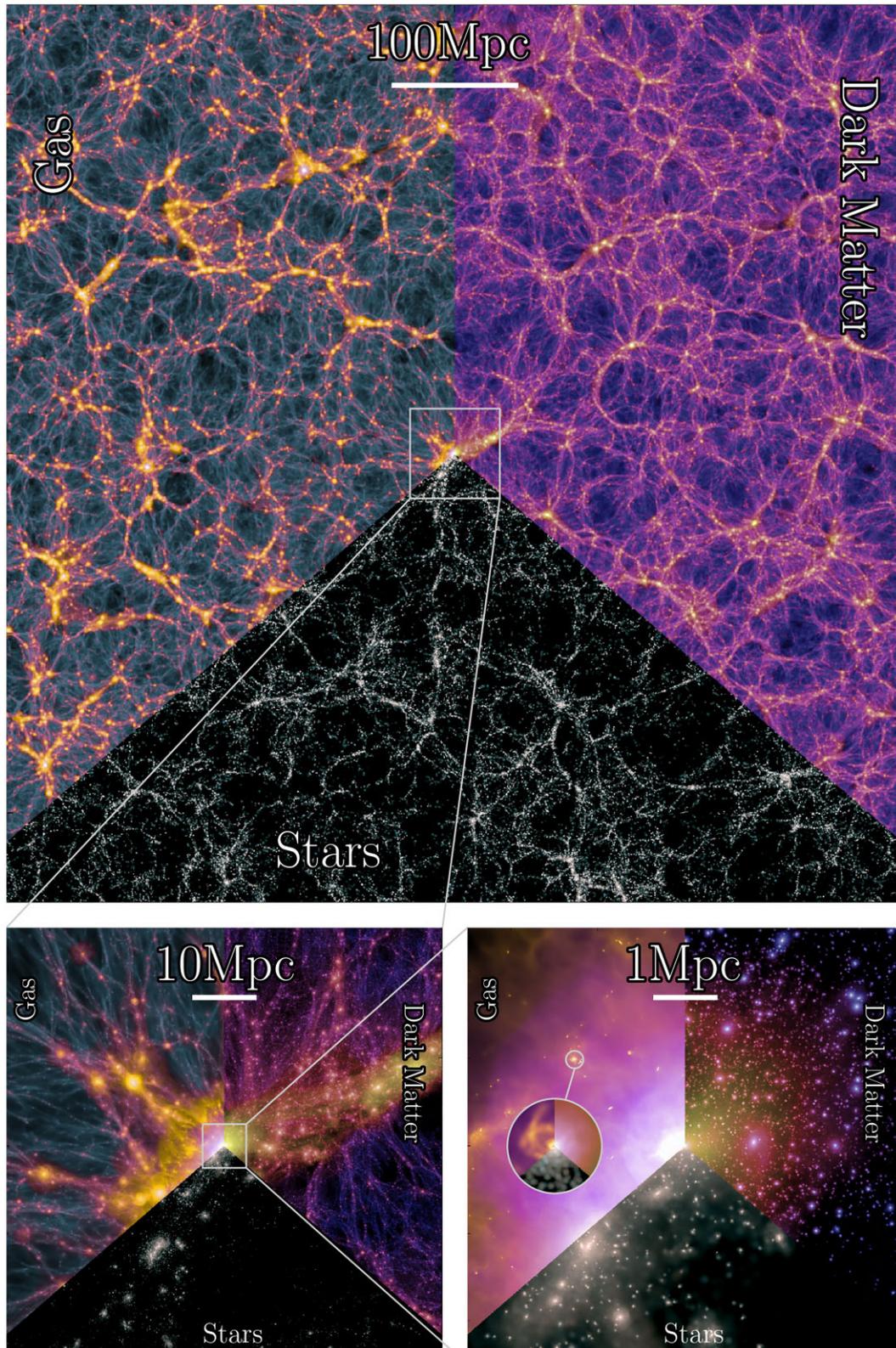
### 3 VERIFICATION

Fig. 1 shows a visual impression of the flagship MTNG740 full physics simulation. To verify the galaxy formation results of the MTNG740 full physics box, we compare them to a set of simulations from the IllustrisTNG project. We specifically compare with TNG100, which has the resolution the galaxy formation physics model was calibrated at (Pillepich et al. 2018a), with TNG300, which is the largest box of IllustrisTNG and has a resolution closest to MTNG740, and with TNG300-2, which has the same box size as TNG300 but features an eight times coarser mass resolution. Thus, TNG300 and TNG300-2 bracket MTNG740 in mass resolution. In Table 1, we list the main numerical parameters of these simulations and contrast them with the MTNG740 full physics box. Notably, MTNG740 simulates a nearly 15 times bigger volume than TNG300 at a mass resolution that is worse by a factor of 2.8. This is made possible by following 5.2 times more resolution elements than TNG300.

To assess the galaxy properties computed by the MTNG740 full physics simulation, we reproduce in Fig. 2 the main calibration plots used for the TNG galaxy formation model (Weinberger et al. 2017; Pillepich et al. 2018a) for the four simulations summarized in Table 1. Where applicable, we bin the simulation data logarithmically with four bins per decade. We show data points for individual objects instead of the median of the bin for bins that contain fewer than 10 data points. We show individual objects as points rather than medians once a bin contains fewer than 10 objects.

To guide the eye and to allow a first rough qualitative assessment, we also include updated observational data in every panel, while for the original data used in the calibration of the model we refer to Pillepich et al. (2018a). Note, however, that all data shown here are quantities inferred from observations based on additional assumptions, they are not direct observables. A more powerful comparison would require elaborate forward modelling of mock observations of the simulations, such that synthetic observables could be contrasted directly with observed quantities. This is beyond the scope of this paper but will be addressed in future work.

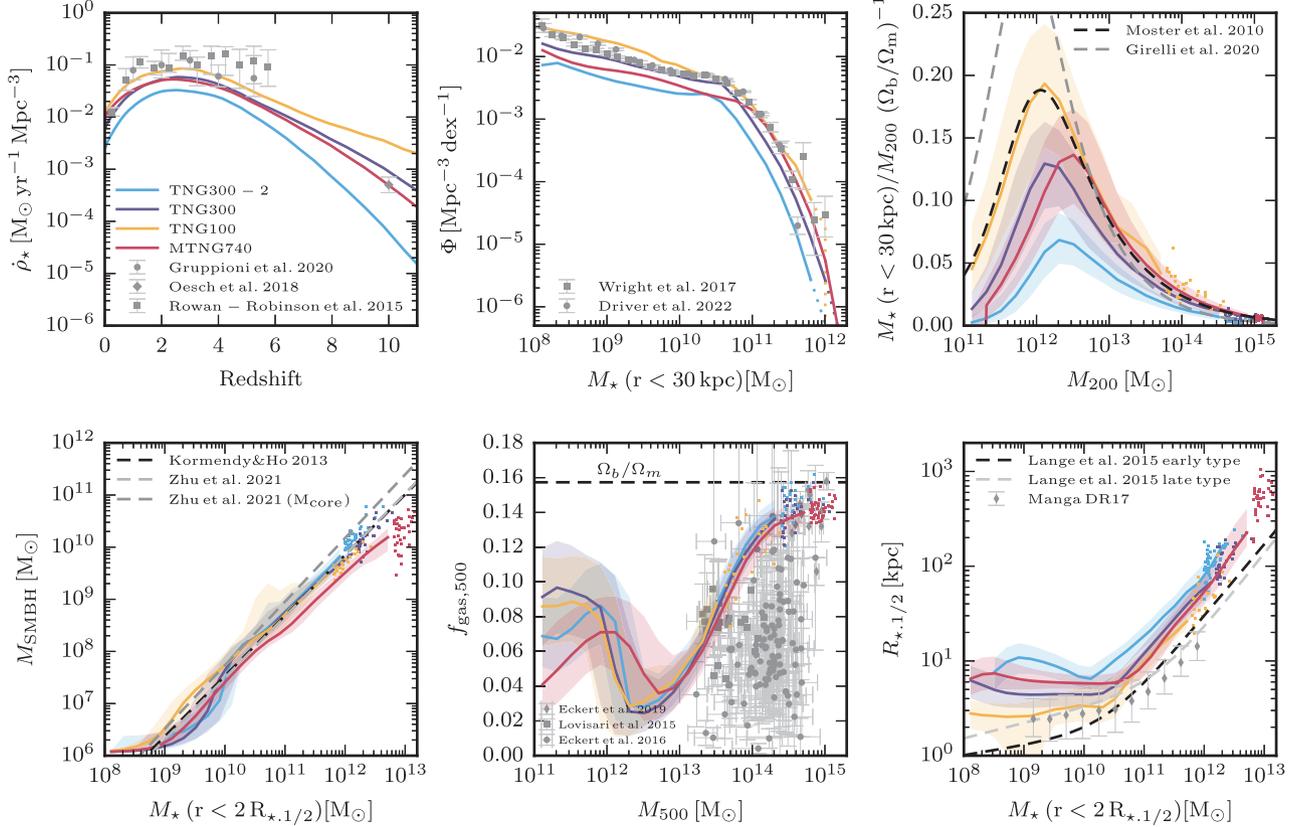
In the top left panel of Fig. 2, we show the evolution of the cosmic star formation rate density,  $\dot{\rho}_*$  [ $M_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}$ ]. We compute it from the initial mass of all star particles formed in the simulations, using 100 logarithmically spaced redshift bins. The star formation rate density of the simulations changes systematically with mass resolution such that simulations with better mass resolution form more stars. MTNG740 lies closest to TNG300, albeit slightly below it at high redshift, and slightly above it at  $z = 0$  where it almost matches TNG100. For comparison, we consider recent results from reconstructions of the star formation rate density of the Universe between  $z = 0.5$  and  $z = 6$  using sub-mm observations (Gruppioni et al. 2020), from infrared observations between  $z = 0.5$  and  $z = 6$  (Rowan-Robinson et al. 2016), and at high redshift from UV data (Oesch et al. 2018). The star formation rate density in MTNG740 is slightly lower than observations for  $z < 6$  but in good agreement with the data points at  $z = 10$ . Note that there are known inconsistencies between the star formation rate density reconstructed from star



**Figure 1.** Thin projections of gas (top left), dark matter (top right), and stellar light (bottom centre) for a depth of 10 Mpc at  $z = 0$ . The projections show the vast physical scales in the simulation from the full 740 Mpc box to an individual spiral galaxy (zoomed inset) with a radius of 50 kpc.

**Table 1.** Basic numerical parameters of the MTNG740 full physics box compared to TNG100, TNG300, and TNG300-2.  $N_{\text{ICs}}$  denotes the number of resolution elements in the initial conditions that include an equal number of dark matter particles and gas cells. The MTNG hydrodynamical model follows a 15 times larger volume than TNG300 at a mass resolution that lies between TNG300 and TNG300-2.

Name	BoxSize [comoving] [ $h^{-1}\text{Mpc}$ ]	BoxSize [phys] [Mpc]	Volume [ $\text{Mpc}^3$ ]	$N_{\text{ICs}}$	$M_{\text{dm}}$ [ $M_{\odot}$ ]	$M_{\text{baryon}}$ [ $M_{\odot}$ ]	$\epsilon_{\text{gas, min}}$ [kpc]	$\epsilon_{\text{dm, *}}$ [kpc]
MTNG740	500	740	$4.0 \times 10^8$	$2 \times 4320^3$	$1.7 \times 10^8$	$3.1 \times 10^7$	0.37	3.7
TNG300	205	300	$2.7 \times 10^7$	$2 \times 2500^3$	$5.9 \times 10^7$	$1.1 \times 10^7$	0.37	1.5
TNG300-2	205	300	$2.7 \times 10^7$	$2 \times 1250^3$	$4.7 \times 10^8$	$8.8 \times 10^7$	0.74	3.0
TNG100	75	110	$1.6 \times 10^6$	$2 \times 1820^3$	$7.5 \times 10^6$	$1.4 \times 10^6$	0.19	0.74



**Figure 2.** Verification plots for MTNG740. We show the quantities used to calibrate the TNG model at the resolution of TNG100 (Pillepich et al. 2018a) and compare to updated observational data. The individual panels show the cosmic star formation rate density (top left), the galaxy stellar mass function (top centre), the stellar mass–halo mass relation (top right), the mass of the central supermassive black hole at fixed host galaxy stellar mass (bottom left), the gas fraction of haloes (bottom centre), and the sizes of galaxies (bottom right). We compare the new MTNG740 simulation with the TNG100, TNG300, and TNG300-2 simulations at  $z = 0$ . Galaxies are identified by SUBFIND and haloes by the FoF algorithm. We show binned medians and 68 per cent percentile bands computed for four bins per decade, while individual objects are displayed instead if a bin contains fewer than 10 objects. We see overall very good agreement of MTNG740 with its predecessor simulations. We also include recent observed relations and data to guide the eye, and to allow a rough qualitative assessment. In particular, at masses  $M_{200c} > 10^{12.5} M_{\odot}$  we see good agreement with observations, except for the hot gas fraction in all but the most massive clusters. Note that, interestingly, MTNG740 agrees best with TNG100 in the stellar mass–halo mass relation, even though it is closest to TNG300 in resolution. This is related to our omission of magnetic fields in MTNG740.

formation rate indicators and the total stellar mass density inferred for the Universe. The latter seems to favour a star formation rate density that is smaller by about a factor of 2 at  $z = 2$  (Wilkins, Trentham & Hopkins 2008; Yu & Wang 2016).

The top centre panel of Fig. 2 shows the galaxy stellar mass function. Here, we measure the stellar mass in a fixed physical radius of 30 kpc that is bound to the galaxy according to SUBFIND. Satellite galaxies are included as well. For comparison, we show the galaxy stellar mass function from the Galaxy And Mass Assembly (GAMA) survey (Wright et al. 2017; Driver et al. 2022).

We find good agreement at the high-mass end,  $M_{*} \gtrsim 10^{11} M_{\odot}$ . MTNG740 is lower by about a factor of 2 at fixed stellar mass for galaxies with around  $M_{*} \sim 5 \times 10^{10} M_{\odot}$  and about 30 per cent lower compared to observations for galaxies with stellar masses  $M_{*} < 10^{10} M_{\odot}$ .

In the top right panel of Fig. 2, we show the stellar mass–halo mass relation for all central galaxies and their dark matter haloes. We compute the ratio of the total stellar mass within a fixed physical aperture of 30 kpc to  $M_{200}$  and divide by the cosmic baryon fraction. We compare the simulations to relations inferred from abundance

matching that are based on galaxy counts and dark-matter-only simulations of a  $\Lambda$  cold dark matter universe (Moster et al. 2010; Girelli et al. 2020). MTNG740 is in excellent agreement with TNG100 and abundance-matching results for massive haloes with  $M_{200} > 10^{12.5} M_{\odot}$ . For less massive haloes with  $M_{200} < 10^{12.5} M_{\odot}$ , the stellar mass of galaxies in MTNG740 is smaller at fixed halo mass than for galaxies in TNG100 but comparable to galaxies in TNG300. Note that at these smaller masses there is also large variation in the estimates from abundance matching and other methods that attempt to reconstruct the relation between stellar mass and halo mass from observations (see e.g. Girelli et al. 2020).

In the bottom left panel of Fig. 2, we show the relation between the mass of the most massive black hole of a galaxy and its central stellar mass. Following Pillepich et al. (2018a), we measure the stellar mass as the total stellar mass within twice the stellar half-mass radius,  $R_{*,1/2}$ . We include central galaxies as well as satellite galaxies. We find good agreement between TNG100, TNG300, and TNG300-2. The black holes in MTNG740 are less massive by up to a factor of 2 compared to the TNG simulations at fixed stellar mass. For comparison, we show the observed relation between black hole mass and stellar mass of the bulge of their host galaxies that was used for calibration of the IllustrisTNG model (Kormendy & Ho 2013), as well as two more recent relations between black hole mass and stellar bulge and core masses (Zhu, Ho & Gao 2021). The new bulge mass relation is essentially identical to the old black hole mass–bulge mass relation. The new core mass relation has the same slope but a slightly higher normalization. We find overall reasonable agreement between MTNG740 and the observed relations. Note, however, that the stellar mass estimate we use here includes a much larger fraction (about 80 per cent) of the total stellar mass of galaxies than observational estimates of bulge mass. Moreover, at the high-mass end, the stellar mass estimate includes significant contributions from the intercluster light component. We discuss this in a bit more detail in the context of the galaxy size estimates shown in the bottom right panel of Fig. 2. Also note that Borrow et al. (2022) recently identified an unintended behaviour of the black hole repositioning scheme in the code that can cause massive galaxies to lose their central black holes when their gas is completely stripped by a more massive object (e.g. within a galaxy cluster). This is also present in MTNG740, but does not affect the median values shown here.

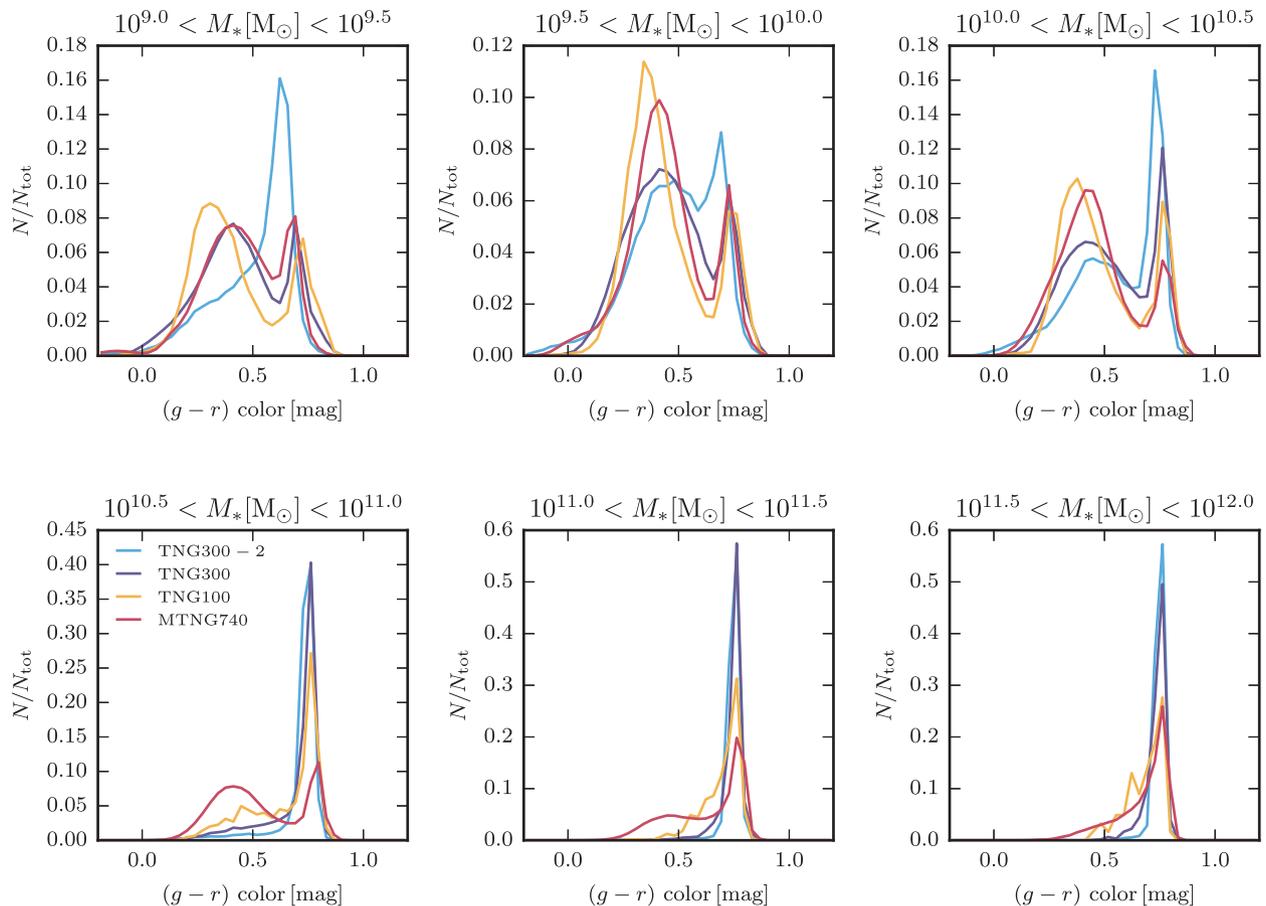
In the bottom centre panel of Fig. 2, we show the gas fraction within  $R_{500}$  and its dependence on  $M_{500}$  for haloes in MTNG740, finding good agreement between the different simulations for  $M_{200} > 5 \times 10^{12} M_{\odot}$ . We compare to observational data for galaxy groups (Lovisari, Reiprich & Schellenberger 2015) and galaxy clusters (Eckert et al. 2016, 2019). We find that MTNG740 is consistent with the latest data points for the most massive clusters, and also agrees on the group scale. For galaxy clusters with masses  $M_{200} \gtrsim 5 \times 10^{13} M_{\odot}$ , however, the inferred gas fractions of most observed clusters are significantly below the gas fractions of the MTNG740 clusters. Moreover, the observational data have a much larger scatter at fixed halo mass than the simulations. A more detailed comparison of the gas fractions in TNG300 with observations using mock X-ray images and estimating the gas mass and total mass of galaxy clusters in the more similar way to observations also finds systematically larger values for TNG300 compared to observations, but much better agreement on the scatter at fixed halo mass (Pop et al. 2022). Additional studies will be needed to understand if the differences at  $M_{500} \gtrsim 10^{14} M_{\odot}$  and in the associated scatter are primarily caused by observational biases, or rather reflect a limitation of the galaxy formation model in the simulations.

Finally, in the bottom right panel of Fig. 2 we show the stellar half-mass radius of galaxies depending on their stellar mass within twice the stellar half-mass radius. The plot includes central as well as satellite galaxies. The stellar half-mass radius is computed for all stars associated with a galaxy by SUBFIND. For the central galaxies of massive haloes this also includes all stars that are part of the halo but not bound to any satellite galaxy, i.e. the intracluster light. This leads to large size estimates for the most massive galaxies. The different simulations are consistent for  $M_{*} > 10^{11} M_{\odot}$ , but they show a systematic trend with mass resolution for lower mass galaxies, where better resolution simulations have smaller galaxies at fixed stellar mass. For comparison, we show data from the GAMA survey (Lange et al. 2015) using their radii computed from  $r$ -band luminosity profiles, and from 10 000 galaxies of the Manga (Bundy et al. 2015) survey released in SDSS DR17 (Abdurro'uf et al. 2022) using stellar mass and effective radius estimates from the PIPE3D pipeline (Sánchez et al. 2016). The simulations are reasonably consistent with the observations, although galaxies in MTNG740 appear systematically larger than the observed galaxies by about a factor of 2 at fixed stellar mass.

An important further diagnostic of the galaxy population is its colour distribution. In particular, it is fundamental to quantify to which degree our simulated galaxy population reproduces the observed mass-dependent bimodal colour distribution. Since galaxy colour depends strongly on the specific star formation rate this is a good measure of quenching. TNG100 in particular has been shown to agree well with the observed colour distribution (Nelson et al. 2018). In Fig. 3, we compare the distribution of  $g - r$  SDSS colours of MTNG740 galaxies to the colour distributions of the TNG simulations. Colours are computed assuming a single stellar population for each star particle with a Chabrier initial mass function (IMF) (Chabrier 2003) and the Bruzual & Charlot (2003) stellar population synthesis model without dust. The luminosity of all star particles bound to a galaxy according to SUBFIND is then summed up to compute the total galaxy luminosity. We find excellent agreement between MTNG740 and TNG100 for low masses  $M_{*} < 10^{10.5} M_{\odot}$  and at the highest masses,  $M_{*} > 10^{11.5} M_{\odot}$ . The sharp transition that is present in all TNG simulations at a mass scale  $M_{*} \sim 5 \times 10^{10.5} M_{\odot}$  is however shifted to higher masses by 0.5 dex in MTNG740, and appears washed out a bit as well. We associate this with differences in the efficiency of the kinetic AGN feedback introduced by the removal of magnetic fields from the TNG physics model (see Pillepich et al. 2018a). The change in colour distribution for galaxies with masses in the range  $10^{10} M_{\odot} < M_{*} < 10^{10.5} M_{\odot}$  directly affects the colour-selected galaxy correlation function in this mass range, which agrees less well with observations than TNG300 (Bose et al. 2022).

As a final piece of validation of the MTNG740 model against IllustrisTNG, we look at the baryonic impact on the total matter power spectrum, which is an important prediction of these non-linear simulations and can affect cosmological constraints. In Fig. 4, we show the ratio of the total matter power spectra of the full physics simulations of MTNG740, TNG100, and TNG300 divided by the total matter power spectra of their dark-matter-only counterparts. We find that the baryonic impact of MTNG740 is almost identical to TNG100 and slightly smaller than in TNG300, with a 1 per cent deviation at  $k = 2 h \text{ Mpc}^{-1}$ , a minimum just above  $k = 10 h \text{ Mpc}^{-1}$ , and a relative difference on this scale of  $\sim 20$  per cent.

Following van Daalen, McCarthy & Schaye (2020), we find a suppression of the power spectrum at  $z = 0$  from baryonic effects at a scale of  $0.5 h \text{ Mpc}^{-1}$  of  $\Delta P / P_{\text{DM only}} = -0.0043$  and an average baryon fraction of all haloes in the range  $6 \times 10^{13} < M_{500c} [M_{\odot}] < 2 \times 10^{14}$  at  $z = 0$  of 0.86 in MTNG740. According to the fit in van



**Figure 3.** Colour distribution of galaxies in the SDSS  $g-r$  colour for different stellar mass bins at  $z=0$ . Colours are computed following Torrey et al. (2014) assuming a single stellar population for each star particle with a Chabrier IMF (Chabrier 2003) and the Bruzual & Charlot (2003) stellar population synthesis model without dust. Each panel shows a normalized colour histogram of all galaxies in the corresponding mass bin, using 40 colour bins. We compare the MTNG740 simulation with the TNG100, TNG300, and TNG300-2 simulations. At the high- and low-mass end, MTNG740 agrees well with the TNG simulations and best with TNG100. For stellar masses between  $10^{10.5} M_{\odot}$  and  $10^{11.5} M_{\odot}$  MTNG740 has a more prominent contribution of blue galaxies compared to the TNG simulations. Thus, the mass scale at which galaxies are quenched moves to slightly higher masses in MTNG740.

Daalen et al. (2020), a suppression of  $\Delta P/P_{\text{DM only}} = -0.0034$  for our baryon fraction would be expected, which is roughly consistent but 25 per cent smaller than our actual suppression value.

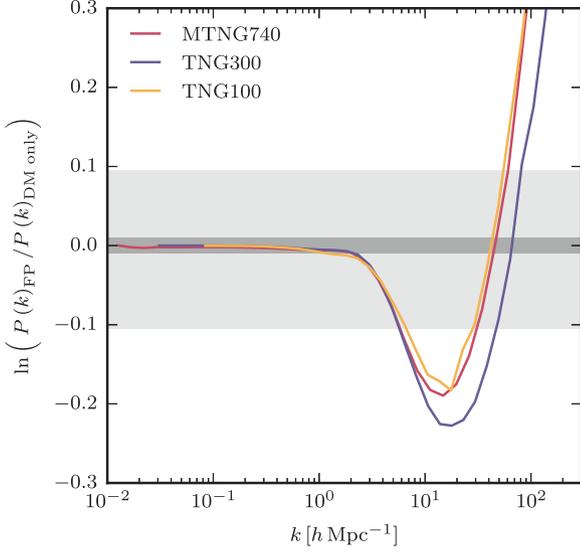
We conclude from the results in this section that there is reassuring agreement between MTNG740 and its TNG predecessors, and that most of the results found for the TNG simulations, in particular for TNG100 and TNG300, can also be expected to hold for MTNG740. At some level, one can therefore view it as a very large volume extension of the TNG suite of simulations, just as we intended.

#### 4 GALAXY CLUSTER SCALING RELATIONS

Owing to its large volume, MTNG740 contains a substantial number of massive galaxy clusters. At  $z=0$  the full physics box encompasses nine galaxy clusters with  $M_{500} > 10^{15} M_{\odot}$ , and more than 2000 galaxy clusters with  $M_{500} > 10^{14} M_{\odot}$ . At  $z=0.25$ , the time of the last full snapshot before  $z=0$ , which is important because observations cover many more clusters at this epoch than at  $z=0$ , the simulation already contains three massive galaxy clusters with  $M_{500} > 10^{15} M_{\odot}$  and more than 1500 galaxy clusters with  $M_{500} > 10^{14} M_{\odot}$ . These numbers allow us to have a detailed look into cluster scaling relations for a representative sample of simulated galaxy clusters, and study their properties in detail.

In Fig. 5, we show six different galaxy cluster scaling relations for all galaxy clusters at  $z=0.25$  with  $M_{500} > 10^{14} M_{\odot}$ . These relations are often used to infer the mass of a galaxy cluster from observables. We compare the galaxy clusters in MTNG740 with a sample of well-studied clusters of the LoCuSS sample (Mulroy et al. 2019) at similar redshifts between  $z=0.15$  and  $z=0.3$ . This sample provides various scaling relations between cluster observables from multiwavelength observations of 41 galaxy clusters and weak-lensing-based mass estimates. In addition to weak lensing maps, the observations used to derive the scaling relations include X-ray, millimetre, optical, and infrared observations. The weak lensing masses inferred for the sample range from  $2 \times 10^{14} M_{\odot}$  to  $2 \times 10^{15} M_{\odot}$ . Note that even though we use the weak lensing masses as baseline for our comparison, here they also still carry significant systematic uncertainties (Ardila et al. 2021). The mass values  $M_{500}$  for the galaxy clusters in MTNG740 are computed based on spherical overdensity measurements around the potential minimum of the galaxy cluster.

The top-left panel of Fig. 5 shows the richness  $\lambda$ . For the MTNG740 galaxy clusters we compute the richness as the number of galaxies within  $R_{500}$  of the galaxy cluster with an intrinsic luminosity in the rest-frame  $r$  band brighter than  $L_r < L_r^* + 1.75 = -19.75$ . Here, we assume  $L_r^* = -21.5$  (Montero-Dorta & Prada 2009) to mimic the luminosity cut used to compute the richness for the



**Figure 4.** Baryonic effects on the total matter power spectrum at  $z = 0$  for MTNG740, TNG100, and TNG300. For each simulation, we show the ratio of the total matter power spectrum of the full physics simulation divided by the total matter power spectrum of its dark-matter-only counterpart. The grey bands show 1 per cent and 10 per cent deviations, respectively. MTNG740 is consistent with the TNG simulations, and is best matched by TNG100.

observed galaxy clusters. At  $z = 0.25$  this roughly corresponds to a stellar mass of  $5 \times 10^9 M_\odot$  for satellite galaxies in MTNG740, which lies in the regime where the stellar mass of galaxies in MTNG740 is already significantly below the expected stellar mass–halo mass relation, as shown in the top-left panel of Fig. 2. The slope of the richness–mass relation of the MTNG740 clusters agrees reasonably well with observations. The normalization of the richness of the MTNG740 clusters at fixed cluster mass is, however, about a factor of 2 lower than for the observed clusters.

The interpretation of this difference is not obvious, because it could be caused either by a difference in the richness in the MTNG model or the observational measurement, or by a bias in the weak lensing mass estimate of the observed clusters. A reason for a difference in the richness estimate might be the smaller number of galaxies in MTNG740 compared to observations around and below the knee of the galaxy luminosity function, as shown in the top centre panel of Fig. 2. Moreover, not including a model for dust may bias the galaxy luminosities of the MTNG740 cluster galaxies. In contrast to the observational method, our richness estimate depends strongly on the exact value of the luminosity cut. If we alleviate the luminosity cut to  $L_r < -18$  instead, we can make the richness scaling relation match with the observed data points. On the other hand, if the weak lensing masses underestimate the true mass of the observed galaxy clusters by a factor of 2, both relations will also match. However, this would also affect the other scaling relations that agree significantly better (see below) and therefore appear to be a much less likely explanation. Plausibly, a combination of several factors adds up to the discrepancy we see.

In the top centre panel of Fig. 5, we show the Sunyaev–Zeldovich (SZ) scaling relation. We compute the integrated Compton- $y$  parameter in a sphere of radius  $R_{500}$  as

$$Y_{500} = \frac{\sigma_T}{m_e c^2} \int_0^{R_{500}} dV P_e = \frac{(\gamma - 1)\sigma_T}{m_e c^2} X_e X \mu E_{\text{gas}}, \quad (1)$$

where  $\sigma_T$  is the Thomson cross-section,  $m_e$  is the electron mass,  $c$  is the speed of light, and  $P_e$  is the electron pressure,  $\gamma = 5/3$  is the adiabatic index,  $X_e = n_e/n_H = 1.158$  is the electron-to-hydrogen number density fraction for a hydrogen mass fraction of  $X = 0.76$ , and  $\mu = 0.588$  is the mean molecular weight for a fully ionized medium with primordial abundances. Here,  $E_{\text{gas}}$  is the total thermal energy of the gas within  $R_{500}$ . For comparison with the observed clusters we show  $D_A^2 Y_{500}$ , where  $D_A$  is the angular diameter distance at  $z = 0.25$  for the MTNG740 cosmology. We find good agreement between MTNG740 and the observed galaxy clusters for both the shape and the normalization of the SZ scaling relation. The MTNG740 clusters show significantly less scatter around the relation than the observed clusters. We postpone a detailed study about the origin of the scatter (Battaglia et al. 2012) and the reason for the deviations from the observational scatter to future work.

The top-right panel of Fig. 5 shows the scaling relation for the core-excised bolometric X-ray luminosity. Detailed X-ray mocks are in principle possible (see e.g. Pop et al. 2022, for TNG300 clusters), but are far beyond the scope of this paper. Here, we compute the bolometric X-ray emissivity of a single cell  $j$  following Kannan et al. (2016) and Marinacci et al. (2018) as

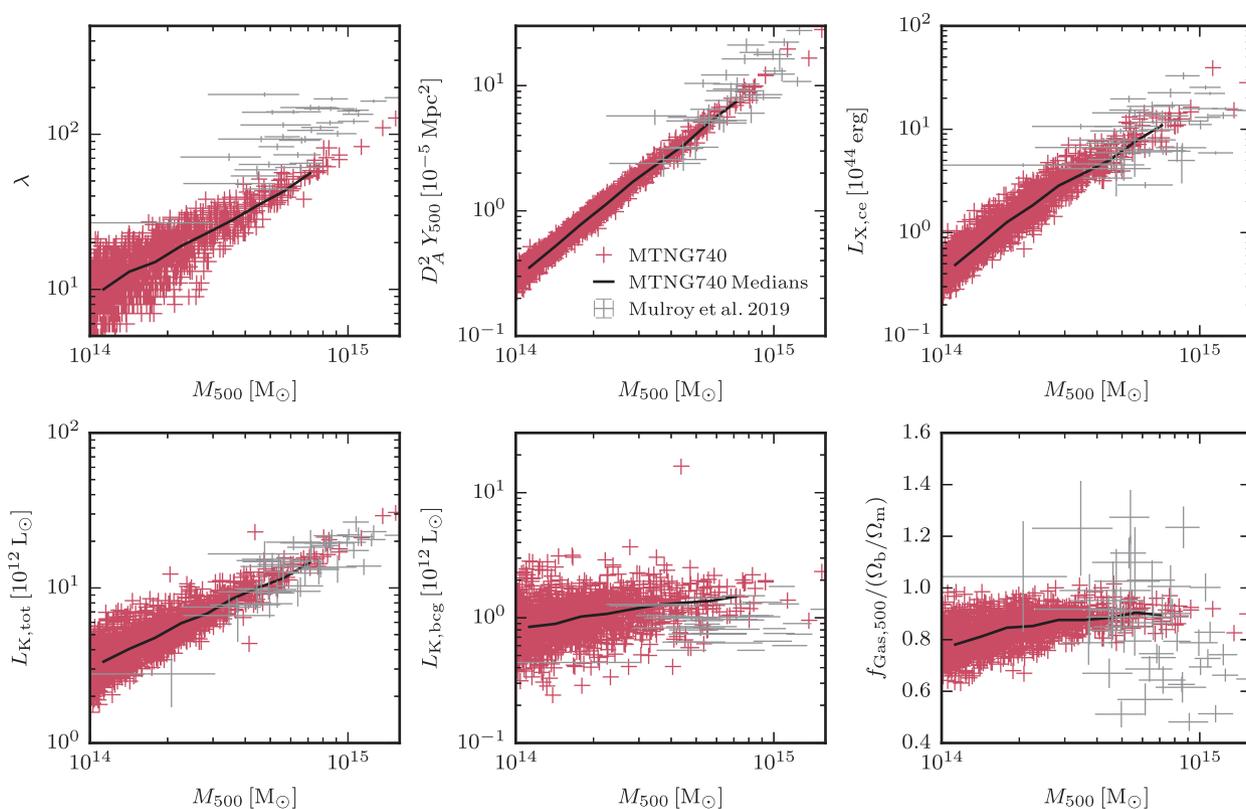
$$L_{j,X} = 6.8 \times 10^{-38} \text{erg K}^{1/2} \text{cm}^3 \frac{k_B T_j^{1/2}}{h} \times \left[ \exp\left(-\frac{E_{\text{low}}}{k_B T_j}\right) - \exp\left(-\frac{E_{\text{high}}}{k_B T_j}\right) \right] \times Z^2 g_{\text{ff}} \frac{X_e}{(X_i + X_e)^2} \left(\frac{\rho_j}{\mu_j m_p}\right)^2 V_j, \quad (2)$$

where  $T$  is the temperature in Kelvin,  $h$  is Planck’s constant (not to be confused with the dimensionless value of the Hubble constant used elsewhere in our work),  $\rho$  the mass density, and  $V$  the volume of a cell (both measured in cgs units). We use  $Z = \sqrt{1.15}$ ,  $g_{\text{ff}} = 1.3$ ,  $X_i = 1.079$ , and  $X_e = 1.16$  for a fully ionized gas of primordial composition (Marinacci et al. 2018). We set the integration limits to  $E_{\text{low}} = 0.7$  keV and  $E_{\text{high}} = 10$  keV to match the observations (Mulroy et al. 2019).

Similar to Pop et al. (2022), we exclude cells that have a temperature below  $10^5$  K, star-forming cells, and cells that have a negative net cooling rate (i.e. that are heated). Our estimated X-ray luminosity scaling relation is roughly consistent with observations. The MTNG740 clusters lie on the observed scaling relation, but with significantly smaller scatter. A significant part of the scatter is likely caused by projection effects that become visible when the simulations are better modelled via mock observables (Pop et al. 2022).

In the bottom-left panel of Fig. 5, we show the scaling relation for the total  $K$ -band luminosity within  $R_{500}$ . We compute the total rest-frame  $K$ -band luminosity of the MTNG740 clusters as the sum of the individual  $K$ -band luminosity of all stars in  $R_{500}$ . We assume a solar luminosity in the  $K$  band of  $M_{K,\odot} = 3.39$  (Johnson et al. 1966) and do not include any dust model accounting for intrinsic dust in the galaxies and galaxy cluster. The observed data points are computed as the sum of the  $K$ -band luminosities of all galaxies of the clusters within a projected radius of  $R_{500}$  that are consistent with being a member of the cluster and brighter than an observational detection threshold. The observed  $K$ -band luminosities are corrected for Milky Way dust. Both scaling relations agree well with each other in normalization and slope. The scatter of the observed clusters is slightly larger than for the MTNG740 clusters.

The good agreement of the total  $K$ -band luminosity indicates that the total stellar mass in  $R_{500}$  matches observations well. Since the richness, however, seems to be smaller than observed, this



**Figure 5.** Scaling relations for all galaxy clusters in MTNG740 with  $M_{500c} > 10^{14} M_{\odot}$  at  $z = 0.25$ . All quantities are computed within  $R_{500}$ . The panels show richness (top left), integrated Compton- $\gamma$  parameter (top centre), core excised ( $0.15 R_{500} < R < R_{500}$ ) bolometric X-ray luminosity (top right), total  $K$ -band luminosity (bottom left),  $K$ -band luminosity of the brightest cluster galaxy (bottom centre), and gas fraction (bottom right) versus  $M_{500}$ . The black line shows the binned median of the MTNG740 clusters. The grey points are data of 41 galaxy clusters between  $z = 0.15$  and  $z = 0.3$  (Mulroy et al. 2019). They are found to be in good agreement with the MTNG740 clusters.

might indicate that galaxies in  $R_{500}$  are stripped too quickly in the simulation, so that they contribute the correct stellar mass into  $R_{500}$ , but at any given time the population of galaxies within  $R_{500}$  is fainter than in observations. Note that satellite stripping is known to depend on numerical resolution with lower resolution simulations showing faster stripping of galaxies when falling into more massive haloes (see e.g. Green, van den Bosch & Jiang 2021).

The bottom centre panel of Fig. 5 shows the scaling relation for the  $K$ -band luminosity of the central brightest cluster galaxy (BCG). For MTNG740 we use the integrated  $K$ -band luminosity of all stars in a fixed physical radius of 30 kpc as a proxy for the luminosity of the central BCG. The BCGs in MTNG740 seem to be brighter in the  $K$  band than the observed BCGs by about a factor of 2. Note that the scatter and the slope of the scaling relations are roughly consistent with observations. The difference is interesting because it might point to shortcomings of the AGN model and its coupling to the central BCG or excessive mergers of the central galaxy with satellite galaxies. In particular, an inaccurate centring of the black hole has been argued to lead to too massive BCGs (Ragone-Figueroa et al. 2018). Owing to our crude definition of the brightness of the simulated BCGs, we refrain from a detailed investigation of the differences here. We leave this to future work that should be based on a faithful forward modelling of MTNG740, and involve mock images of the simulated clusters including a model for intrinsic dust.

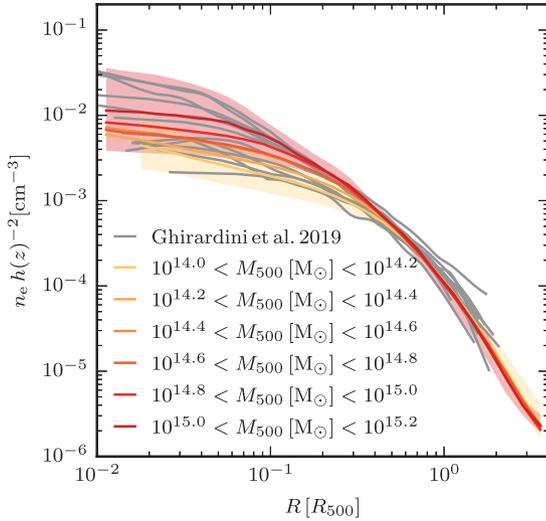
Finally, we show the gas fractions relative to the cosmic baryon fraction in the lower right panel of Fig. 5. For the simulations we compute the gas fraction as the total gas mass in  $R_{500}$  divided by

total mass in  $R_{500}$  (i.e.  $M_{500}$ ). The panel is essentially a zoom-in of the bottom centre panel of Fig. 2 with a different set of data points. The gas fractions of the MTNG740 clusters lie right in the middle of the distribution of the observed data points, but the relation is much tighter for the MTNG740 clusters than for the observed clusters.

For most of the scaling relations we find that the scatter in the observed relations is significantly larger than for the MTNG740 clusters. One potential avenue to explain the differences is observational uncertainties. These contribute to the mass estimate as well as to the observed quantity of the scaling relation, but the error bars usually only include statistical contributions and at best the well-known systematic errors. Moreover, projection effects can play a significant role for the scatter of various cluster observables (Debackere, Hoekstra & Schaye 2022; Pop et al. 2022). Nevertheless, it is also possible that at least some part of this scatter at fixed mass is real and reflects physics not adequately captured in our simulation. In this case, qualitative changes to our physics model will likely be required to significantly increase the object-to-object scatter at fixed mass, for example, for the gas fraction or Compton- $\gamma$  parameter. We emphasize that both better observations and more realistic comparisons between simulations and observations are required to better understand all the sources of the scatter.

## 5 GALAXY CLUSTER PROFILES

After discussing integrated quantities of galaxy clusters, we now go one step further and focus on their internal radial profiles. We select

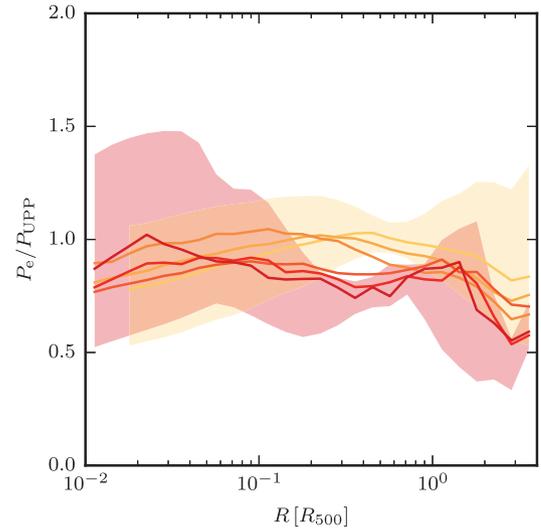
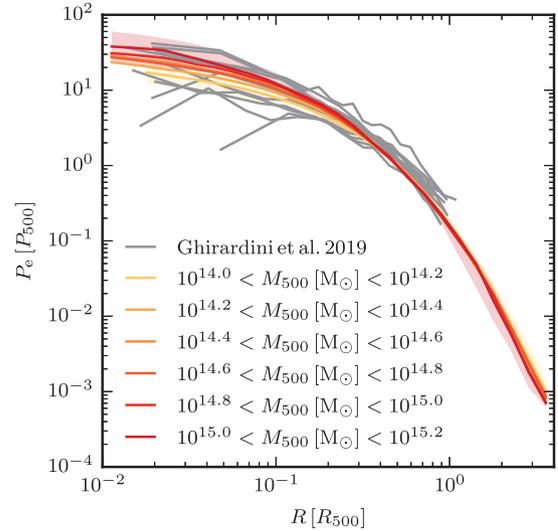


**Figure 6.** Median profiles of the electron density of galaxy clusters with  $M_{500c} > 10^{14} M_{\odot}$  at  $z = 0$  in various narrow mass bins, as labelled. The bands show 10 per cent and 90 per cent percentiles for the lowest and highest mass bins. We use  $h(z) = H(z)/H_0$  for the MTNG740 cosmology. For comparison, we show the profiles of observed galaxy clusters from the X-COP project in grey (Ghirardini et al. 2019). We see good agreement between the MTNG740 clusters and observations. The MTNG740 clusters tend to have a lower central electron density compared to the observed cluster profiles. More massive clusters have a higher central electron density at fixed relative radius in MTNG740.

all 2359 clusters at  $z = 0$  with  $M_{500} > 10^{14} M_{\odot}$  and compare them to observed profiles of nearby galaxy clusters from the X-COP project (Eckert et al. 2019; Ettori et al. 2019; Ghirardini et al. 2019). X-COP provides thermodynamic profiles as well as metallicity and stellar mass profiles (Ghizzardi et al. 2021) for 12 nearby galaxy clusters at redshifts  $0.04 < z < 0.1$ , with estimated masses of  $4 \times 10^{14} M_{\odot} \lesssim M_{500} \lesssim 10^{15} M_{\odot}$ . The results are based on deep *XMM-Newton* X-ray as well as millimetre observations to reconstruct high-precision profiles of the cluster gas. The 12 galaxy clusters of the X-COP survey are selected from the Planck all-sky SZ map and constitute the most significant detections in the corresponding redshift range.

We first look at different thermodynamical profiles of the gas in clusters. Even though these profiles are strongly correlated, they allow us different insights into the thermodynamical state of the gas in the MTNG740 galaxy clusters.

We start with the electron density profile in Fig. 6. We compute the volume-averaged electron density profile in spherical shells and group the MTNG740 clusters into six logarithmic mass bins in the range  $14 < \log_{10}(M_{500}/M_{\odot}) < 15.2$ . We scale the profiles with  $R_{500}$  of each cluster, then combine them. We then show the median electron density profile in each mass bin with 80 per cent percentile bands, and we compare to the observed cluster profiles from the X-COP project (Ghirardini et al. 2019). All cluster profiles compare the 3D profiles of the MTNG740 clusters with the 3D profiles inferred from observed 2D profiles. We find a reassuring agreement between MTNG740 and the X-COP profiles. The typical scatter of the MTNG740 cluster profiles around the median profile is a factor of a few in the centre, and almost zero at radii equal or larger than  $R_{500}$ . The central electron densities of the median profiles of the MTNG740 clusters tend to be slightly smaller compared to the observed clusters. However, the most extreme clusters of MTNG740 more than cover the range of the observed profiles. The central electron density of MTNG740 clusters increases systematically with cluster mass at fixed scaled



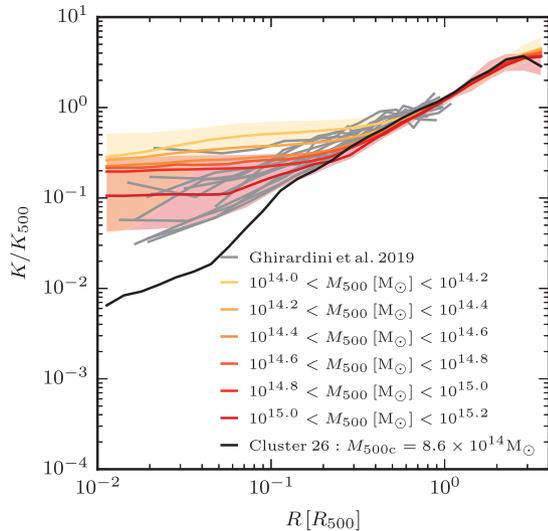
**Figure 7.** Median pressure profiles of galaxy clusters with  $M_{500c} > 10^{14} M_{\odot}$  at  $z = 0$  in bins of similar mass (top panel). The bands show 25 per cent and 75 per cent percentiles for the lowest and highest mass bins. We show the profiles of observed galaxy clusters from the X-COP project in grey for comparison (Ghirardini et al. 2019). The bottom panel gives the MTNG740 cluster profiles divided by the universal pressure profile for galaxy clusters (Arnaud et al. 2010). The shape of the profiles of the MTNG740 clusters agrees well with the shape of the universal pressure profile up to  $R_{500}$ . For massive clusters in MTNG740, the normalization of the profile is  $\sim 20$  per cent lower at  $R_{500}$  than predicted by the universal pressure profile.

radius, but reaches the same value independently of cluster mass in the outer parts, for  $R \gtrsim 0.4R_{500}$ , with very small scatter.

We consider pressure profiles in the top panel of Fig. 7, relative to  $P_{500}$  defined as

$$P_{500} = 1.65 \times 10^{-3} h(z)^{8/3} \left[ \frac{M_{500}}{3 \times 10^{14} h_70^{-1} M_{\odot}} \right]^{2/3} h_70^2 \text{ keV cm}^{-3}, \quad (3)$$

where  $h_70 = H_0/(70 \text{ km s}^{-1} \text{ Mpc}^{-1})$  and  $h(z) = H(z)/H_0$ , i.e.  $h(z = 0) = 1$  (Nagai, Kravtsov & Vikhlinin 2007). We compute the median pressure profile in the same way as the profile of the electron density, and again compare to the clusters of the X-COP survey (Ghirardini et al. 2019). We find good agreement between the profiles of the MTNG740 clusters and the observed clusters. There is no systematic



**Figure 8.** Median entropy profiles of galaxy clusters with  $M_{500} > 10^{14} M_{\odot}$  at  $z = 0$  in bins of similar total mass. The bands show 10 per cent and 90 per cent percentiles for the lowest and highest mass bins. The black line shows the profile of one extreme individual cool-core cluster with  $K < 30 \text{ keV cm}^2$  for  $R/R_{500} < 0.05$ .

We compare to the entropy profiles of observed galaxy clusters from the X-COP project in grey (Ghirardini et al. 2019). We find satisfactory agreement between the MTNG740 clusters and the observed profiles. More massive clusters in MTNG740 have, however, systematically lower entropy at fixed relative radius.

trend with cluster mass after expressing the pressure in terms of  $P_{500}$ . The scatter in the profiles of the MTNG740 clusters around  $R_{500}$  is slightly smaller than in the observations, but the latter is likely dominated by observational systematics.

In the bottom panel of Fig. 7, we divide the pressure profile by the universal pressure profile (Arnaud et al. 2010) that is often assumed when converting cluster observables to physical properties of the cluster. The electron pressure profiles of the MTNG740 clusters are well described by the universal pressure profile out to  $R_{500}$ , with deviations smaller than 20 per cent. However, at larger radii the universal pressure profile systematically underestimates the pressure in the MTNG740 clusters.

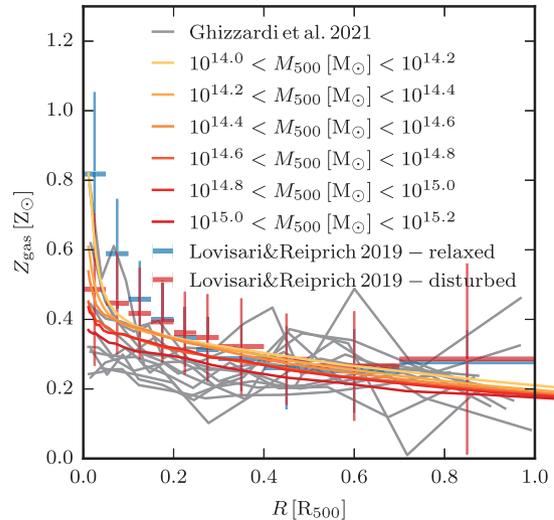
We consider the entropy profile in Fig. 8. For the MTNG740 clusters, we define the entropy as  $K = k_B T / n_e^{2/3}$  (Henden et al. 2018) and express it in units of the characteristic entropy  $K_{500} = k_B T_{500} / n_{e,500}^{2/3}$ , using

$$T_{500} = \frac{\mu m_p G M_{500}}{2 k_B R_{500}} \quad (4)$$

and

$$n_{e,500} = \frac{500 f_b \rho_c}{\mu_e m_p}, \quad (5)$$

where  $k_B$  is the Boltzmann constant,  $m_p$  is the mass of the proton,  $f_b = \Omega_b / \Omega_M$  is the baryon fraction of the Universe, and  $\rho_c$  is the critical density of the Universe at  $z = 0$ . We again compare to the profiles from the X-COP survey. We find that the median central entropy is significantly lower for more massive clusters in MTNG740. Overall, the entropy profiles of the MTNG740 clusters agree reasonably well with the observed profiles, but the median central entropy of the simulated clusters is systematically higher than the observed central entropy profiles. This can partly be understood as reflecting the differences in the central electron density profile



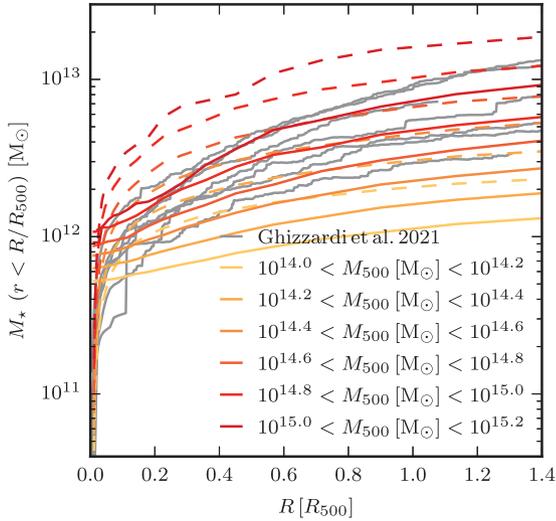
**Figure 9.** Median metallicity profiles of galaxy clusters with  $M_{500} > 10^{14} M_{\odot}$  at  $z = 0$  in bins of similar mass for a solar metallicity of  $Z_{\odot} = 0.0127$ . For comparison, we show observed profiles of two sets of nearby galaxy clusters (Lovisari & Reiprich 2019; Ghizzardi et al. 2021). The metallicity profiles of the MTNG740 galaxy clusters are in broad agreement with observed metallicity profiles of galaxy clusters.

that is lower for the MTNG740 clusters compared to the observed clusters. The kinetic AGN feedback in TNG quickly dissipates and increases the central entropy on the sound crossing time-scale. As a consequence, the cluster population can mostly be described by non-cool-core clusters (with central entropies  $> 30 \text{ keV cm}^2$ ), even though a small number of cool-core clusters exists in MTNG740. Nevertheless, MTNG740 misses out on a large fraction of cool-core clusters (as shown for the TNG300 sample by Barnes et al. 2018). Fitting an entropy profile of the form

$$K(r) = K_0 + K_{100} \left( \frac{r}{100 \text{ kpc}} \right)^a \quad (6)$$

in the range  $10^{-2} R_{500} - R_{500}$  (Cavagnolo et al. 2009; Barnes et al. 2018) we find that 196 or 8 per cent of all clusters at  $z = 0$  with  $M_{500} > 10^{14} M_{\odot}$  have  $K_0 < 30 \text{ keV cm}^2$  and can be classified as cool-core clusters, consistent with TNG300 (Barnes et al. 2018). None of the nine clusters with  $M_{500} > 10^{15} M_{\odot}$  fulfil this criterion. The lack of cool-core clusters is similar to purely thermal AGN feedback models (Altamura et al. 2022), in contrast to light AGN jet feedback that is able to maintain the cool core while self-regulating the cooling ICM (Ehlert et al. 2022; Weinberger et al. 2022).

In Fig. 9, we look at the metallicity profiles of the intracluster gas. For the MTNG740 clusters we first compute the mass-weighted average metallicity in spherical shells for each cluster, then combine the profiles to obtain median profiles in six mass bins. We compare to the observed metallicity profiles by Lovisari & Reiprich (2019) and Ghizzardi et al. (2021). We note that the observations typically measure iron abundance as a proxy for the metallicity of the gas in galaxy clusters. However, we only track the total metal abundance in MTNG740 (forced by memory constraints), and thus use it for the comparison. The median metallicity profiles of MTNG740 clusters for different mass bins are very similar. They show only a small but still systematic trend with the mass of the clusters. Specifically, more massive galaxy clusters have slightly lower gas metallicity at fixed relative radius. This is consistent with basic expectations because the fraction of baryons that is converted to stars and eventually produces



**Figure 10.** Median cumulative stellar mass profiles of galaxy clusters with  $M_{500c} > 10^{14} M_{\odot}$  at  $z = 0$  in bins of similar mass. The solid lines show the stellar mass within a certain radius in the BCG and satellite galaxies, and the dashed lines show the total stellar mass within a certain radius including intracluster light. The bands show 10 per cent and 90 per cent percentiles. For comparison we show observed profiles of nearby galaxy clusters from the X-COP project computed from all galaxies in the clusters (Ghizzardi et al. 2021). The BCG- and satellite-galaxy-based stellar mass profiles are in good agreement with the observed profiles generated in the same way. The total stellar mass profiles of the MTNG740 clusters are significantly higher owing to the a significant contribution of diffuse intracluster light.

metals decreases with halo mass (see also the top right-hand panel of Fig. 2). We find overall good agreement with the observed metallicity profiles. The inner slope of the observed clusters, most notably for the sample of relaxed clusters, seems to be slightly steeper than for the MTNG740 clusters, though a similar trend is not obvious in the X-COP sample. We leave a more detailed analysis that splits the MTNG740 galaxy clusters into relaxed and disturbed clusters in a similar way as done for observed clusters to future work.

Finally, we show cumulative stellar mass profiles in Fig. 10 and again compare to the observed profiles of the clusters of the X-COP project (Ghizzardi et al. 2021). The stellar mass profiles of the MTNG740 clusters computed from their BCG and satellite galaxies look consistent with the stellar mass profiles of the observed clusters that are generated in a similar way. The observed stellar mass profiles are slightly steeper at large radii.

The total stellar mass profiles of the MTNG740 clusters, however, are significantly higher showing that there is a significant amount of stellar mass in a diffuse intracluster light component. This component is missed when only galaxies in the cluster are counted. Interestingly, the total  $K$ -band luminosity of the MTNG740 clusters seems to agree well with the scaling relations shown in Section 4 even though the total  $K$ -band luminosity of the observed clusters was also calculated from cluster galaxies. We leave this apparent discrepancy for future work.

We find that the median cumulative stellar mass profiles are almost self-similar for different mass bins, with a normalization that increases with cluster mass. There is reasonable agreement of the outer shape of the stellar mass profiles between MTNG740 and observations. However, the inner stellar mass profiles are significantly steeper for the MTNG740 clusters, and their stellar mass is more centrally concentrated than for the observed profiles. The total stellar mass in the cluster is larger in MTNG740 clusters, but the

stellar mass in galaxies in MTNG740 clusters is lower than observed. This difference indicates that galaxies are stripped more quickly in MTNG740 than in the real Universe, which can increase the intracluster light contribution in MTNG740 at the expense of the individual galaxies. However, Ahad et al. (2021) argue that, if anything, galaxies are stripped too little in the Hydrangea simulations, which have a roughly similar mass resolution as MTNG740. We compare the properties of the cluster galaxies in MTNG740 with observations in Section 7. Nevertheless, as shown in previous simulations, the intracluster light can contribute a substantial fraction of the stellar mass outside the central BCG (Puchwein et al. 2010). A similar result has later been found in observations with sufficiently deep surface brightness sensitivity (Presotto et al. 2014; de Oliveira, Jiménez-Teja & Dupke 2022). Future work should mimic the stellar light of the simulated MTNG740 clusters and apply the same observational cuts and limitations to facilitate a proper comparison.

## 6 THE SZ-SIGNAL ON THE LIGHT-CONE

While spherical profiles and integrated quantities on spherical apertures greatly assist physical understanding, these measures cannot be observed directly. Instead, we can only access projections of an object on the sky. More strictly speaking, we only see our past backwards light-cone. To understand the relevance of this difference for SZ observations, we first consider the Compton- $y$  parameter for a full, deep light-cone output of our simulation in Fig. 11. This light-cone extends to redshift  $z = 5$  and covers a square area of  $10^{\circ} \times 10^{\circ}$  on the sky.

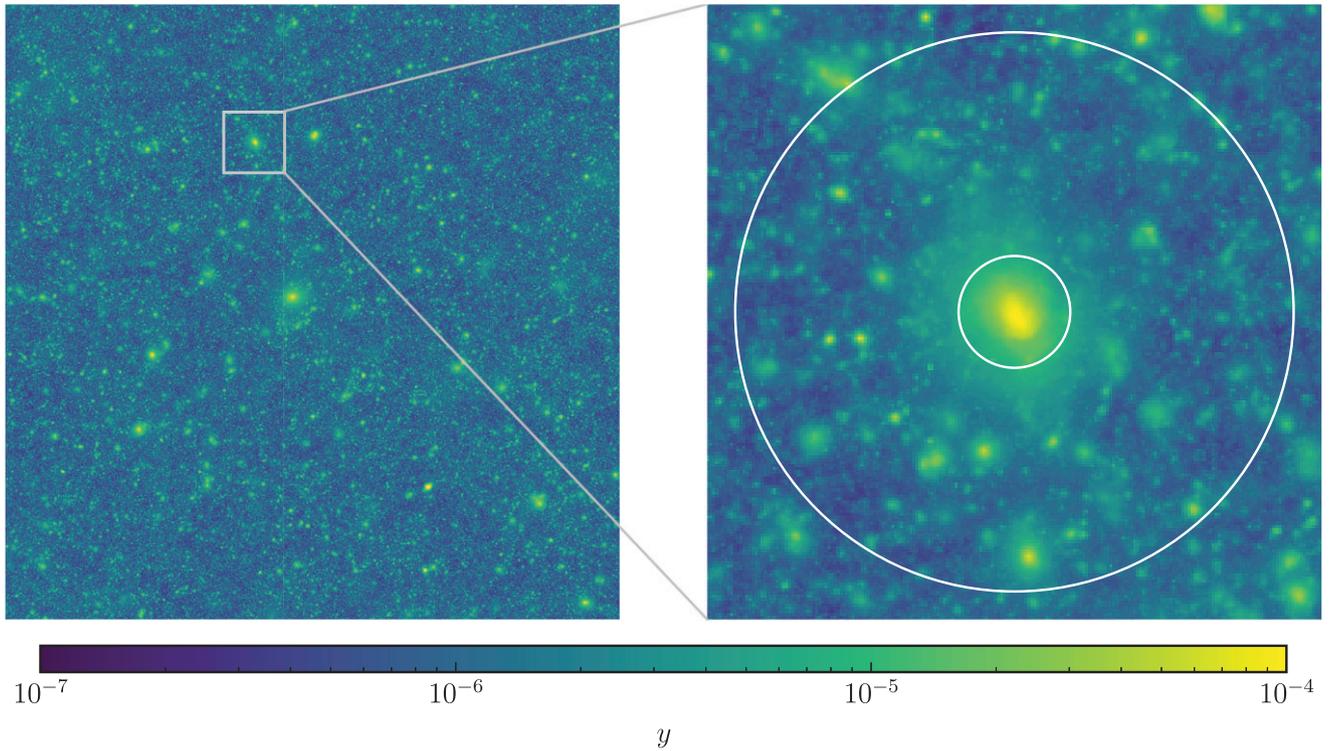
We compute the Compton- $y$  of a pixel on the light-cone in a similar way as  $Y_{500}$  in equation (1) from the total thermal energy of the electrons in the volume covered by pixel on the sky as

$$y(\theta) = \frac{k_B \sigma_T}{m_e c^2} \int_0^D n_e T_e dl = \frac{(\gamma - 1) \sigma_T}{m_e c^2} \bar{x}_e X \mu \frac{1}{\Omega} \sum_i \frac{E_i}{D_{\Lambda,i}^2}, \quad (7)$$

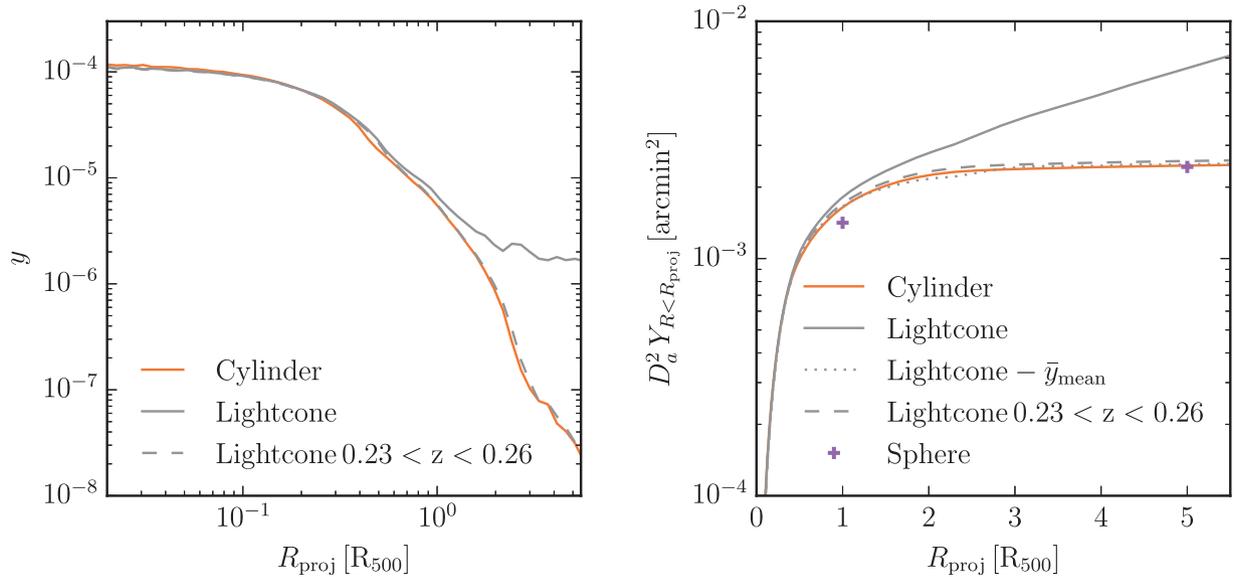
where  $E_i$  is the thermal energy of a cell and  $D_{\Lambda,i}$  the angular diameter distance at the redshift when the cell crosses the light-cone,  $\Omega$  is the solid angle of the pixel on the sky, and  $\Sigma_i$  includes all light-cone cells whose centres lie on the pixel.

Massive galaxy clusters clearly stick out with a high central Compton- $y$  up to  $10^{-4}$ . We find a mean background value on the full light-cone map of  $\bar{y} = 1.68 \times 10^{-6}$  (mean) and  $\bar{y} = 1.12 \times 10^{-6}$  (median). This background level is still consistent with the upper bound from Planck after removing galaxy clusters with  $\bar{y} < 2.2 \times 10^{-6}$  (Khatri & Sunyaev 2015). Note, however, that the latter bound was computed for a slightly different cosmology. Our background level is also consistent with previous estimates from Springel, White & Hernquist (2001), the Magneticum simulation ( $\bar{y} < 1.18 \times 10^{-6}$ , Dolag et al. 2016), and with estimates from semi-analytical models on top of dark-matter-only simulations (Osato et al. 2018).

In the right panel of Fig. 11, we zoom in on the most massive galaxy cluster at  $z = 0.25$  in the deep light-cone map. This galaxy cluster has a mass of  $M_{500} = 7 \times 10^{14} M_{\odot}$  at  $z = 0.25$ , and the white circles in Fig. 11 denote its  $R_{500}$  and  $5R_{500}$  radii. These are typical aperture radii used to estimate the integrated Compton- $y$  parameter of galaxy clusters. Interestingly, we see a great deal of substructure projected on to the cluster, in particular between  $R_{500}$  and  $5R_{500}$ , though most of this substructure is in the foreground and background and not in the immediate vicinity of the cluster, as we will show in the following text.



**Figure 11.** Map of the Compton- $y$  parameter for a deep  $10^\circ \times 10^\circ$  light-cone that extends to  $z = 5$ . The right panel shows a zoom-in on the largest galaxy cluster with mass  $M_{500} = 7 \times 10^{14} M_\odot$  at  $z = 0.25$  contained on this light-cone. The white circles show  $R_{500}$  and  $5 R_{500}$  of this cluster.



**Figure 12.** Compton- $y$  profile of the cluster shown on the right of Fig. 11 at  $z = 0.25$ . The left panel of the present figure shows the radially averaged profile computed from the full light-cone (solid grey line) as well as the profile computed only from the cells contributing to the light-cone between  $z = 0.23$  and  $z = 0.26$  (dashed grey line). The orange line gives the same profile but computed from a cylinder with a depth of 100 Mpc from the snapshot data at  $z = 0.25$ . The right panel plots the integrated Compton- $y$  parameter computed from the light-cone and cylinder projections shown in the left panel, and compares them to the integrated Compton- $y$  parameter computed for a spherical aperture at  $R_{500}$  and  $5 R_{500}$ . Additionally, we show the integrated Compton- $y$  parameter from the full light-cone after subtracting the mean Compton- $y$  from the full light-cone shown in Fig. 11. Background subtraction is crucial to reconstruct the integrated Compton- $y$  of the cluster.

To examine this point further, we consider the differences between measuring the Compton- $y$  on the deep light-cone and from a local aperture that considers the cluster and its large-scale environment extending over 100 Mpc. We compute the projected Compton- $y$

profile of the cluster on the light-cone as shown in the left-hand panel of Fig. 12. This is compared to the projected Compton- $y$  profile of the same cluster when seen in a cylindrical projection of the snapshot time-slice at  $z = 0.25$ , using a total depth of  $\sim 100$  Mpc around the

cluster centre and the same viewing direction as for the light-cone observer. Both resulting radial profiles are contrasted in the left panel of Fig. 12. Additionally, we also include separately a projected profile of the light-cone, but with contributions only from gas between  $z = 0.23$  and  $z = 0.26$ , i.e. to local contributions around the redshift of the cluster with a depth similar to the cylindrical projection.

We find that the profiles computed from the cylindrical aperture or from the local light-cone contributions are essentially identical. In contrast, the radial profile of the full light-cone starts to significantly deviate at a projected radius of  $\sim R_{500}$ , when the Compton- $y$  approaches the mean background level of the light-cone.

To quantify this difference better we show the integrated Compton- $y$  parameter for the galaxy cluster out to a given radius in the right panel of Fig. 12. We see that  $Y_{500, \text{Cylinder}}$  is 10 per cent smaller than  $Y_{500, \text{Light-cone}}$ , while  $Y_{500, \text{Sphere}}$  that we used in Fig. 5 is 20 per cent smaller than  $Y_{500, \text{Light-cone}}$ .

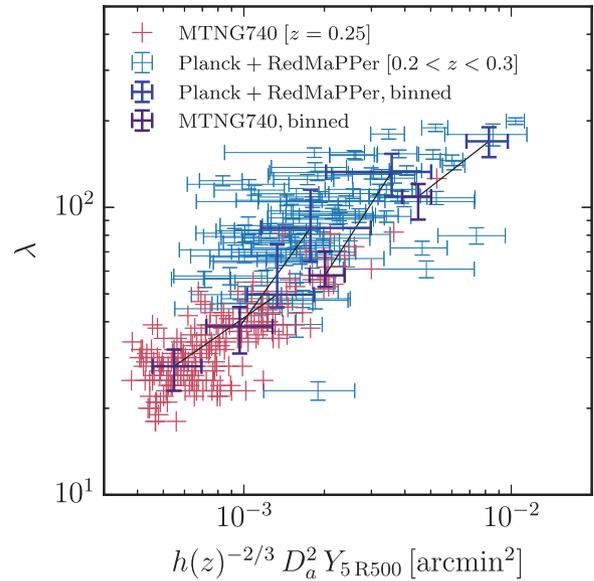
For a radius of  $5 R_{500}$  as used by Planck (Planck Collaboration XXVII 2016b) the situation is different.  $Y_{5R500}$  is essentially identical for the spherical and cylindrical apertures. Thus, the local background that is included in the cylindrical aperture but not the spherical aperture does not significantly contribute to the total signal. However,  $Y_{5R500}$  measured on the light-cone is 2.5 times larger than measured for either local aperture, as the distant background that can be seen in Fig. 11 contributes very significantly in this case. Hence, a careful background removal as done via matched or scale-adaptive filtering (Tegmark & de Oliveira-Costa 1998; Sanz, Herranz & Martínez-González 2001; Herranz et al. 2002a, b; Schäfer et al. 2006; Kay et al. 2012) is crucial. Future studies can use MTNG740 to test the full procedure used by SZ surveys and to identify potential systematics still present in these procedures.

## 7 GALAXY CLUSTER OBSERVABLES

For all comparisons of galaxy cluster properties with observations discussed so far we have assumed that we precisely know the mass of the observed galaxy clusters. However, mass is unfortunately not a direct observable for galaxy clusters, so mass estimates always require additional assumptions about the galaxy clusters that all introduce significant systematic uncertainties.

We can avoid these assumptions by comparing to observations fully in observational space. Here, we present a first example of this approach with MTNG740 by comparing two directly observable quantities of the MTNG740 clusters with corresponding measurements made for observed galaxy clusters. For definiteness, we chose the richness and the Compton- $y$  parameter. Both quantities are almost completely independent probes of the state and properties of a galaxy cluster. Here, we do not attempt to construct fully synthetic observations that are analysed in the same way as observational data, but we instead compare the theoretical quantities  $Y_{5R500}$ , i.e. the integrated Compton- $y$  parameter in a sphere with a radius  $5 R_{500}$ , and the richness  $\lambda$ . Both quantities are well defined and easily determinable for the MTNG740 clusters, but they still require some modest assumptions to be made when measuring them from observations. Specifically, we compare the MTNG740 clusters to a sample of galaxy clusters that have both SZ measurements from Planck (Planck Collaboration XXVII 2016b; Planck Collaboration 2017), and a richness measurement from the RedMaPPer catalogue based on the SDSS-8 galaxy catalogue (Rykoff et al. 2014, 2016).

We compare the MTNG740 clusters to the observed sample in Fig. 13 for all MTNG740 clusters at  $z = 0.25$  with minimum mass of  $M_{500} > 10^{14.4} M_{\odot}$ , roughly matching the detection limit of galaxy clusters in Planck at this redshift. We compare to the sample of



**Figure 13.** Richness versus Compton- $y$ . The light blue crosses show galaxy clusters in MTNG740 in the  $z = 0.25$  snapshot with  $M_{500} > 10^{14.4} M_{\odot}$ . The orange data points give all galaxy clusters between  $z = 0.2$  and  $z = 0.3$  that are both in the Planck SZ catalogue (Planck Collaboration XXVII 2016b, 2017) and the RedMaPPer SDSS-8 catalogue (Rykoff et al. 2014). We also show median values and error bars for 16 per cent and 84 per cent percentiles in four logarithmic mass bins for MTNG740 clusters (blue) and observed clusters (brown). The black lines connect points of clusters of the same mass. The individual MTNG740 clusters lie at the lower end within the parameter space of observed clusters. The binned distributions seem however to be consistent, except for a systematically lower cluster mass in MTNG740 at fixed richness and Compton- $y$ . The scatter between MTNG740 clusters is smaller than the scatter between observed clusters.

clusters that are detected in Planck and that are part of the RedMaPPer SDSS8 catalogue in the redshift range  $z = 0.2$  and  $z = 0.3$ . We compute the richness and integrated Compton- $y$  parameter or the MTNG740 clusters as described in Section 4.

The MTNG740 clusters fall well within the parameter space covered by the observed galaxy clusters. The scatter at fixed richness or fixed Compton- $y$  parameter seems to be larger for the observed clusters than for the MTNG740 clusters. This difference is most easily explained by observational systematics that contribute to the scatter. The MTNG740 clusters lie at the low-richness end of the distribution of observed clusters. However, this offset is much smaller than what we found for the scaling relations (see Fig. 5), which may indicate that the weak lensing mass estimates that we used for the observed clusters could be biased and be responsible for at least parts of the discrepancy.

Importantly, MTNG740 lacks galaxy clusters with the largest values of richness and Compton- $y$  found in the Planck + RedMaPPer sample. A simple explanation for this discrepancy could be that the most massive MTNG740 clusters are significantly less massive than the clusters in the observed sample due to the still limited volume of our simulation. This explanation is consistent with a rough estimate of the observational volume of SDSS-8 that covers roughly 25 per cent of the sky, which is equivalent to volume about four times larger between  $z = 0.2$  and  $z = 0.3$  than the MTNG740 volume.

To further test this hypothesis, we group both sets of galaxy clusters into four logarithmic mass bins from  $10^{14.4} M_{\odot}$  to  $10^{15.2} M_{\odot}$  with a width of 0.2 dex each. We find object counts of the Planck + RedMaPPer sample of 9, 60, 14, and 2, whereas the

corresponding counts for the MTNG740 clusters are 191, 72, 19, and 3, for bins of increasing cluster mass. Surprisingly, we find a slightly larger number of objects in the MTNG740 cluster sample at the high-mass end. However, the richness and Compton- $y$  values for MTNG740 clusters are significantly smaller than the values found for Planck + RedMaPPer clusters at a similar estimated mass.

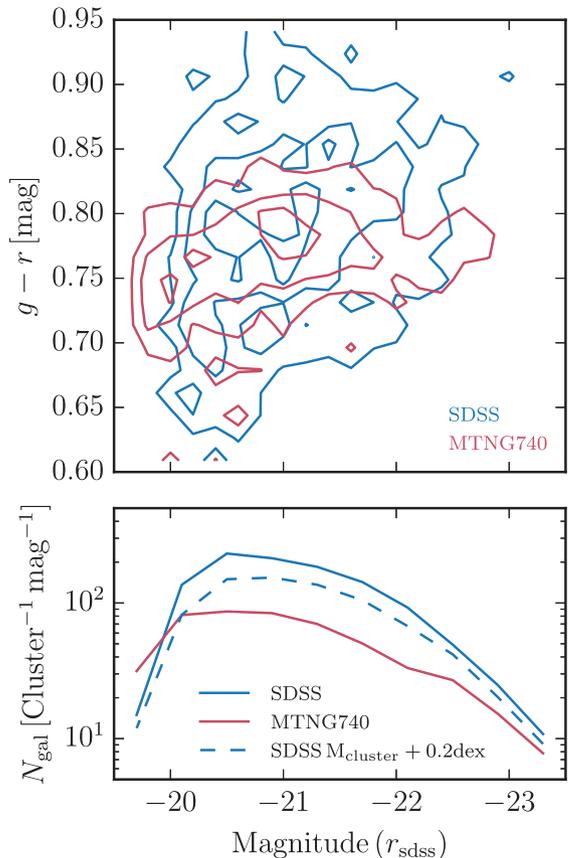
There are two obvious ways to explain the discrepancies. First, if we assume that the SZ-based mass estimate from the Planck SZ catalogue is underestimating the true mass by  $\sim 0.2$  dex we find almost perfect agreement between both samples. In this case the richness, the integrated Compton- $y$  values, as well as the number of clusters in the highest mass bins are consistent within statistical scatter between both sets of clusters after taking into account the differences in probed volumes.

Remarkably, this factor of  $\sim 0.2$  dex to correct the mass estimates of the Planck clusters is very similar to the empirical correction to the SZ mass estimates applied in the ACT SZ catalogue (Hilton et al. 2021). Their correction is based on the observed richness of a subset of the ACT clusters, and a relation between richness and weak lensing mass estimates (McClintock et al. 2019).

One problem with this interpretation is that the MTNG740 clusters do not lie on the richness scaling relation shown in Section 4. This is not completely unexpected as MTNG740 galaxies around and below the knee of the stellar mass function are not massive enough as shown in Fig. 2, even though it is not clear how much this changes the richness estimate. However, it may also indicate more complicated interpretations of the richness estimate of the MTNG740 cluster for different brightness limits of cluster galaxies or possibly an observational bias in the richness estimate of the LoCuSS sample (Mulroy et al. 2019).

As a consistency check of the richness of the MTNG740 clusters, we select all galaxies of all clusters with a mass larger than  $10^{14.6} M_{\odot}$  that contribute to the richness estimate of Fig. 13 and show their distribution in absolute rest-frame  $r$ -band magnitude versus rest-frame SDSS  $g - r$  colour space in the top panel of Fig. 14. For the Planck + RedMaPPer sample, we include all SDSS galaxies with a non-zero cluster membership probability and weight them with their individual probability. Moreover, we weight every galaxy with the inverse of the number of clusters in the cluster mass bin of its parent cluster. Here, we use the same four mass bins as in Fig. 13. Note that the SDSS galaxies are de-reddened for Milky Way dust, but not for intrinsic dust. We apply  $K$ -corrections (Chilingarian, Melchior & Zolotukhin 2010) to obtain rest-frame  $r$ -band magnitudes of the observed galaxies and use the redshift of their parent galaxy cluster to transform them to absolute magnitudes. This yields good agreement in colour and brightness between the MTNG740 cluster galaxies and cluster galaxies of the Planck + RedMaPPer sample, only the scatter in colour is smaller for the MTNG740 cluster galaxies. In the bottom panel of Fig. 14 we show the average number of galaxies per cluster and decade of magnitude for MTNG740 and the Planck + RedMaPPer sample. They are similar at the bright end, but there are progressively fewer galaxies in MTNG740 clusters compared to the Planck clusters the fainter the galaxies become. The lower panel of Fig. 14 also shows that this difference becomes smaller but does not vanish (dashed line) when we assume that the Planck clusters are 0.2 dex more massive, which changes which clusters are matched to the MTNG740 clusters. While this is mostly expected from Fig. 13, it could still be explained by statistical variance.

If we alternatively assume that the SZ-based mass estimates of the Planck clusters are correct, we need to conclude that the MTNG740 clusters systematically underestimate the richness as well as the



**Figure 14.** Galaxy distributions in rest-frame  $g - r$  colour versus apparent rest-frame  $r$ -band magnitude (upper panel). We use all galaxies with an apparent  $i$ -band magnitude  $m_i < 21$  at  $z = 0.25$  within  $R_{500}$  of the MTNG740 clusters shown in Fig. 13 that have a mass larger than  $10^{14.6} M_{\odot}$ . We compare to all galaxies that have a membership probability larger than zero for one of the clusters of the Planck + RedMaPPer sample used in Fig. 13. Observed clusters are weighted with their membership probability. The colour and brightness distributions of both galaxy populations are consistent, with slightly larger scatter in colour in the observed galaxies. Note that no dust was included to compute the MTNG740 galaxy magnitudes. The lower panel shows the number of galaxies per cluster and magnitude in the  $i$  band for both samples. The dashed line shows the observed galaxies from all Planck clusters with the same mass cut, but assuming that their masses are 0.2 dex larger. There is a clear difference between MTNG740 and the Planck cluster galaxies that increases towards fainter galaxies. It is reduced but still there if the Planck clusters are assumed to be more massive.

integrated Compton- $y$ . The former is consistent with the richness scaling relation, but the latter would break the SZ scaling relation as shown in Fig. 5. Because Planck measures  $Y_{5R_{500}}$  rather than  $Y_{500}$  as shown in the SZ scaling relation, this would require significantly stronger AGN-driven outflows that heat the gas around massive galaxy clusters on scales well beyond  $R_{500}$  without changing the properties of the gas within  $R_{500}$ . Moreover, the Planck + RedMaPPer sample would need to lack a significant number of the most massive clusters expected to be found in the survey volume.

We conclude that comparisons between simulations and observations in observational space are a crucial and promising test of simulations, and they may also help to understand and interpret observational systematics. We thus envision to carry out a full forward modelling of mock light-cone data, and a detailed comparison of the SZ signal of MTNG740 clusters on the light-cone with observations, in future work.

## 8 SUMMARY AND OUTLOOK

In this introductory paper of the MillenniumTNG project, we have focused on the flagship full physics simulation that evolves a  $500 h^{-1}$  Mpc (740 Mpc) cosmological box to  $z = 0$  with a baryonic mass resolution of  $3.1 \times 10^7 M_{\odot}$  and a physics model close to the IllustrisTNG model. In Section 3, we established that the MTNG740 hydrodynamical simulation is consistent with its IllustrisTNG predecessors as well as with recent observational data (see Fig. 2). In particular, the colour bimodality of galaxies is still present despite the slightly reduced resolution compared to TNG300. The transition from star-forming blue galaxies to quenched red galaxies occurs at a slightly higher stellar mass in MTNG740 than in TNG (see Fig. 3). The baryonic impact on the total matter power spectrum in MTNG740 is essentially identical to TNG100 as shown in Fig. 4.

We then presented a first analysis of the integrated properties of galaxy clusters in MTNG740. In Section 4, we discussed different galaxy cluster scaling relations in MTNG740 and compared them to the nearby galaxy cluster sample by Mulroy et al. (2019). We found generally good agreement between MTNG740 and observed clusters. The normalization of the richness of the observed galaxy clusters seems to be about a factor of 2 higher than the richness of the MTNG740 at fixed mass. However, an important caveat is that the mass estimates of the observed clusters are based on weak gravitational lensing mass estimates that might be biased.

We then looked at the internal structure of the MTNG740 clusters in Section 5. We examined internal profiles of hydrodynamical quantities as well as metallicity and stellar mass of the MTNG740 galaxy clusters and compared them to observed profiles of nearby well-observed galaxy clusters from the X-COP survey (Eckert et al. 2019; Ettori et al. 2019; Ghirardini et al. 2019). We again found satisfactory agreement between the profiles of the MTNG740 clusters and observations. We found the main discrepancy in the inner part of the hydrodynamical profiles of clusters that are able to reproduce non-cool-core clusters, but not cool-core clusters, perhaps requiring a more physical AGN feedback model that reproduces both classes of clusters as observed. A possible refinement to the TNG galaxy model could be to include large-scale jets for the most massive systems.

In Section 6, we considered the Compton- $y$  map of a deep light-cone up to redshift  $z = 5$ . The average background value of the Compton- $y$  parameter we find is consistent with the latest constraints from Planck. We then compared the Compton- $y$  of a galaxy cluster at  $z = 0.25$  measured in different apertures. We showed that neither spherical nor cylindrical apertures fully describe the signal observed on the light-cone. In particular, for large apertures ( $5 R_{500}$ ) the Compton- $y$  background significantly contributes so that it cannot be ignored and requires matched filtering to derive an unbiased estimate of the cluster-intrinsic Compton- $y$  signal.

Having shown that the MTNG740 galaxy cluster population is overall consistent with observed galaxy clusters but shows some interesting discrepancies, we turned in Section 7 to a comparison that stays fully in observational space and sidesteps the thorny issue of cluster mass estimates. We compared the galaxy clusters in MTNG740 with a set of galaxy clusters found in both, the Planck SZ cluster catalogue, and the RedMaPPer SDSS-8 catalogue. We compared richness and Compton- $y$ , which are both observables that can also be directly extracted from the simulation. Overall, we found quite good agreement in Fig. 13. Moreover, we confirmed in Fig. 14 that the population of cluster galaxies contributing to the richness estimate is consistent between simulation and observations. However, at the same time this has highlighted that at similar richness and Compton- $y$  parameter, the SZ-based  $M_{500}$  mass estimates of the

Planck cluster catalogue are 0.2 dex higher than the masses measured for the MTNG740 clusters. We discussed possible explanations for this offset and remark that this discrepancy is completely consistent with the weak lensing and richness-based mass correction applied in more recent SZ cluster surveys (Hilton et al. 2021).

We conclude that MTNG740 offers a great opportunity to study and understand galaxies and galaxy clusters in their large-scale cosmological context. Also, the MTNG740 simulations can be used to interpret data from future large cosmological surveys, and to quantify and potentially correct for observational biases in cosmological measurements. We give further examples of such applications in our set of companion papers, and presently work on additional research in this direction in forthcoming studies.

## ACKNOWLEDGEMENTS

We thank the anonymous referee for helpful comments that improved the paper. RP thanks Wilma Trick for help with the colour scheme of the paper. The authors gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time on the GCS Supercomputer SuperMUC-NG at the Leibniz Supercomputing Centre (LRZ) in Garching, Germany, under project pn34mo. This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility, with equipment funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1, and ST/S002502/1, Durham University, and STFC operations grant ST/R000832/1. CH-A acknowledges support from the Excellence Cluster ORIGINS which is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2094-390783311. VS and LH acknowledge support by the Simons Collaboration on ‘Learning the Universe’. LH is supported by NSF grant AST-1815978. CP acknowledges support by the European Research Council under ERC-AdG grant PICO GAL-101019746. SB is supported by the UK Research and Innovation (UKRI) Future Leaders Fellowship (grant number MR/V023381/1).

## DATA AVAILABILITY

The MillenniumTNG simulations will be made fully publicly available at <https://www.mtng-project.org/> in 2024. The data shown in the figures of this article will be shared upon reasonable request to the corresponding author.

## REFERENCES

- Abdurro'uf et al., 2022, *ApJS*, 259, 35  
 Ahad S. L., Bahé Y. M., Hoekstra H., van der Burg R. F. J., Muzzin A., 2021, *MNRAS*, 504, 1999  
 Altamura E., Kay S. T., Bower R. G., Schaller M., Bahé Y. M., Schaye J., Borrow J., Towler I., 2022, preprint ([arXiv:2210.09978](https://arxiv.org/abs/2210.09978))  
 Angulo R. E., Hahn O., 2022, *Living Rev. Comput. Astrophys.*, 8, 1  
 Angulo R. E., Pontzen A., 2016, *MNRAS*, 462, L1  
 Ardila F. et al., 2021, *MNRAS*, 500, 432  
 Arnaud M., Pratt G. W., Piffaretti R., Böhringer H., Croston J. H., Pointecouteau E., 2010, *A&A*, 517, A92  
 Bahé Y. M. et al., 2017, *MNRAS*, 470, 4186  
 Barnes D. J., Kay S. T., Henson M. A., McCarthy I. G., Schaye J., Jenkins A., 2017, *MNRAS*, 465, 213  
 Barnes D. J. et al., 2018, *MNRAS*, 481, 1809  
 Barrera M. et al., 2022, preprint ([arXiv:2210.10419](https://arxiv.org/abs/2210.10419))  
 Battaglia N., Bond J. R., Pfrommer C., Sievers J. L., 2012, *ApJ*, 758, 74

- Bird S., Ni Y., Di Matteo T., Croft R., Feng Y., Chen N., 2022, *MNRAS*, 512, 3703
- Borrow J., Vogelsberger M., O'Neil S., McDonald M. A., Smith A., 2022, preprint (arXiv:2205.10376)
- Bose S. et al., 2022, preprint (arXiv:2210.10065)
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Bundy K. et al., 2015, *ApJ*, 798, 7
- Cavagnolo K. W., Donahue M., Voit G. M., Sun M., 2009, *ApJS*, 182, 12
- Chabrier G., 2003, *PASP*, 115, 763
- Chilingarian I. V., Melchior A.-L., Zolotukhin I. Y., 2010, *MNRAS*, 405, 1409
- Contreras S. et al., 2022, preprint (arXiv:2210.10075)
- Cui W. et al., 2018, *MNRAS*, 480, 2898
- Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, *MNRAS*, 486, 2827
- de Oliveira N. O. L., Jiménez-Teja Y., Dupke R., 2022, *MNRAS*, 512, 1916
- Debackere S. N. B., Hoekstra H., Schaye J., 2022, *MNRAS*, 515, 6023
- Dolag K., Komatsu E., Sunyaev R., 2016, *MNRAS*, 463, 1797
- Driver S. P. et al., 2022, *MNRAS*, 513, 439
- Dubois Y., Peirani S., Pichon C., Devriendt J., Gavazzi R., Welker C., Volonteri M., 2016, *MNRAS*, 463, 3948
- Dubois Y. et al., 2021, *A&A*, 651, A109
- Eckert D. et al., 2016, *A&A*, 592, A12
- Eckert D. et al., 2019, *A&A*, 621, A40
- Ehlert K., Weinberger R., Pfrommer C., Pakmor R., Springel V., 2022, *MNRAS*, 518, 4622
- Ettori S. et al., 2019, *A&A*, 621, A39
- Ghirardini V. et al., 2019, *A&A*, 621, A41
- Ghizzardi S. et al., 2021, *A&A*, 646, A92
- Girelli G., Pozzetti L., Bolzonella M., Giocoli C., Marulli F., Baldi M., 2020, *A&A*, 634, A135
- Green S. B., van den Bosch F. C., Jiang F., 2021, *MNRAS*, 503, 4075
- Gruppioni C. et al., 2020, *A&A*, 643, A8
- Hadzhiyska B. et al., 2022a, preprint (arXiv:2210.10068)
- Hadzhiyska B. et al., 2022b, preprint (arXiv:2210.10072)
- Henden N. A., Puchwein E., Shen S., Sijacki D., 2018, *MNRAS*, 479, 5385
- Hernández-Aguayo C. et al., 2022, preprint (arXiv:2210.10059)
- Herranz D., Sanz J. L., Hobson M. P., Barreiro R. B., Diego J. M., Martínez-González E., Lasenby A. N., 2002a, *MNRAS*, 336, 1057
- Herranz D., Sanz J. L., Barreiro R. B., Martínez-González E., 2002b, *ApJ*, 580, 610
- Hilton M. et al., 2021, *ApJS*, 253, 3
- Johnson H. L., Mitchell R. I., Iriarte B., Wisniewski W. Z., 1966, *Commun. Lunar Planet. Lab.*, 4, 99
- Kannan R., Springel V., Pakmor R., Marinacci F., Vogelsberger M., 2016, *MNRAS*, 458, 410
- Kannan R. et al., 2022a, preprint (arXiv:2210.10066)
- Kannan R., Garaldi E., Smith A., Pakmor R., Springel V., Vogelsberger M., Hernquist L., 2022b, *MNRAS*, 511, 4005
- Kay S. T., Peel M. W., Short C. J., Thomas P. A., Young O. E., Battye R. A., Liddle A. R., Pearce F. R., 2012, *MNRAS*, 422, 1999
- Khatri R., Sunyaev R., 2015, *J. Cosmol. Astropart. Phys.*, 2015, 013
- Kormendy J., Ho L. C., 2013, *ARA&A*, 51, 511
- Lange R. et al., 2015, *MNRAS*, 447, 2603
- Lovisari L., Reiprich T. H., 2019, *MNRAS*, 483, 540
- Lovisari L., Reiprich T. H., Schellenberger G., 2015, *A&A*, 573, A118
- Marinacci F. et al., 2018, *MNRAS*, 480, 5113
- McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, *MNRAS*, 465, 2936
- McClintock T. et al., 2019, *MNRAS*, 482, 1352
- Montero-Dorta A. D., Prada F., 2009, *MNRAS*, 399, 1106
- Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, *ApJ*, 710, 903
- Mulroy S. L. et al., 2019, *MNRAS*, 484, 60
- Nagai D., Kravtsov A. V., Vikhlinin A., 2007, *ApJ*, 668, 1
- Naiman J. P. et al., 2018, *MNRAS*, 477, 1206
- Nelson D. et al., 2018, *MNRAS*, 475, 624
- Nelson D. et al., 2019a, *Comput. Astrophys. Cosmol.*, 6, 2
- Nelson D. et al., 2019b, *MNRAS*, 490, 3234
- Ni Y. et al., 2022, *MNRAS*, 513, 670
- Oesch P. A., Bouwens R. J., Illingworth G. D., Labbé I., Stefanon M., 2018, *ApJ*, 855, 105
- Osato K., Flender S., Nagai D., Shirasaki M., Yoshida N., 2018, *MNRAS*, 475, 532
- Pakmor R., Springel V., Bauer A., Mocz P., Munoz D. J., Ohlmann S. T., Schaal K., Zhu C., 2016, *MNRAS*, 455, 1134
- Pillepich A. et al., 2018a, *MNRAS*, 473, 4077
- Pillepich A. et al., 2018b, *MNRAS*, 475, 648
- Pillepich A. et al., 2019, *MNRAS*, 490, 3196
- Planck Collaboration XIII, 2016a, *A&A*, 594, A13
- Planck Collaboration XXVII, 2016b, *A&A*, 594, A27
- Planck Collaboration et al., 2017, *A&A*, 594, A27
- Pop A.-R. et al., 2022, preprint (arXiv:2205.11528)
- Potter D., Stadel J., Teyssier R., 2017, *Comput. Astrophys. Cosmol.*, 4, 2
- Presotto V. et al., 2014, *A&A*, 565, A126
- Puchwein E., Springel V., Sijacki D., Dolag K., 2010, *MNRAS*, 406, 936
- Ragone-Figueroa C., Granato G. L., Ferraro M. E., Murante G., Biffi V., Borgani S., Planelles S., Rasia E., 2018, *MNRAS*, 479, 1125
- Rowan-Robinson M. et al., 2016, *MNRAS*, 461, 1100
- Rykoff E. S. et al., 2014, *ApJ*, 785, 104
- Rykoff E. S. et al., 2016, *ApJS*, 224, 1
- Sánchez S. F. et al., 2016, *Rev. Mex. Astron. Astrofis.*, 52, 171
- Sanz J. L., Herranz D., Martínez-González E., 2001, *ApJ*, 552, 484
- Schäfer B. M., Pfrommer C., Bartelmann M., Springel V., Hernquist L., 2006, *MNRAS*, 370, 1309
- Schaye J. et al., 2015, *MNRAS*, 446, 521
- Springel V., 2010, *MNRAS*, 401, 791
- Springel V., Hernquist L., 2003, *MNRAS*, 339, 289
- Springel V., White M., Hernquist L., 2001, *ApJ*, 549, 681
- Springel V. et al., 2018, *MNRAS*, 475, 676
- Springel V., Pakmor R., Zier O., Reinecke M., 2021, *MNRAS*, 506, 2871
- Tegmark M., de Oliveira-Costa A., 1998, *ApJ*, 500, L83
- Torrey P., Vogelsberger M., Genel S., Sijacki D., Springel V., Hernquist L., 2014, *MNRAS*, 438, 1985
- van Daalen M. P., McCarthy I. G., Schaye J., 2020, *MNRAS*, 491, 2424
- Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, *MNRAS*, 436, 3031
- Vogelsberger M. et al., 2014a, *MNRAS*, 444, 1518
- Vogelsberger M. et al., 2014b, *Nature*, 509, 177
- Vogelsberger M., Marinacci F., Torrey P., Puchwein E., 2020, *Nat. Rev. Phys.*, 2, 42
- Weinberger R. et al., 2017, *MNRAS*, 465, 3291
- Weinberger R., Springel V., Pakmor R., 2020, *ApJS*, 248, 32
- Weinberger R. et al., 2022, preprint (arXiv:2211.11771)
- Wilkins S. M., Trentham N., Hopkins A. M., 2008, *MNRAS*, 385, 687
- Wright A. H. et al., 2017, *MNRAS*, 470, 283
- Yu H., Wang F. Y., 2016, *ApJ*, 820, 114
- Zhu P., Ho L. C., Gao H., 2021, *ApJ*, 907, 6

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